Roadmap of Mechanical Systems Design for the Semiconductor Automatic Test Equipment Industry

by

Michael Alec Chiu

B.S. Mechanical Engineering, University of Minnesota, Minneapolis, 1991
M.S. Mechanical Engineering, Massachusetts Institute of Technology, 1994

SUBMITTED TO THE DEPARTMENT OF MECHANICAL ENGINEERING IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN MECHANICAL ENGINEERING
AT THE
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

FEBRUARY 1998

Signature of Author: _____
Department of Mechanical Engineering
January 5, 1998

Certified by:
Alexander H. Slocum
Alex and Brit d'Arbeloff Professor of Mechanical Engineering
Thesis Supervisor

Accepted by: _____
Ain A. Sonin
Chairman, Departmental Committee on Graduate Studies

Roadmap of Mechanical Systems Design for the Semiconductor Automatic Test Equipment Industry

By
Michael Alec Chiu

Submitted to the Department of Mechanical Engineering
On January 5, 1998 in partial fulfillment of the
Requirements for the degree of
Doctor of Philosophy in Mechanical Engineering

Abstract

Since the development of the first integrated circuit in 1958, semiconductor products have advanced at an unprecedented rate. Given such rapid technological progress, there is a corresponding need for continuous, rapid improvement in semiconductor processing and manufacturing equipment. Like many other segments of the semiconductor equipment market, automatic test equipment (ATE) manufacturers struggle with the growing mechanical costs and complexity of what is fundamentally a non-mechanical system. Increasing mechanical accuracy, thermal, packaging and cleanliness issues are shifting the architecture of ATE from a test electronics-driven design to one that fully integrates the mechanical and electronic elements. This shift in architecture challenges the ATE industry, which has traditionally not focused on the mechanical expertise and tools required to understand and address these forthcoming issues.

The fundamental contributions of this thesis include a set of models that correlate ATE mechanical performance requirements--such as positional accuracy, signal density and thermal management--to device under test (DUT) parameters such as pin count, bus speed and power consumption. These models, combined with existing semiconductor device roadmaps, are used to generate a prediction of future ATE mechanical development requirements. ATE manufacturers can then use these models and predictions to focus and implement their development efforts.

As an example to illustrate the benefit of focused mechanical expertise, a further contribution of this thesis is the development of a novel kinematic fixturing system that functions as a modular element that is applied to interfacing of test equipment. The design process is described from concept development to product introduction; in addition, test data from prototype and production systems is presented. The performance gains of this new system are then interpreted within the context of the above models to illustrate the benefits of both the design and analysis.

Thesis Supervisor: Alexander H. Slocum
Title: Alex and Brit d'Arbeloff Professor of Mechanical Engineering

Acknowledgments

I would like to thank the many people who have helped and supported me in my efforts. In particular, my thesis advisor Prof. Alex Slocum, who provided the ideas, opportunities and freedom to pursue this work, as well as both Maureen Lynch and Leslie Regan, who often helped in many small ways. Alex d'Arbeloff for his support and all the people at Teradyne who contributed to this effort. I also would like to thank my parents for their encouragement and I would like to dedicate this work to the memory of my father. Most of all I would like to thank my friend and wonderful wife, Maren for her patience, understanding and support during this seemingly never-ending effort.

Michael Chiu
Massachusetts Institute of Technology
January, 1998

This work was funded, in part, by the department of Energy Computational Science Graduate Fellowship

Table of Contents

List of Figures

List of Charts

1 Introduction and Thesis Overview

1.1 Introduction

Since the development of the first integrated circuit in 1958, semiconductor products and processes have advanced at a blinding pace. In 1964, Gordon Moore of Intel observed that the number of transistors per chip, and thus the complexity and performance, has historically doubled every year. He saw no physical limitations that would prevent the continuation of this trend for the foreseeable future, and thus Moore's law was formed. While this pace has since slowed to a mere doubling every 18 months, this simple insight has become the driving goal of the semiconductor industry[1]. Conservative estimates predict that the industry is capable of maintaining this pace for at least 10 more years. Figure 1.1 shows data that supports this observation.

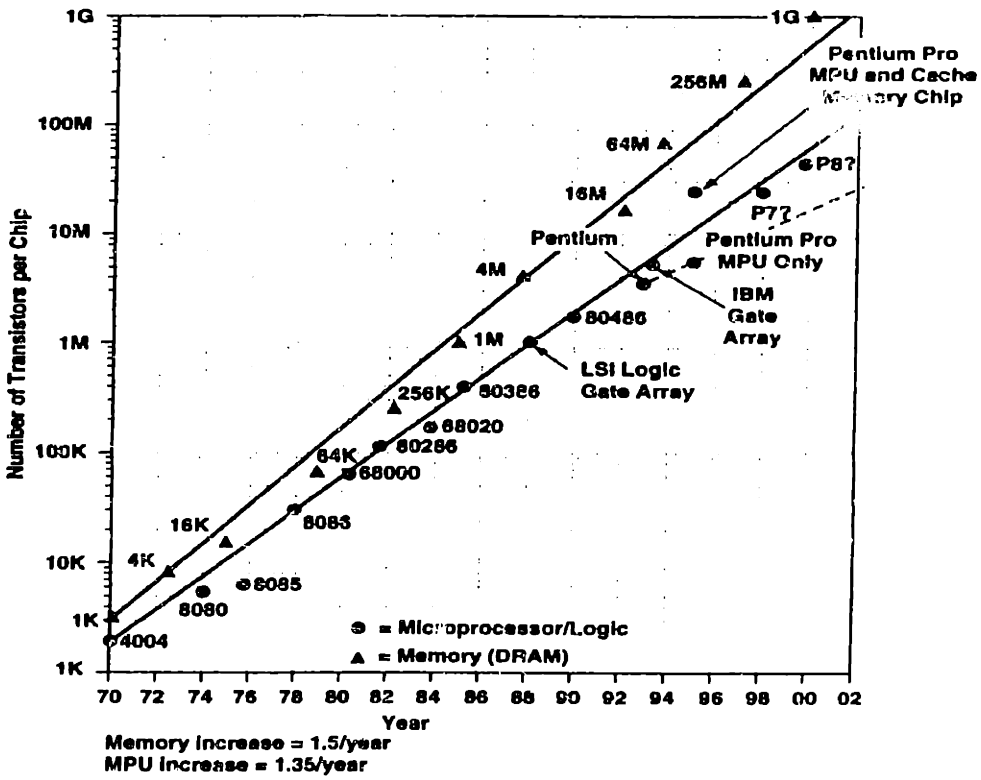


Figure 1.1: Moore's Law[2]

[1] Hutchinson, 1996

[2] Source: Integrated Circuit Engineering

Because of this rapid advance of technology, there is a corresponding need for continuous, rapid progress in semiconductor processing and manufacturing equipment. The design of this equipment is generally driven by the critical process technologies. For example, stepper design is dominated by optics, wet process is driven by chemistry and wire bonding by mechanics. As semiconductor technology marches forward, this critical-process approach limits the equipment performance and a more interdisciplinary, systems-level approach is required to solve the next layer of problems. This limit is most evident in the Automatic Test Equipment (ATE) industry, which traditionally focuses on test pin electronics, with mechanical design as afterthought. Shrinking packages and increasing pin counts are causing mechanical accuracy problems, higher clock speeds are causing thermal and signal transmission problems, and Lilliputian size and power are creating handling and cleanliness problems. Because of these issues, mechanical systems design will soon equal the challenges of electronics design in driving ATE systems architecture.

Due to its focus on test electronics, the ATE industry has an accurate understanding of future electrical needs and challenges and has developed detailed understanding of how they will be addressed. However, this understanding does not include the growing mechanical challenges and currently few tools exist to better evaluate this area. Ultimately, ATE design needs to be driven not by the test electronics or by the mechanics, but instead, the design should reflect the requirements of the device under test (DUT) itself. What is required to accomplish this from a mechanical perspective is an understanding of how device parameters such as pin count, speed, power and feature size, drive the overall mechanical design of ATE.

This thesis presents a body of work that addresses this need, including a high level analysis of ATE mechanical systems architecture requirements, a detailed analysis and model that can be applied to the design of ATE signal transmission systems and a novel mechanical interface design that addresses a critical problem in current ATE systems. This chapter starts with a description of the fundamental contributions of this work. Background information on the semiconductor industry, as well as a more detailed discussion of the ATE industry is then presented. This is followed by a summary of test electronics and test cell mechanics and the chapter closes with an overview of the structure of the thesis.

1.2 Thesis Objectives and Contributions

The objective of this thesis is to address the fundamental mechanical systems issues that face the semiconductor ATE industry. Of primary concern is the lack of a high-level, device-driven understanding of the ATE test cell requirements. Little insight into the needs and trends of ATE mechanics exists which results in an inappropriate allocation of development resources and a mismatch between the performance offered by ATE and the needs of the test community. The pin electronics-driven ATE industry also lacks a detailed understanding of the relationships between high frequency signal transmission lines and mechanical interface design. Finally, there are many pressing mechanical design problems that face the ATE industry. Of particular concern is the design of the interfaces between the tester and device handler/wafer prober because it's performance can have a direct effect on throughput yield and electrical interface reliability. The work described in this thesis addresses all of the above problems. The specific contributions of this thesis are as follows:

1. A set of models and correlations were developed that relate the parameters of semiconductor devices, such as pin count and clock frequency to the performance required of ATE mechanics, such as mechanical accuracy, thermal dissipation, cost and signal density. These device-driven models can be combined with existing device roadmaps to get a better understanding of the future requirements of ATE mechanical systems and to help with the formation of a coherent strategy to address these needs. These models can also be used as tools to optimize the design of current tester architectures by providing insight into the principles that affect tester performance.

2. A detailed model of signal transmission systems was developed that relates their mechanical implementation to electrical performance. This model lends insight into a highly coupled problem that has historically been poorly understood by both the electrical and mechanical design communities. The model can be used to:

- Gain intuitive insight into the relationships that govern transmission lines
- Optimize the design of existing or proposed signal transmission systems
- Investigate the sensitivity of mechanical tolerances such as geometry, accuracy and temperature, on the electrical performance of the transmission line
- Predict the limitations of existing technology

3. A novel interface product, the *Kinematic Docking System* (KDS), was developed from concept through product introduction. This interface was developed using the

deterministic principles and tools that are common in the design of instruments and precision machine tools. In particular, the design is based on the kinematic coupling and thus provides a highly repeatable, perfectly constrained interface between the test head and device handler/wafer prober. This product represents a significant improvement over the current state-of-the-art as evidenced by test data, product sales and industry recognition.

These contributions are significant for several reasons. First, the semiconductor industry plays a significant role in the economy and in our everyday lives and the performance of automatic test equipment plays an integral role in the success of this industry. The models and roadmaps described in this thesis will help semiconductor ATE manufacturers meet the growing demands of the industry. Second, the proliferation of high-speed digital devices and microwave communication products is pushing the limits of existing signal transmission line structures. The analyses developed in this thesis represent a first step in understanding the limitations of the current ATE signal transmission fixturing technology and provides a structure in which new technologies can be evaluated. Finally, the *Kinematic Docking System* solves a relevant and important problem that faces the ATE industry by using sound engineering principles and practices from the precision machine design industry.

1.3 Overview of the Semiconductor Industry

In the 40 years since the invention of the integrated circuit, the semiconductor industry has grown to the point where it impacts almost every aspect of modern life. Semiconductors are ubiquitous; desktop computers are now a required workplace tool, micro-controllers are scattered throughout automobiles and home appliances, toys can now think and speak, and communication devices such as pagers and cell phones are now wearable and integrated into watches, headsets and even jewelry. In the past 20 years, the worldwide semiconductor industry has grown from less than $10 billion in 1980, to over $150B today. Forecasts predict that this industry will continue to grow and will exceed $300B by the turn of the century[3]. Figure 1.2 shows the historical and forecasted progress of the semiconductor market, including the increasing electronic content of consumer products.

[3] Source: Dataquest

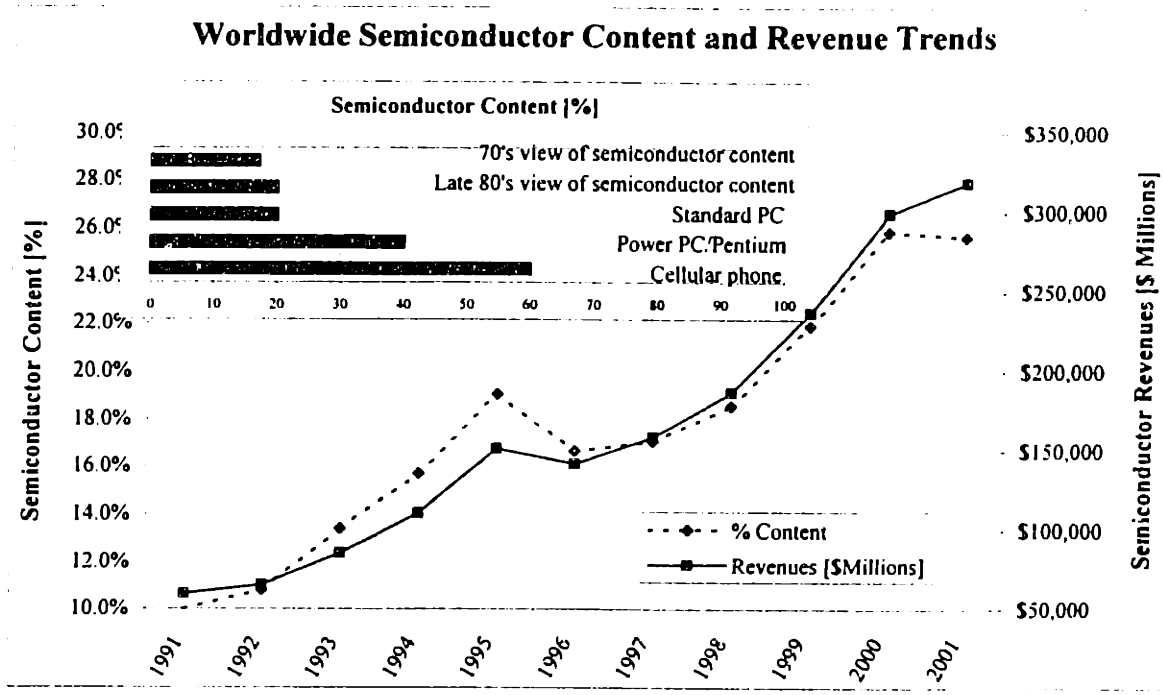


Figure 1.2: Semiconductor Content and Revenue Forecasts

The mechanics of early semiconductor fabrication equipment are relatively simple and have not changed substantially in the past 20 years. The rapid progress of the semiconductor industry has been enabled primarily by advances in solid state physics, electronics, chemistry and manufacturing engineering, and these advances have driven the gains in feature density, productivity and cost reduction that fuel the proliferation of semiconductors. Although the fundamental semiconductor fabrication processes are based on non-mechanical principles such as optics, wet chemistry and vapor deposition, their manufacture still requires a tremendous amount of mechanical equipment and processes. Most of these processes utilize equipment whose mechanical performance in some way limits the overall process capability. For example:

- Minimum feature size of circuits is limited by the mechanical accuracy of wafer steppers and optical focusing mechanics.
- Multi-layer interconnects are limited by coating and planarization processes.
- Defect rates are limited by particulate contamination.
- Packaging designs are limited by wire-bonder speed and accuracy.

Figure 1.3 shows the typical process flow of semiconductors and mechanical operations, from the fabrication (front-end) processes, through the test, assembly and package (back-end) processes[4].

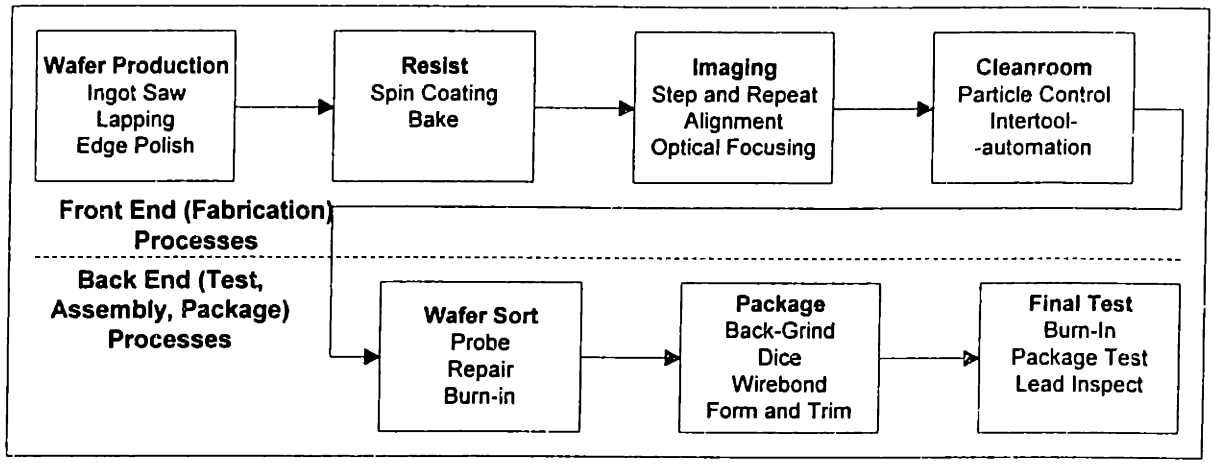


Figure 1.3: Semiconductor Manufacturing Process Flow

As device feature sizes approached the wavelength of the light used in the lithography process[5], the easy gains of the past became much more difficult to realize due to increasing mechanical complexity. Figure 1.4 shows the historical and predicted evolution of device feature size. As device feature size approaches 1 nm, the benefits of smaller devices diminish when compared to the requisite effort and expense. As evidence of this, it is estimated that a new 0.25 micron, 300mm wafer fab will cost in excess of $2 billion dollars.

[4] Elliot, 1986

[5] The wavelength range of visible to deep UV light is approximately 1 micron to 1 nanomenter

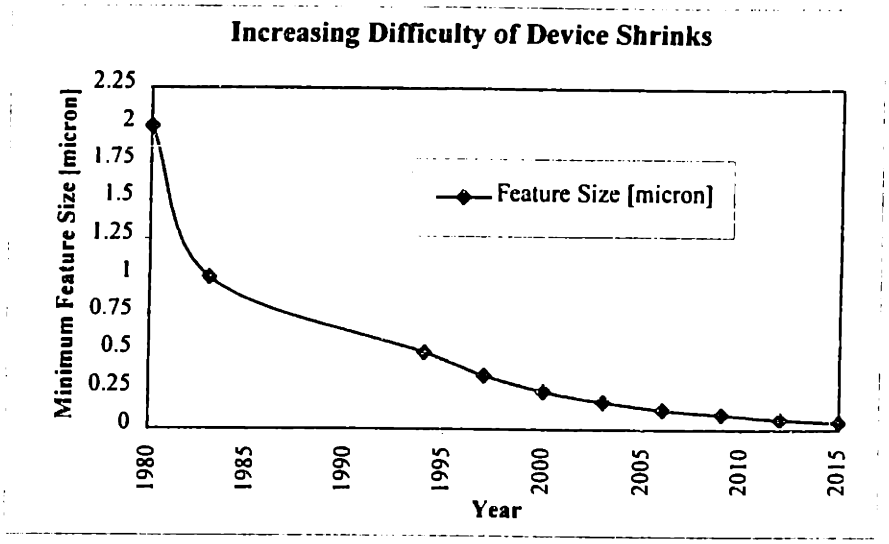


Figure 1.4: Evolution of Minimum Feature Size

In addition to optical limitations, device feature size is also constrained by the mechanical precision of the wafer stepper and mechanical planarization methods. This flattening of the S-curve has required front-end equipment manufacturers to focus more effort and innovation on the mechanical development of process tools. In general, front end process equipment development has seen tremendous growth in the past 10 years. Diminishing productivity and performance gains from feature shrinks have been offset by increased mechanical accuracy, increases in wafer sizes, inter and intra-bay automation and improved cleanroom practices.

While front-end equipment has already benefited from this shift in emphasis, many back-end processes are still crossing this discontinuity between purely electrical systems to more integrated electro-mechanical designs. Most back-end processes deal with mechanical features such as bond pads and package leads that are orders of magnitude larger than lithographic features. Therefore this time lag between front and back end mechanical development is expected, but it is none-the-less inevitable because device and package features are quickly approaching the limitations of back end processes.

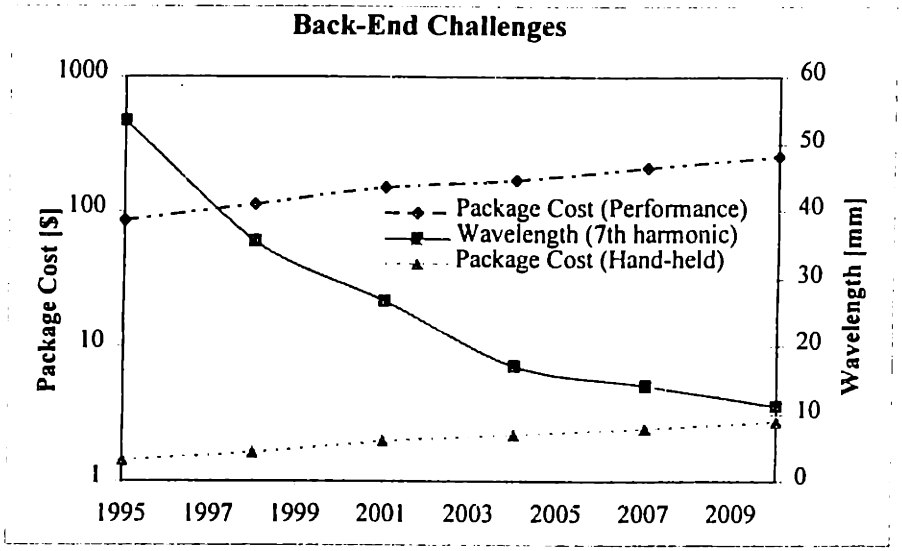


Figure 1.5: Graph of Back-End Challenges

Figure 1.5 shows data that supports this assertion. Packaging costs are approaching fabrication costs due to the increase in pin counts and the inability to reduce cost per pin. Device clock speeds are also increasing to the point that the wavelength of the signal is approaching the scale of the device. Both of these trends will drive back-end process equipment manufacturers to increase the mechanical focus of their product development.

1.4 Semiconductor ATE Industry

The semiconductor automatic test equipment is a $1.4 billion dollar industry[6] that is responsible for the design and manufacture of instrumentation and automation equipment used to test semiconductor devices in the final stages of manufacturing. Virtually all integrated circuits are tested in some way, most are tested as a final verification before shipment and additional tests are also run to grade or bin parts according to speed, efficiency or other parameters. Testing is also used to identify defective devices while still in wafer form to allow for the repair or rejection of bad die before the cost of packaging.

[6] Source: Dataquest

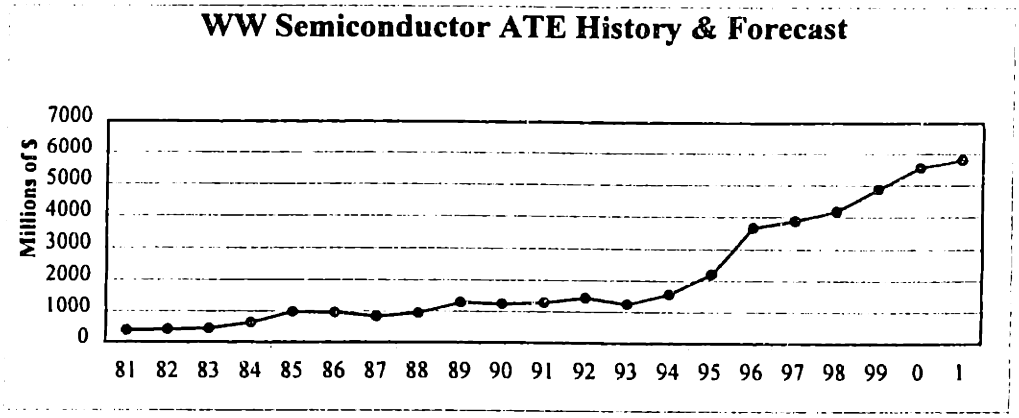


Figure 1.6: ATE Market History and Forecast

The ATE market had its beginnings in the 1960's, before the explosive growth of IC's and before the benefits of quality control and test automation was recognized. Prior to this, discrete components such as diodes and transistors were manually tested for functionality and sorted by leakage current and breakdown voltage or other parameters. Although early test systems still required a human operator, they improved productivity by combining several manual tasks into a single, reliable, calibrated instrument. An early model of one of these instruments, shown in Figure 1.7, was a simple box that tested leaded devices such as diodes and transistors. The interface on these early testers consisted of two slots into which device leads are inserted and the tests were essentially performed at DC.

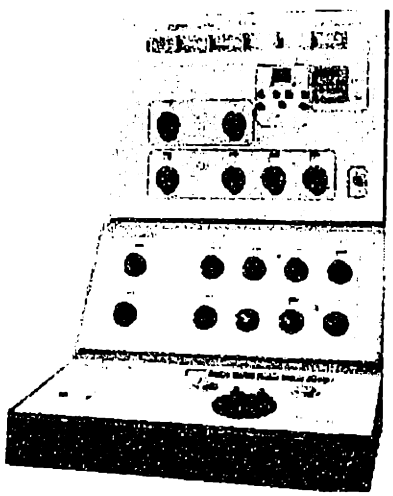


Figure 1.7: Photograph of an Early Diode Tester

There are a number of distinct segments to the ATE industry. Pin electronics, or tester manufacturers develop the electronics instrumentation responsible for testing the DUT.

Prober manufacturers design the mechanical, optical and thermal systems that deliver the DUT to the pin electronics while it is in wafer form. Handler manufacturers produce the mechanics that handle the DUT when it is in package form and a cottage industry of small vendors exists that builds the tooling, interfaces and software that allow this equipment to work together on a production floor.

1.4.1 ATE Test Electronics

The complexity of test equipment has grown along with the complexity of integrated circuits. Overall, however, the fundamental mechanics of test systems have changed little in the past 20 years; a tester remains essentially an assembly of electronics, and the mechanics function merely as an enclosure. Cooling, packaging and cost concerns have arisen, but they are generally addressed by evolutionary means. As testers evolved, sockets were added to test integrated DIP packages, cables were added to interface with automatic device handlers and speeds crept into the kHz range.

This started to change in the late 70's. At this time, most integrated circuits were based on standard bi-polar transistor technology, which is characterized by low output impedances and generally low speeds. Rising popularity of portable electronics, such as transistor radios and walkmans, drove the demand for low power products and increased the popularity of digital complimentary metal-oxide semiconductor (CMOS) devices. CMOS combines low power dissipation with stiff outputs that swing the full supply range, allowing the design digital, battery powered products such as laptop computers, camcorders and cell-phones. However, the benefits that make CMOS attractive also create problems for the test engineer; the fast edges and low power consumption accentuate the detrimental effects of stray capacitance in the signal lines between the DUT and the measurement instrument. At high frequencies, these lines behave as low-pass filters, and the CMOS drivers on the DUT were not designed to drive the capacitances that existed in the long cables of ATE systems.

This problem was also addressed through evolutionary means. At first, a small contingent of additional electronics, primarily drivers, amplifiers and relays, were added at the end of the test wires, close to the DUT. At the time, this solution was practical because the typical device had at most a few dozen pins and the collection of electronics near the device was quite small. Over time, this solution evolved into the current architecture; a tester

that is segmented into a small test head connected to a larger mainframe via a cable bundle. The test head contains drivers, relays and utility electronics that need to be close to the device, as well as additional space for custom, device-specific fixturing and electronics. The mainframe houses the bulk of the instrumentation: power supplies, drivers, pattern generators, measurement systems, computers etc... A modern test system, such as the one shown in Figure 1.8 below, can easily exceed $1M and often approaches $2M.

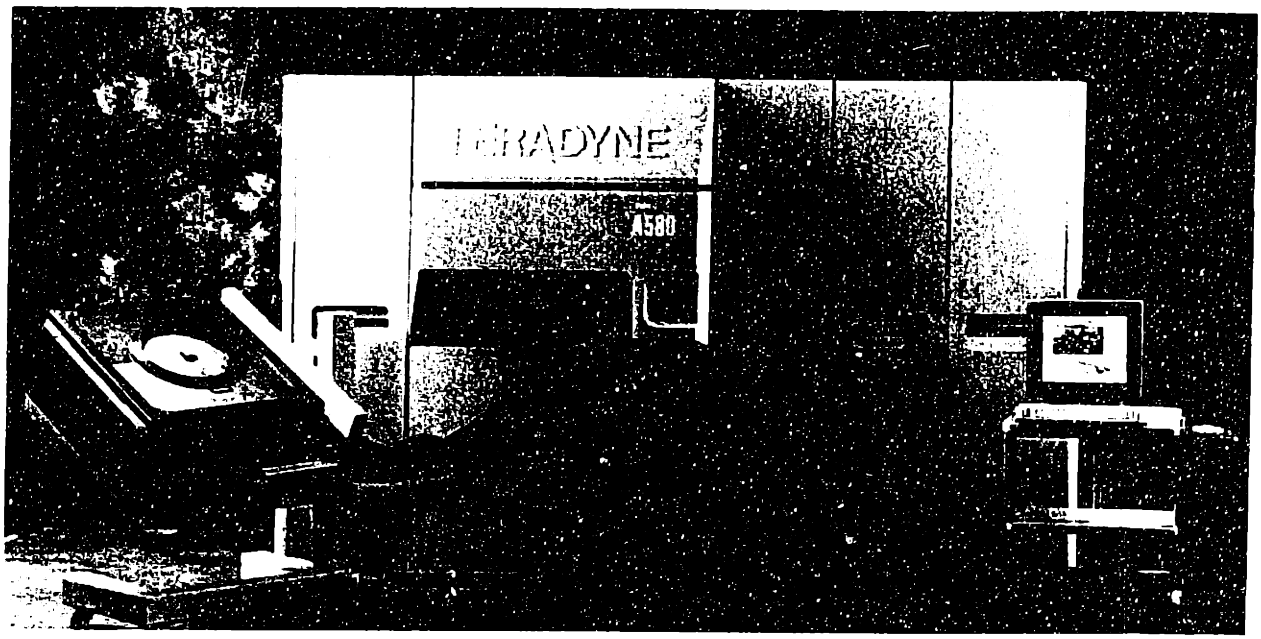


Figure 1.8: Modern Semiconductor ATE system

As semiconductors devices evolved, the test electronics industry migrated into three primary groups: memory, logic/VLSI[7] and analog/mixed-signal. Each group has a unique set of device specific problems and therefore have somewhat different mechanical systems needs. For example:

- Analog testing has the widest range of products spanning all frequencies from DC power, to audio frequency (20-20K) to RF and microwave (1-20GHz) for communication. Analog devices are generally low in volume and high in variety and therefore require a flexible system that can be reconfigured easily and frequently.
- VLSI devices, mainly microprocessors and digital ASICS[8], are characterized by high speeds, many pins and a mix between stable product (microprocessors) and a low volume products (ASICS). Because VLSI devices have high pin counts, but

[7] Very Large Scale Integration
[8] Application Specific Integrated Circuit

little variety, VLSI ATE requires dense fixturing with relatively small amounts of space for custom circuitry.

- Memories are high volume commodity products with long test times. Memory testers increase throughput by sharing resources and testing multiple devices in parallel. High volume and complex parallel test fixturing results in relatively inflexible memory test solutions.

While this historic segregation of device applications is convenient for test equipment designers, it may soon be outmoded. The growing demand for digital electronics that interact with an analog world is driving the need for devices that combine digital to analog (DAC) and analog to digital (ADC) circuitry with microprocessors, memory and communications circuitry. This trend is fueled by demand for portable, high performance products such as laptop computers with cell modems, digital video cameras and interactive cars. These 'system on a chip' (SOC) products will necessitate a new type of tester that is capable of economically testing memory, digital and analog circuitry on the same instrument. Therefore, a new mechanical systems approach will be required that accommodates the mechanical requirements of each of these device groups.

1.4.2 ATE Test Mechanics

Device testing is performed at several stages in the back-end of semiconductor manufacturing. The number and type of test varies with device type, but generally, devices are tested at least twice; once while in wafer form and again after packaging. Certain devices such as memories and some ASICs may also undergo a repair/trim operation as well as an elevated temperature burn-in operation, but these tests generally require a separate set of test electronics.

1.4.2.1 Wafer Probers

Testing devices in the wafer form requires an additional piece of equipment called a wafer prober, which contains the mechanics that delivers the wafer to the pin electronics contained in the test head. The test head is generally fixtured, or docked, to the prober, as shown in Figure 1.9.

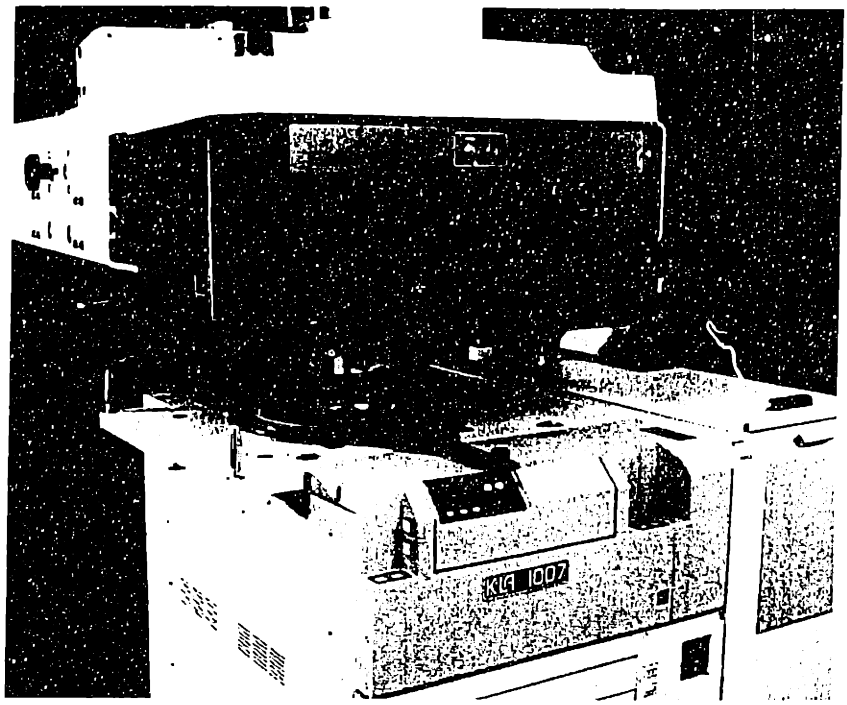


Figure 1.9: Test Head Docking to a Wafer Prober

A prober consists of several sub-systems, including:

- Stage: X-Y-Z motion + metrology capable of micron resolution.
- Chuck: yaw motion, vacuum and temperature control –80/+200°C.
- Optics: needed for device to pin alignment.
- Environmental: thermal conditioning and air filtering.
- Controls: user interface, prober-tester interface, subsystem control.

Probing interfaces directly to the bond pads of a die, which are on a 50-100 micron pitch. Because of this, the mechanics involved in this operation are quite delicate and are sensitive to vibrations, thermal growth and external forces. Due to the delicate nature of the probing operation, it is often difficult to fully test a complex device at this stage and tests are often run at reduced speed. Generally, the probing operation is used to identify and eliminate non-functioning, non-repairable devices so that they can be discarded prior to the packaging process, which often represent a significant portion of the total cost of the device. Testing is usually done with 100% coverage, with yields ranging in the 60-99% depending on the maturity and complexity of the device. Complete, at-speed tests are run on the packaged device to sort the product by function and performance.

New technologies such as flip chip and multi-chip modules (MCM) eliminate the packaging step by attaching a bare die directly to a printed circuit board (PCB) or within a small sub-assembly. These technologies requires the use of known good die (KGD) and therefore drive the need for full function, full speed testing at probe. Probers start at $200K and can often exceed $500K if options such as optical wafer alignment, auto probe card changing and air filtration systems are added.

1.4.2.2 Package Handlers

After the probe operation has sorted the die and, in some cases, minor flaws are repaired, the wafer is diced and the good die are packaged. After package, many products are subjected to an extended 'burn-in' test conducted at elevated temperature to eliminate the devices that have marginal performance. After all the processing and packaging processes are complete, a final test is performed at-speed to sort the devices by performance and to insure that no bad die are shipped to the end user. This test is generally done using the same test program and test system, but mated to a specialized handler that is designed to handle the specific type of package being tested. There are many different models of these machines designed to handle the growing array of package types.

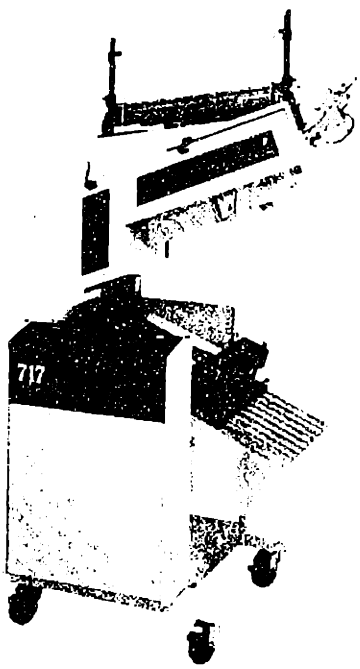


Figure 1.10: Gravity-Fed Handler

There are two primary types of handlers; gravity feed and pick-and-place. Gravity feed handlers, like the one shown in Figure 1.10, can only accept devices in tubes (often referred to as 'sticks'). Therefore, it is limited to the lower pin density parts such as through-hole DIP and J-leaded QFP[9] packages. Of greater interest are the newer pick-and-place and tray handlers that are capable of accepting almost any package type, including perimeter leaded surface mount (SMT) and area leaded pin grid array (PGA) and ball grid array (BGA) devices. Lead pitches on packages are historically 1.25 – 2.5 mm (0.050" – 0.100") and thus positioning tolerances are not as tight. These handlers manipulate each device discretely or in small batches and contain similar sub-systems as a prober, including:

- Stage: X-Y-Z motion + metrology capable of mil resolution
- Optics: device lead alignment and inspection
- Environmental: thermal soak capable of –80/+200°C
- Controls: user interface, handler-tester interface, subsystem control

A simple, single-site gravity-fed handler can cost less than $50K. However, a more advanced multi-site pick-and-place handler with thermal chamber and lead inspection can exceed $200K. A typical pick-and-place handler is shown in Figure 1.11.

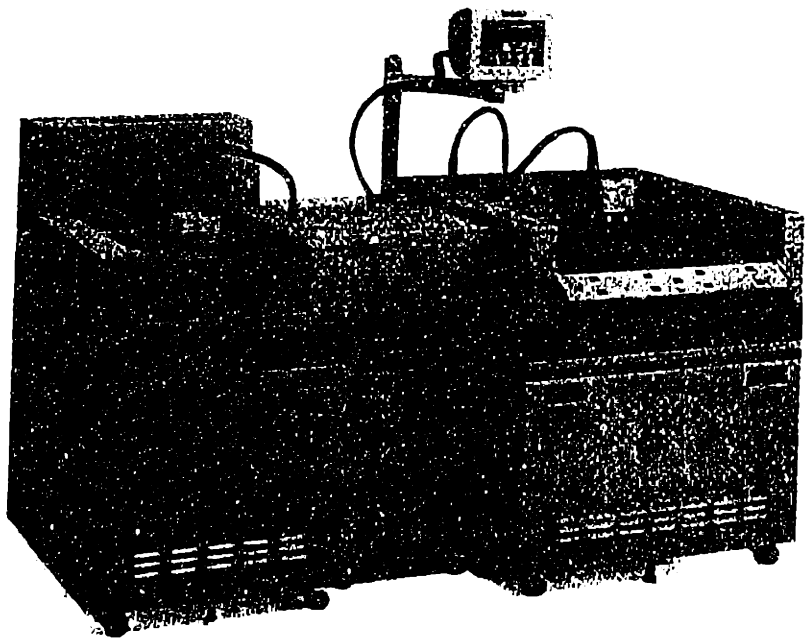


Figure 1.11: Pick-and-Place Handler

[9] Dual Inline Package, Quad Flat Pack

Handling is a less delicate process than probing; the mechanical scale of the components is greater, the devices are less fragile and the cleanliness requirements are not as strict. However, newer, high-density packages are now available with 0.4 mm (0.016") pitch and flip chip pad array device may have pitches as small as 0.12 mm (0.005"). These finer pitch devices will make high-speed handling increasingly difficult.

1.4.2.3 Interfaces and Manipulators

The last segment of ATE mechanics deals with the interfaces between the test head and the wafer probers and device handlers described above. With the exception of high volume memory and microprocessor testing, test equipment is generally not dedicated to a particular product or package and therefore test floors are usually in a state of continuous flux. Also, many on-shore manufacturing facilities share test equipment with engineering, who use it during the day while production is run on second and third shifts. Therefore, test head-hander and test head-prober setups often need to be reconfigured several times per day, and interface tooling and support equipment are designed to make this task as robust as possible. There are several primary components required for a successful interface, including:

- Tooling: Device specific components used in testing such as package test sockets, wafer probe fixtures and custom circuitry.
- Inner Dock: Provides alignment and interconnection between tooling and test head.
- Outer Dock: Provides alignment and support between test head and prober or handler.
- Manipulator: used to support and move test head during docking.
- Tooling Changer: Occasionally used in prober to automatically change tooling without undocking the test head.
- Peripherals: Additional equipment such as microscopes, inkers and thermal chambers.

Tooling costs can vary from $10K to $50K per device type and docking and manipulator costs can easily exceed $100K per test head. Tooling changers and peripheral equipment are typically included in the cost of the wafer prober.

Tooling generally consists of a collection of custom, device specific circuitry, a test socket or probe needle array, and a PCB 'space transformation' that connects the signal pins on the test head to the pins on the socket. Often, this tooling is separated into 2 or more PCBs that are interconnected by an array of compressible pins. These PCBs, called device

interface boards (DIBs) are often attached to a metal stiffener to maintain board planarity. Figure 1.12 shows some typical tooling, and Chapter 4 discusses this in greater detail.

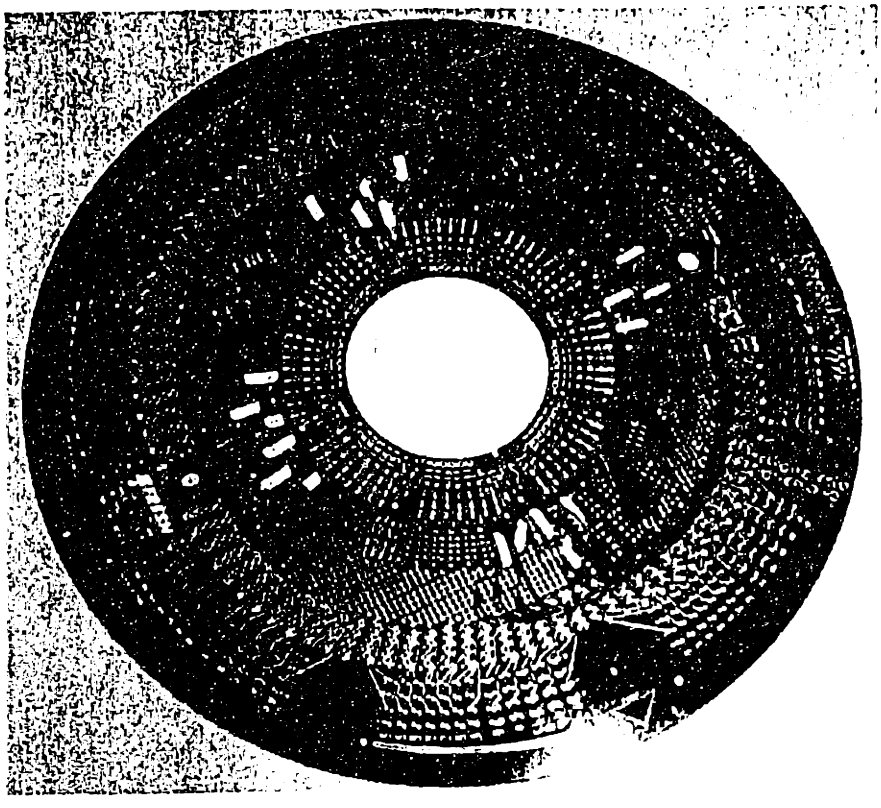
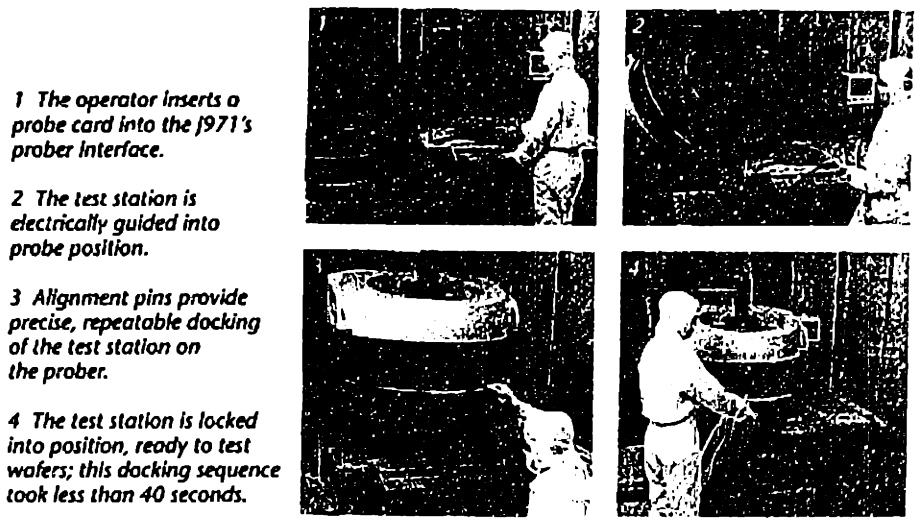


Figure 1.12: Device Interface Board (DIB) Tooling

Figure 1.13: Inner Docking Mechanism

The inner and outer docking mechanism align and interlock mating test cell components such as DIB's to the testhead and the testhead to the device handler. There are many variations of inner and outer dock mechanisms. Historically, the two functions were combined into a single mechanism that accomplished both simultaneously with a combination of progressively tighter pin-bushing alignment features and a cam-follower or vacuum-assisted mechanisms to pull the two components together. Figure 1.13 shows the DIB and other tooling which is generally sandwiched within this inner docking mechanism, and Figure 1.14 shows the operation of this docking process. Because the tooling is trapped between these two pieces of equipment, the test head usually needs to be undocked and moved away from the prober/handler to change tooling.



1 The operator inserts a probe card into the J971's prober interface.

2 The test station is electrically guided into probe position.

3 Alignment pins provide precise, repeatable docking of the test station on the prober.

4 The test station is locked into position, ready to test wafers; this docking sequence took less than 40 seconds.

Figure 1.14: Process of Docking a Testhead to a Prober

The primary task of the manipulator is to support the weight of the test head and cable bundle so that they can be easily moved by an operator and so that they do not transfer significant loads to the prober, handler or interface. This is usually accomplished using counterweights and pivots located near the test head center of mass. The size and weight of the test head and cable bundle has increased with each generation and current systems may exceed 1000 lbs. This has made counterbalancing and manipulation increasingly difficult. A typical manipulator and test head are shown in Figure 1.15.

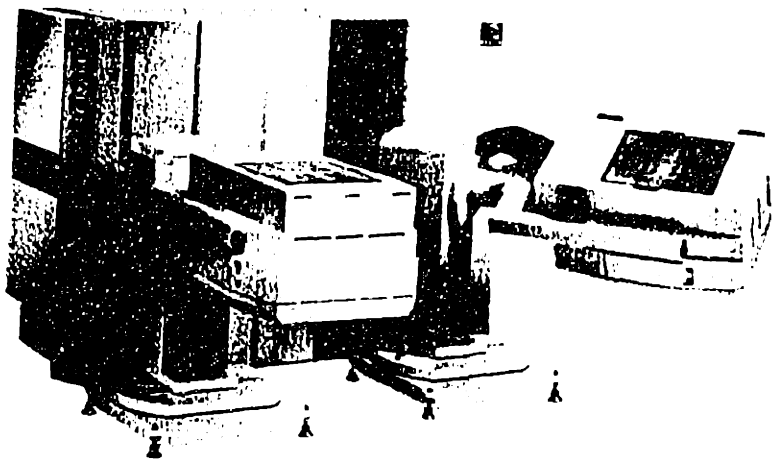


Figure 1.15: Test Heads with Manipulators

The remaining interface components; tooling changers, inkers, microscopes, etc... are often used to increase the functionality of the test setup, but are generally not required to perform the test. For example, tooling changers that automatically change the DIB or probe fixture are used to eliminate the need to undock the test head when changing tooling and microscopes are often used to help with probe setups. Most of these features are being integrated into newer test equipment.

1.4.3 Definition of an ATE Test Cell

Much like manufacturing test cells of the machine tool industry, test equipment is often arranged in small groups, or cells of commonly used equipment. For the purposes of this study, a test cell is defined as all of the equipment required to transform untested devices into tested, binned product. The primary component of a test cell is the tester itself, which constitutes the vast majority of the capitol cost of the system. In addition to this, a prober or handler(s), tooling, manipulator, software and peripheral equipment may be required. Currently, wafer test cells are somewhat distinct from handling test cells. Wafer cells are generally operated in a class 1000 or better cleanroom and a test head is dedicated to a single prober. The test head is undocked when product is changed, but equipment is moved only for repair or maintenance. Handler cells are usually much more flexible. Often a single device will be sold in multiple package types and therefore require several different handler models. Ideally, a single handler cell contains all equipment required to test packaged devices, but often some equipment is shared between cells.

1.4.4 Need for Automatic Test Equipment

Testing can be done at any stage of the manufacturing process. The historic view of quality control dictated that a final test or sample is required before product is shipped. Among other things, the total quality management (TQM) movement of the late 1980's suggested that this is generally not the best approach. Instead, it is better to model and understand the manufacturing process and eliminate defects, thus eliminating waste and the need for testing. This concept has been embraced by the semiconductor industry and great effort is spent improving yields. However, despite these efforts, yields on mature integrated product, such as a microprocessor or micro controller, generally do not exceed 95%. The primary reasons for this that the continuous advance in process technology and the short life cycles of products that prevent the industry from developing a mature process.

If one considers that testing is a capitol intensive, non-value-added process, an argument could be made for the elimination of testing of high-yield parts. However, the nature of how semiconductor components are used dictates that the 1-5% failure rate will incur more costs than just that of the scrapped parts. This is because these devices are generally only one component in a much larger assembly, and even a small failure rate at the component level can quickly grow to have unacceptable results downstream for a moderately complex assembly[10].

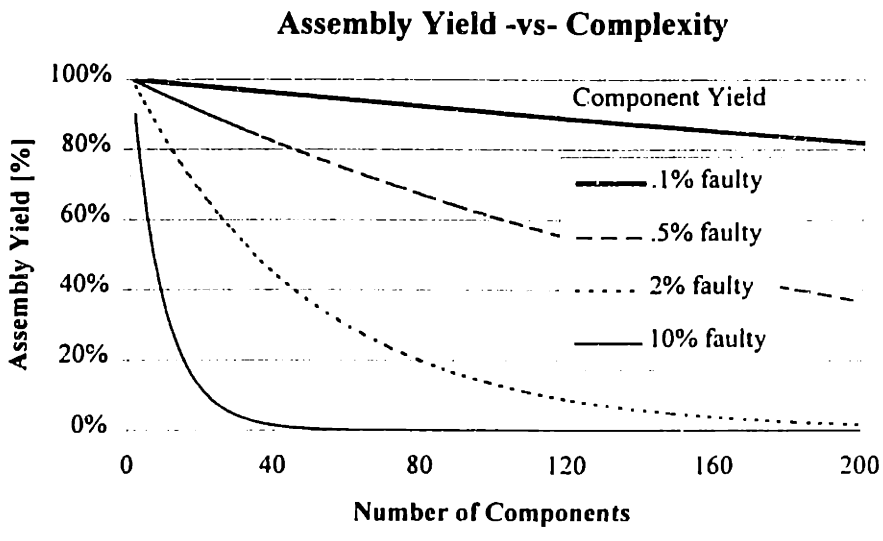


Figure 1.16: Relationship Between Component and Assembly Yield

[10] Davis, 1982

This relationship is shown in Figure 1.16; a small component failure rate of 0.5% has the potential to cause system failure rates in excess of 50% for a relatively complex assembly.

The approach of using testing to reduce downstream costs is also applied within semiconductor manufacturing process. Historically, greater emphasis has been placed upon package test as a final quality control step and wafer sort was conservatively guard-banded to eliminate all bad and marginal parts from the packaging process. As yields increased from 20% in the 60's, to nearly 95% today, packaging processes became more robust, and the emphasis of testing has shifted to wafer sort. Today, yields at package test generally exceed 95% and therefore the scrap cost of packaging bad die has been eliminated at probe. If fab and packaging yields continue to improve, and devices can be fully tested at probe, it is possible that the final package test step can be reduced to a mere statistical sampling. However, until the fab processes and product designs mature, it is unlikely that testing for improperly packaged devices can be eliminated altogether.

1.5 Thesis Structure

In addition to the introductory remarks in Chapter 1 and conclusions in Chapter 8, this thesis is divided into three, two chapter sections.

The first section contains a high level view of the semiconductor market trends, how it will affect test cell mechanics, and some predictions and recommendations on how the ATE industry should address these changes. Chapter 2 discusses the device roadmap data collected by the Semiconductor Industry Association (SIA). This data represents an industry consensus on the overall direction of the semiconductor market. Data from this roadmap is interpreted within the context of how it will influence the design of ATE mechanics. Chapter 3 explicitly interprets how trends shown in this data will affect the mechanics of semiconductor ATE. First-order models of the relationships between roadmapped device parameters and ATE system design are presented. These models attempt to predict the future requirements of test cells, such as mechanical accuracy, cooling capacity and signal density, by relating them to device parameters, such as pin count, clock speed and power consumption.

The second section discusses the development and application of a model that can be used to analyze signal transmission systems, a critical component of ATE systems design. Chapter 4 begins by reviewing the underlying theory of high frequency digital signal

transmission and some of the fundamental difficulties in the design of these systems. This is followed by a description of the physical implementation of signal transmission lines in ATE and other electronic systems. The chapter continues with an overview of existing microwave modeling methods, covering the S-parameter methodology in particular and concludes with a description of how these tools are combined to build a model relating the physical implementation of these transmission lines to the electrical performance. Chapter 5 discusses the use of the model and presents a number of sample analyses, including a baseline model of an existing system, data relating mechanical accuracy to signal performance and a sensitivity analysis using manufacturing tolerances. It then uses this model to predict some of the limitations that the mechanical implementation will place on the design of future ATE systems.

The third section describes the design, development and verification of the *Kinematic Docking System.* Chapter 6 starts by discussing the problem definition, including the functional requirements, constraints and goal of the design effort. This is followed by a brief description of the principles of kinematics as well as some preliminary analysis. The chapter continues with a detailed description of the KDS product and closes with some proposals for future improvement. Chapter 7 presents the outcome of the KDS development project. This includes performance data from all stages of development, including field tests of prototype units, life cycle tests of production units and verification tests in a controlled environment.

Chapter 8 concludes with a summary of the fundamental contributions of this thesis work, as well as some discussion of how this work may be pursued in the future.

1.6 Summary of Chapter 1

This chapter opened with a brief history of semiconductors and a preview of the significant contributions of this thesis. This is followed by a more detailed description of automatic test equipment segment of this industry, including a discussion when and why devices are tested and a description of the mechanical and electronic systems that are used in these tests. The chapter closes with an outline of the remaining chapters.

2 Device Roadmaps: The Future of Semiconductor ATE

2.1 Introduction

"The complexity of semiconductor technology is increasing at such a rapid pace that there is significant danger that the U.S. industry will not be able to continue its historical rate of progress without a common vision and increased cooperation in precompetitive research and development."[11]

What is true for semiconductor fabrication is also true for semiconductor ATE. But in some ways, the test challenge is even more daunting for the designers of test equipment. Wafer fabrication process equipment manufacturers use leading edge technologies in optics, deposition and etching to determine the pace of semiconductor technology. However, ATE manufacturers must build their instruments using commercially available semiconductor technology, and these systems must be capable of testing the devices created five years in the future! Therefore, it is critical that ATE manufacturers understand the needs and trends of the semiconductor industry and how these trends will define the requirements of future ATE systems.

This chapter discusses the first half of this effort to understand and roadmap ATE requirements. It begins by defining technology roadmaps and their value and limitations. This is followed by a summary of the SIA device roadmap results, with some comments on their accuracy and limitations. Finally, these device roadmap predictions are interpreted within the context of ATE design requirements, that is, which data is important and why. Chapter 3 will discuss the ramifications of these device trends in greater detail and make some suggestions for potential ATE architectural approaches/solutions.

2.2 Technology Roadmap Defined

Roadmaps are used to plan a journey and navigate through uncharted terrain. A well-planned roadmap must contain both the final destination as well as the expected route, alternate paths and anticipated detours. Analogous to a cartographers roadmap, a technology roadmap is a tool that is used to describe the direction and anticipated route for the development of a given technology. For example, a roadmap for the auto industry might

include the engine fuel efficiency goals, accompanied by existing, prototype and proposed engine technologies that are capable of meeting those goals within the required timeframe. A roadmap is not a definitive plan, rather it is an industry's best guess of the those technologies that are most likely to succeed. Armed with this understanding of future needs and potential solutions, the auto industry can make relatively informed decisions regarding distribution resources to the maintenance of existing products, development of new products and research of new technologies. Thought of another way, a roadmap is an attempt to understand the position on and shape of the S-curve[12] that the technology is following. This understanding will help the industry predict and plan for discontinuities, which are a major source of failure of established firms.

2.3 The SIA Technology Roadmap

The Semiconductor Industry Association (SIA)- a consortium of equipment and device manufacturers, end users and academia- develops a biennial roadmap that describes the progress of semiconductor technology. This roadmap encompasses all aspects of the design, development and manufacture of integrated circuits including design tools, lithography, materials, packaging, test and environmental safety. The fundamental premise of the SIA roadmap is a continuation of the historical gains predicted by Moore's law. This suggests that the industry overall is still on the steep slope of the S-curve and has not yet reached the flattened crest of diminishing returns.

With the assumption that semiconductor density will continue to double every two years, the SIA roadmap lays out the device parameters required to meet this goal, including feature sizes, clock speeds, voltage levels and package densities. This is followed by a detailed description the processes required in each area to meet this goal. For example, minimum feature sizes are expected to shrink from 0.50 micron today to 0.05 micron by 2013. The specific advances in lithography, mask, etch and polishing technologies and levels of development required to meet this goal are then described.

Much of the SIA roadmap data is market specific and is sorted into the market segments shown in Table 2.1.

---

[11] Foreword to The National Technology Roadmap for Semiconductors, 1994
[12] Foster, 1986

Commodity	Consumer products, microcontrollers, disk drives, displays.	General focus on cost System < $300
Hand-held	Battery-powered and hand-held products, cellular phones.	Power consumption and size System < $1000
Cost/Performance	Laptop and desktop and industrial microprocessors.	System < $3000
High-performance	High end workstations, servers, supercomputers, military	High cost (system > $3000), low volume
Automotive	Analog, hostile environment, higher power	Moderate cost, moderate volume
Memory	DRAMs, SRAMs	Cost sensitive commodity

Table 2.1: Device Market Segments

Each of these market segments imposes different constraints on the design of the test cell. For example a high performance ASIC may require a large number of signal pins but have low power consumption and an automotive device may only have a handful of pins, but dissipate a large amount of power. Fortunately, it is unlikely that a semiconductor manufacturer would need to test both of these devices on the same test floor. This explains the historic separation of analog, VLSI and memory test markets. However, the growing demand for system-on-chip (SOC) devices mentioned in Chapter 1 is the exception to this rule and these devices often have high performance VLSI, memory and analog communications content on the same silicon. Therefore, the SOC represents the worst-case device test scenario and it is this set of roadmap parameters that will be used to define the functional requirements of future test cells. The SIA roadmap does not explicitly define a category for SOC devices, but it can be approximated by the trends of the cost/performance and high performance market segments.

2.4 ATE Roadmap Data

Data describing the needs and goals of the ATE industry is also presented in the SIA roadmap. However, little direction is given regarding what technologies are required to meet these needs in general and no specific insight is given into the mechanical technologies and development required. Despite the lack of specific focus on ATE mechanics, the roadmap data is still useful to the extent that it provides some bounds for the definition and resolution of ATE mechanical systems challenges. The graphs and charts that follow

represent the 1994 SIA roadmap device data that is relevant to ATE mechanical systems development.

2.4.1 Device Geometry, Pincount and Package Density

The most obvious device parameters of interest to ATE mechanical systems design are those that relate to device geometry and packaging. The primary parameters of concern are the total device pin count and the device size. These two parameters determine the mechanical accuracy and signal packing density required to contact the DUT. There are several ways to view these parameters. The most useful for understanding the mechanical accuracy needs are to look at separate cases for wafer probe and package handling.

The primary mechanical drivers for wafer probe are pin density and chip size. The density, or pad pitch, depends on whether the pads are arranged in a perimeter or area array. Figure 2.1 shows the raw data provided by the roadmap, chip size and pad count. Figure 2.2 displays the same data transformed into pad pitch. The perimeter and area pitches represent upper and lower bounds for these values.

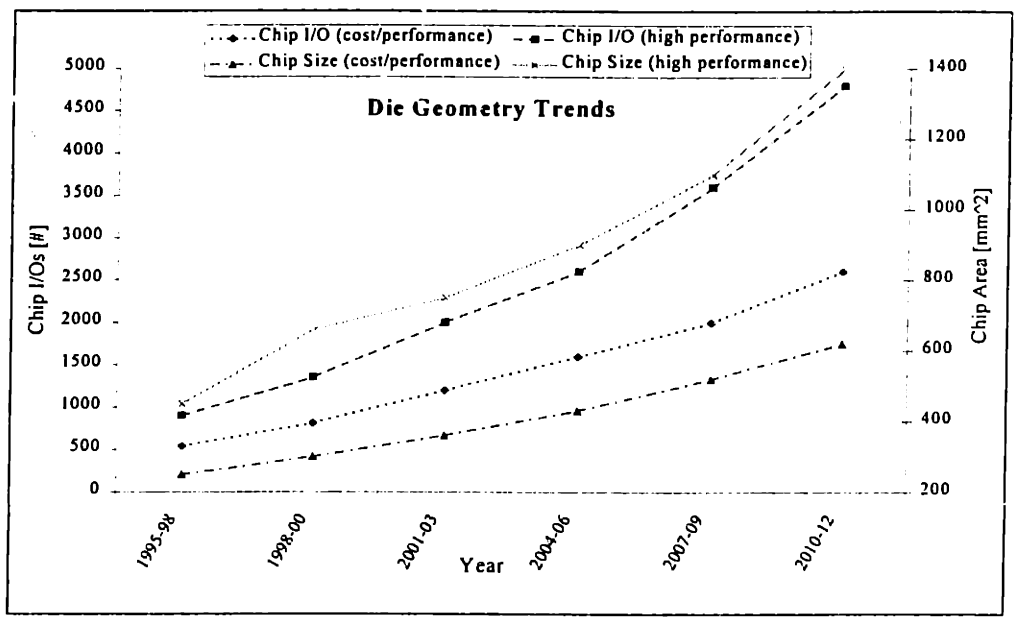


Figure 2.1: Die Pincount and Area Roadmap Data

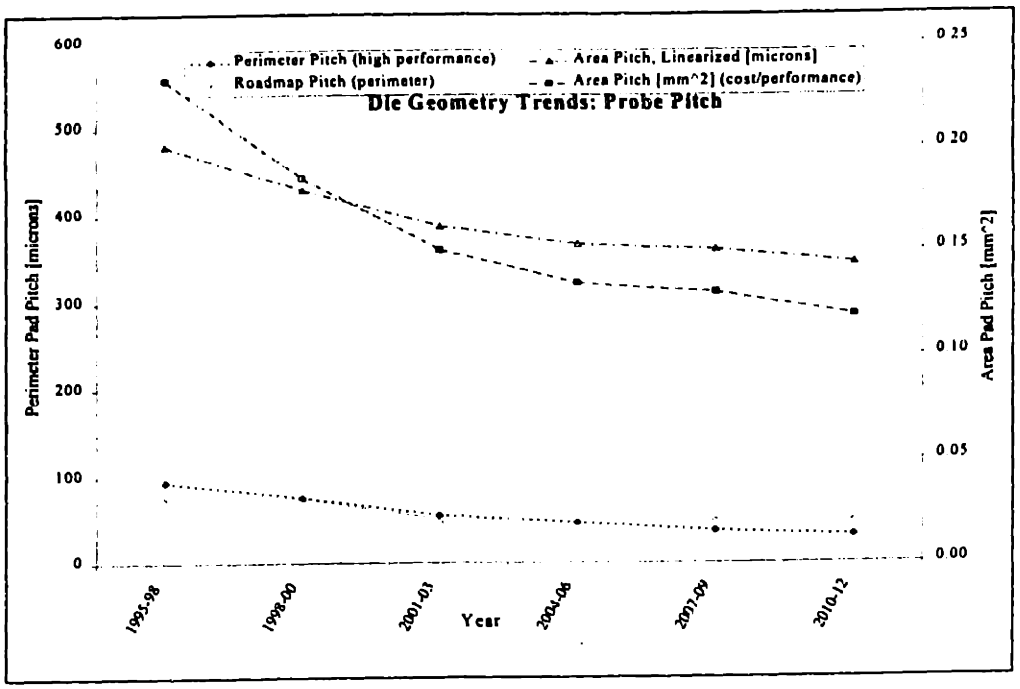


Figure 2.2: Die Pad Pitch Roadmap Data

The drivers for device handling are similar to probe, but the roadmap explicitly reports pin count and pad pitch. Device size is less important in handling applications because the mechanical alignment and contacting of package leads in handling is done differently than in probe. This data is shown in Figure 2.3.

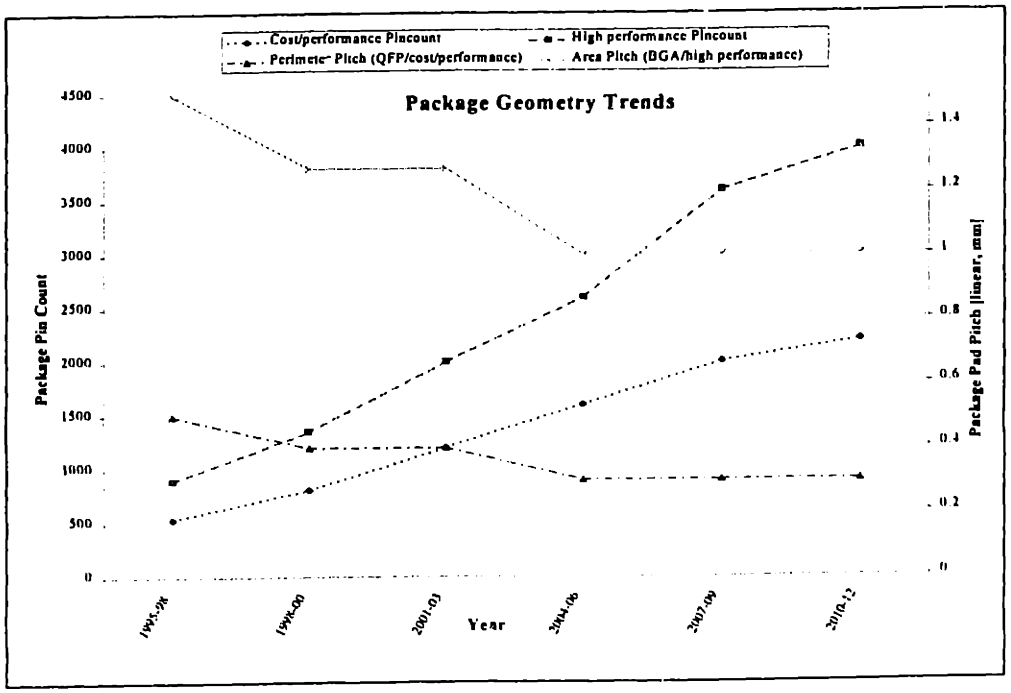


Figure 2.3: Package Pin Count and Pitch Roadmap Data

There are two additional parameters mentioned by the SIA roadmap that may have some effect on the above data. The first is the prevalence of boundary scan, design for test (DFT) and built-in self test (BIST) circuitry on the DUT. These technologies propose to integrate test resources directly on to the DUT, during the design stage, that will be used solely for test purposes. This sacrifice in usable device area will result in simpler test equipment that will require fewer pins. To date, designers have been slow to adopt these techniques and thus the roadmap has been overly aggressive in it's predictions. However, increasing pin counts and test complexity will naturally drive the adoption of these or similar methods. Regardless of how successful these methods become, the baseline predictions of pin count will represent a conservative bound of the overall test need.

Figure 2.4 shows two BIST trends suggested by the SIA roadmap. Percent of average IC tested roughly represents the percent reduction in pins required to test a given device. Percent of chips using scan port represents the expected level of industry acceptance. It should be noted that regardless of what percentage of pins are eliminated by BIST, the pad or lead pitch will not change and until all devices adopt this technology, ATE vendors still need to design against the worst case of no BIST functionality.

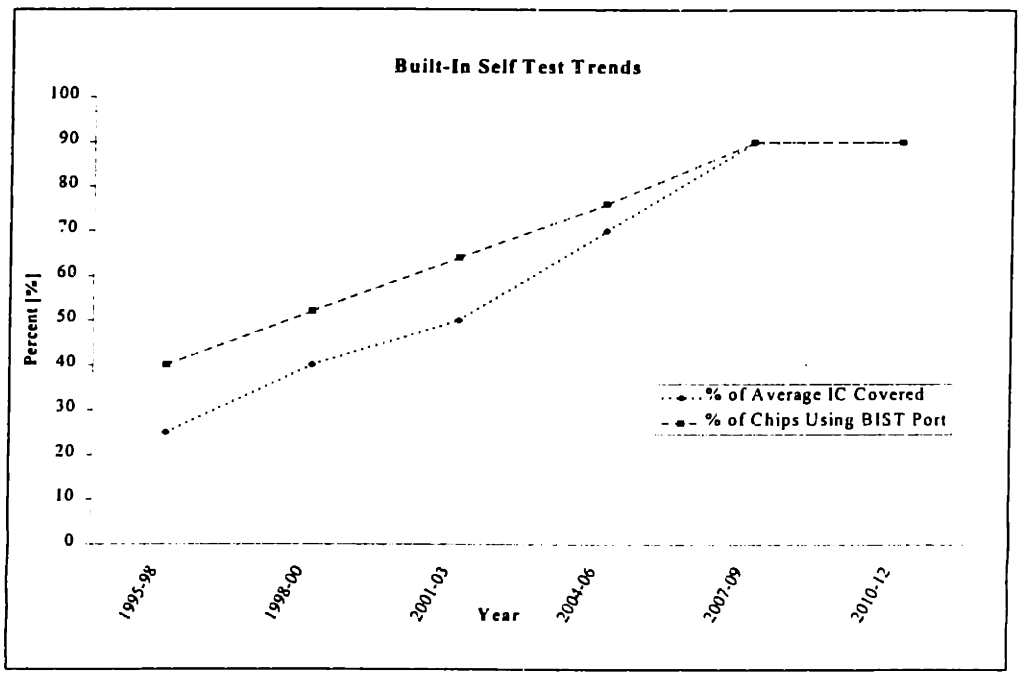


Figure 2.4: Built-In Self Test Adoption Roadmap Data

The second trend suggested by the roadmap is the demand for known good die (KGD). No numerical data is given, only the fact that bare die handlers will be needed for testing. If this trend becomes significant, the tight accuracy requirements of probing will be imposed on the handler environment, significantly increasing the difficulty of this task.

2.4.2 Clock Frequency

A second device parameter that affects mechanical systems design is the device operating frequency or data rate. There are two values given for this parameter, on and off chip, or internal and bus speeds. The internal (on-chip) clock determines the speed at which computation is done within the device and has little effect on tester design other than the requirement that the tester provide a single or small number of synchronized pins at this clock rate. The external (off-chip) bus speed is the more important parameter because it determines the rate at which data is passed into and out of the device. To properly evaluate the device, the tester must be able to read, write and process data at this speed. Moving digital data at high speeds is more than an electrical and data I/O problem. At high enough frequencies, such as those suggested in the roadmap, the physical implementation of the signal lines can have a significant effect on the test performance. Therefore, even from a mechanical systems perspective, it is important to understand the trends in bus frequency.

Historically, the roadmap has underestimated the internal clock frequencies and overestimated the bus frequencies. High-end desktop computers are currently running at 300-400 MHz, with plans to reach 900 MHz by 1999. Bus speeds on these same computers are at 66-80 MHz, with only a few specialized machines breaking the 100 MHz barrier. The 1994 roadmap data shown below predicts that on-chip clocks will not reach 800 MHz until 2007 and off-chip speeds will exceed 150MHz by 1998. Figure 2.5 shows a historical view of processor clock frequencies and performance and Figure 2.6 shows the SIA roadmap predictions for both on and off chip clock rate[13].

[13] 1994 and 1997 off-chip roadmap data are similar but the '97 data projects further into the future

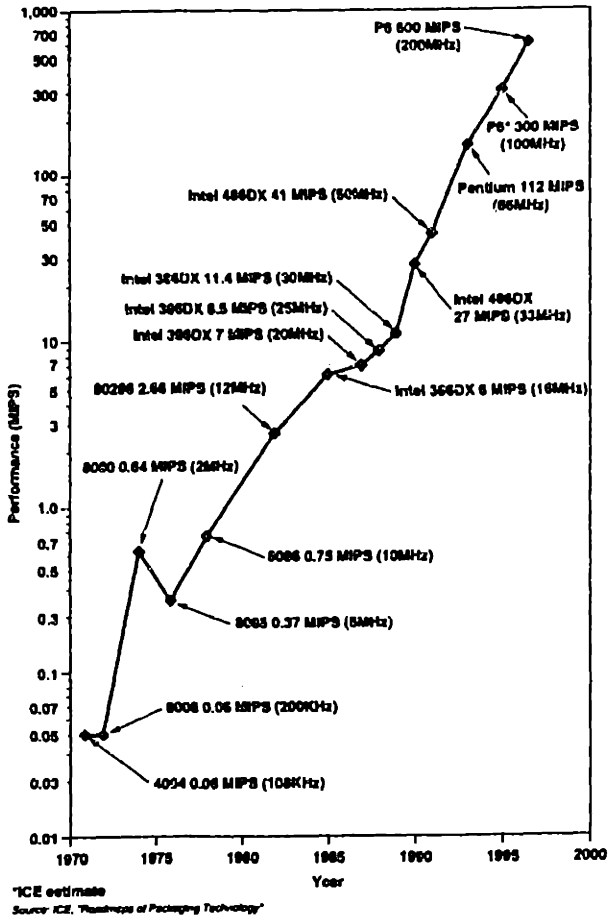


Figure 2.5: Historical Clock Rates for Intel Processors[14]

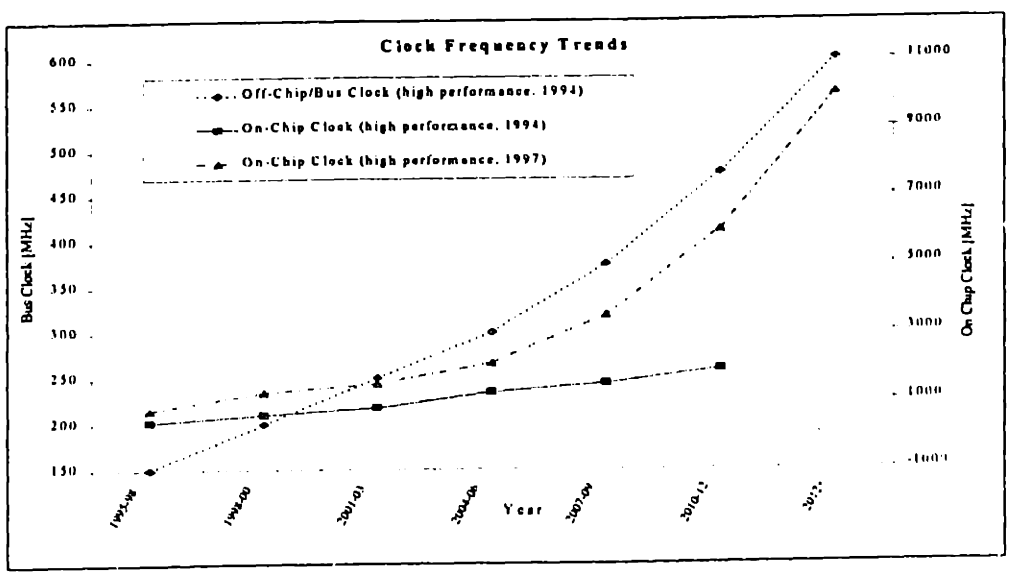


Figure 2.6: Clock Frequency Roadmap Data

[14] Integrated Circuit Engineering, 1997

2.4.3 Voltage and Power

Device voltage levels and power consumption are being driven by two factors. Demand for longer use of battery powered devices is the primary driver and to this end, any reduction in power is beneficial. The industry has a goal of 0.9V TTL operating voltage, which represents the practical end-of-life limit of common battery cells. The second factor is the maximum power dissipation of device packages, which dictates the junction temperature, which in turn affects device reliability. The Arrhenius rate equation, based on statistical mechanics, suggests that every 10°C increase in temperature will double the device failure rate.[15]

Lower voltages have several significant effects on devices and their application. Lower signal levels imply lower signal-to-noise ratios and therefore test instrumentation and signal transmission systems must be capable of detecting and discriminating signals at higher noise levels. In addition, at constant power levels, a lower voltage implies higher currents, which results in increased I^2R losses and voltage sag. Both of these have historically been addressed by increasing the number of power and ground pins, further exacerbating the growing pin-count problems described in Section 2.4.1. For high power, high pin count devices, power and ground pins may represent up to 40% of the die and package pins.

Larger die, denser features and packaging and higher currents all suggest that total device power dissipation will continue to increase. At some point, passive cooling methods are insufficient and the addition of heatsinks, fans, heat pipes or liquid cooling technologies will be required, all of which add cost and volume to the device. This is true for both the DUT as well as the devices used in the test instrumentation.

[15] Lienhard, 1995

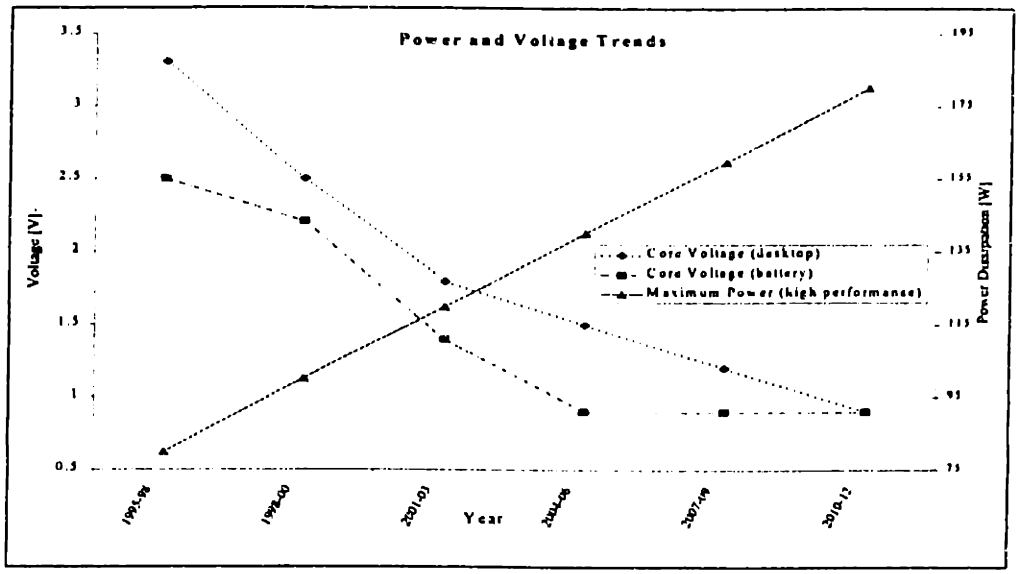


Figure 2.7: Device Power Dissipation and Voltage Roadmap Data

Figure 2.7 shows the SIA roadmap data for device power dissipation and supply voltage. Power values are separated into three groups; high performance with heatsink, logic without heatsink and battery powered. The devices that employ heatsinks draw considerably more power and represent the worst-case to test. Voltage levels are separated into desktop and battery applications, with the only distinction being desktop devices are lagging battery-powered by several generations.

2.4.4 Cost

The industry has historically separated the mechanics and pin electronics of ATE and this distinction is apparent in the roadmap. The data on test cost-per-pin includes only the capitol cost of the test electronics and does not include device handling and fixturing capitol costs, nor does it include per-test charges. Considerable focus has been given to cost of test economic models and equipment utilization, but the roadmap does not present specific data on the cost per test or test cost as a percentage of device cost. Part of the reason for this is that device manufacturers consider this data proprietary and part of the reason is that this area is still very poorly understood.

Figure 2.8 shows the SIA roadmap data on tester cost/pin. The overall tester cost is also approximated by multiplying the cost/pin by a fraction of the total number of pins for a high

performance device.[16] While the specific data is useful only in the design of testers and pin electronics, the overall data trend and slope can be used as indicators of the total test cell and device cost-of-test trends.

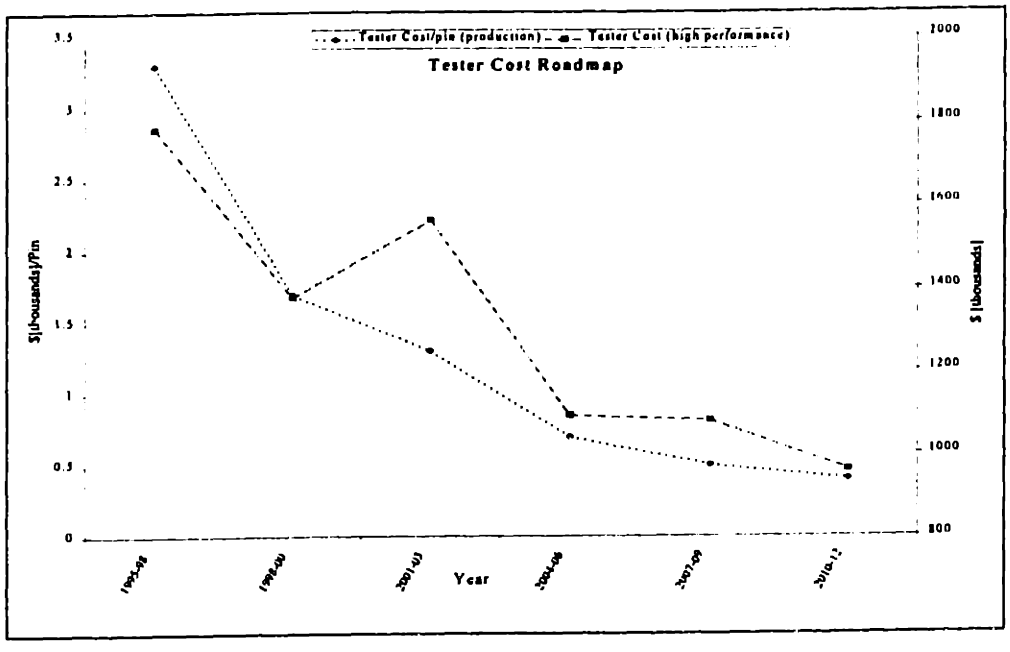


Figure 2.8: Tester Cost Roadmap Data

2.5 Summary of Device Roadmap Data

The SIA roadmaps additional parameters for ATE that do not directly affect the mechanical systems design, including the software, simulation and test vector requirements. The data presented above is summarized in Table 2.2 below.

[16] It is assumed that 40% of the pins are for power and ground and therefore are not included in the total pin count when calculating cost.

Timeframe ->		1995-98	1998-00	2001-03	2004-06	2007-09	2010-12
Technology ->		0.35 micron	0.25 micron	0.18 micron	0.13 micron	0.10 micron	0.07 micron
1a. Number of Chip I/O's	[#]						
cost/performance (microproc.)		540	810	1200	1800	2000	2800
high performance (ASIC)		900	1350	2000	2800	3600	4800
1b. Package pin count	[#]						
cost/performance (uproc.)		540	810	1200	1800	2000	2200
high performance (ASIC)		900	1350	2000	2600	3600	4000
2a. Chip Size	[mm^2]						
cost/performance (uproc.)		250	300	360	430	520	620
high performance (ASIC)		450	660	750	900	1100	1400
2b. Package Pad Pitch	[mm]						
QFP/Perimeter (cost/performance)		0.5	0.4	0.4	0.3	0.3	0.3
Area/BGA (high performance)		1.5	1.27	1.27	1	1	1
3a. On-Chip frequency	[MHz]						
cost/performance		150	200	300	400	500	625
high performance		400	600	800	1250	1500	1900
3b. Chip to board bus	[MHz]						
cost/performance		50	66	100	100	125	150
high performance		150	200	250	300	375	475
4a. Core Voltage	[V]						
desktop		3.3	2.5	1.8	1.5	1.2	0.9
battery		2.5	1.8-2.5	0.9-1.8	0.9	0.9	0.9
4b. Max Power	[W]						
cost/performance		16	18	22	28	37	55
high performance		80	100	120	140	160	180
5. Tester Cost	[$/pin]						
production		3.3K	1.7K	1.3K	0.7K	0.5K	0.4K

Table 2.2: Summary of SIA Device Roadmap Data Relevant to ATE

The data above does not present a complete picture of the needs of ATE manufacturers. Additional interpretation of this data is required to understand the quantitative relationships between these device parameters and the mechanical design requirements of ATE systems. Chapter 3 investigates these relationships in detail and attempts to make some broad predictions of the future needs of ATE mechanics.

2.6 Roadmap Limitations

While the SIA roadmap represents a consensus of industry experts, it is important to note that the data presented above are only predictions, not facts. There are several limitations to this data. First, the roadmaps contain the explicit assumption of Moore's Law. Many of the data points are simply linear fits to the curve dictated by this assumption and it is possible that the industry has already entered the top of the S-curve, distinguished by diminishing returns on effort. Second, the experts polled for this data are, to a large degree, all entrenched in the same industry. This is exactly the type of scenario that fails to predict drastic changes in technology trajectories. Finally, the roadmap has been proven significantly wrong in the past. This includes both over and underestimation of critical parameters.

While these caveats limit the scope of how this roadmap data may be applied, they do not invalidate its application in developing a technology strategy. Chapter 3 uses this data to make some high level predictions of the mechanical systems requirements of future ATE test cells. In light of the above limitations, it is important to remember that these predictions can only be as good as the roadmap data upon which they are based.

2.7 Summary of Chapter 2

This chapter presented a general definition of technology roadmaps and an introduction to the SIA semiconductor roadmap. Data from this roadmap is presented that describes the overall industry trends as well as the specific goals of the ATE market. Data relevant to the ATE industry is presented, however, none of this data explicitly describes, or proposes solutions to the mechanical systems challenges that face the ATE industry.

3 ATE architectural roadmap

3.1 Introduction

The device roadmaps described in Chapter 2 were developed to help the semiconductor industry understand and plan for future needs. These high-level trends are the first step in developing a description of the mechanical systems design requirements of future automatic test equipment. The next step is to translate these device trends into design requirements that can be used to understand the mechanical systems needs at a more detailed level. This understanding is especially critical in the area of ATE mechanics because there has historically been little focus or expertise in this rapidly changing field.

The goal of this chapter is to identify high-level issues regarding the mechanical requirements of future test systems. This will be accomplished by examining the basic relationships between the device parameters roadmapped in Chapter 2 and the critical ATE test cell mechanical parameters such as accuracy, thermal performance, signal delivery performance and cost. This chapter is arranged as follows: section 2 discusses some assumptions regarding what is modeled and what is important, sections 3-5 describe the specific models, results and recommendations, Section 6 summarizes the results and closes with some concluding remarks.

3.2 Modeling Assumptions

The ATE market spans a broad range of device types and performance ranges. To perform meaningful analysis, some assumptions must be made and constraints placed on the scope of device types. The primary assumption made in this study relates to what devices and applications will drive the development of future test systems. Insight into the test challenges of the future can be found by looking at today's leading edge devices. A test cell capable of supporting these lead devices will also have the performance to test simpler devices. This assumes that performance, not cost should be the dominant driver for future test systems.

The class of devices that represent the greatest test challenge is that which contains elements from all three device markets, analog, VLSI and memory. Examples of these system-on-chip (SOC) devices of the future, currently in their infancy, can be found in the

multimedia and handheld products of today. For example, Figure 3.1 is an excerpt from an advertisement for a modern system-on-chip device that is designed for hand-held computer products. This device contains the following components:

- VLSI: 66 MHz CPU
- Analog: Codec/telephone, speaker, IR, touch panel, A/D converter
- Embedded memory: 5 KB on-board cache
- Digital data busses: (32 bit, 20 MHz) for RAM, ROM, PCMCIA, RS232, PC, keyboard, LED and LCD displays

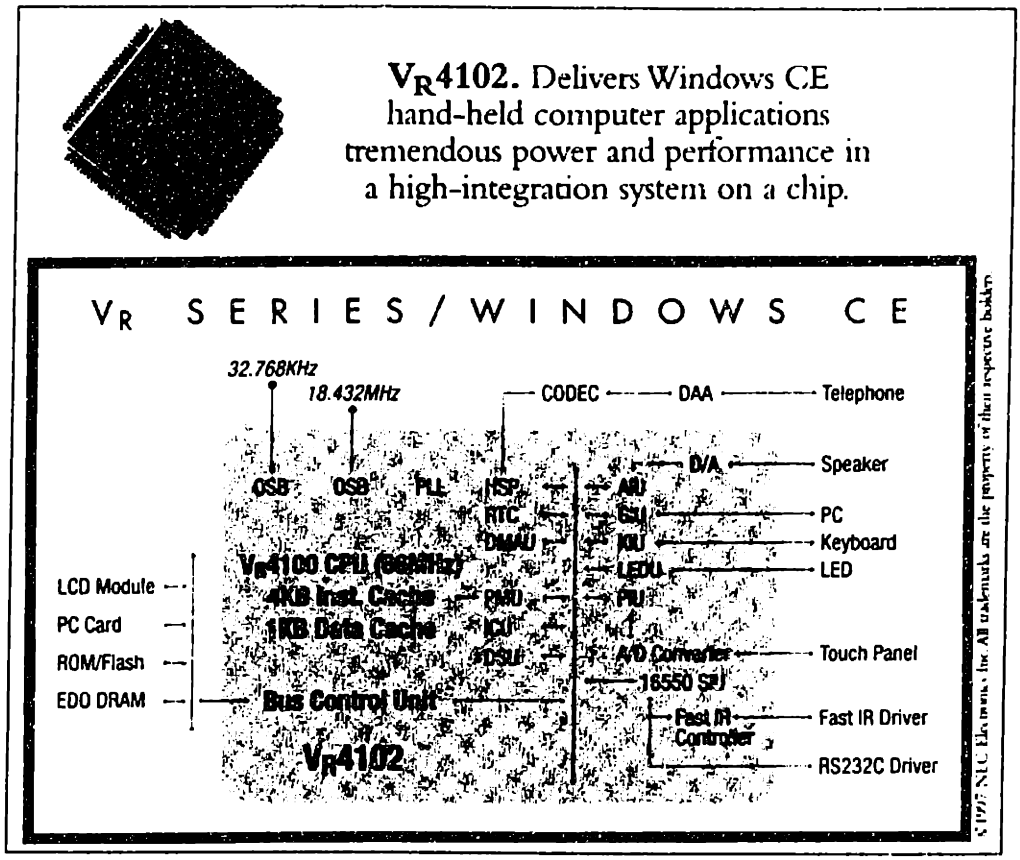


Figure 3.1: Modern System on a Chip Device

The specifications shown above does not represent the state of the art in any of the three categories. Today, high-end VLSI processors currently run at 500 MHz, memories are made in 64-megabit blocks and analog design is far beyond CODEC and 12-bit A/D conversion. There will always be specialized devices that have higher performance than a SOC device in a particular category and these devices will continue to incrementally push the upper boundary of ATE performance. However, as SOC devices evolve, this performance gap will narrow. The challenge will then shift to understanding and managing the trade-offs between

analog, VLSI and memory test architectures. Therefore, the parameters of future SOC devices should loosely track those of the SIA roadmap presented above. Specifically, the quantitative assumptions are as follows:

- Pin Count: Pin counts will reside between the cost/performance and high performance curves and therefore should increase from about 750 today to 3500 by 2012. Digital I/O will account for the majority of these pins, approximately 50%. Power and ground pins will follow with about 40% and analog content will represent 10% balance of pins.
- Clock and Bus Speed: One of the primary advantages of integrated silicon is the elimination of external busses, with all blocks running off the same internal clock. However, it is likely that a external, high speed expansion or memory bus will always be required and therefore bus speeds should fairly closely track those of the roadmap, increasing from 150 MHz today to 600 MHz in 2012.
- Supply Voltage and Power: Voltage levels should track the hand-held device roadmap, dropping to 0.9 V sometime in the next 5-7 years. Maximum power levels should follow the high performance device data, linearly increasing from 80 to 180 Watts over the next 15 years.
- Embedded Memory: Embedded memory size will remain small, but important. No trending data is available on this aspect of device design, but it is assumed that any memory that takes more than 2 seconds to test will require a separate test/burn-in process. If bus speeds keep pace with internal memory size, this assumption will remain true.

The sections that follow will discuss how the above trends will likely affect the design of future ATE test cells. Specifically, the impact of these trends on cooling systems, mechanical accuracy and fixturing as well as total test cell cost will be modeled and the results will be interpreted within the context of the capabilities of existing technologies.

3.3 Test Cell Thermal Systems Roadmap

There are two independent benefits to reducing the overall operating temperature of an electronics instrument. Device reliability is the most common reason cited, as failure rates increase at higher temperatures. The Arrhenius rate equation, a statistical mechanics description of diffusion at the atomic level, is often used as a basis for understanding the relationship between temperature and semiconductor failure rates. At nominal operating conditions, this equation roughly predicts that failures due to thermal phenomenon will double every 10 degree Celsius rise in temperature. While there are many disputes regarding

the accuracy of this prediction, there is general agreement that a strong relationship exists between temperature and reliability.[17]

Device performance is the second common objective of temperature control. Performance gains are manifested in two different phenomenon. First, the maximum operating speeds of complimentary metal oxide semiconductor (CMOS) and other semiconductor devices exhibit temperature dependence; the colder the silicon, the shorter the gate and RC delays. This phenomenon can yield a 2-5X increase in clock speed if cooled to cryogenic temperatures and at more typical operating conditions, it ties instrument accuracy directly to temperature. Data illustrating this effect is shown in Figure 3.2.

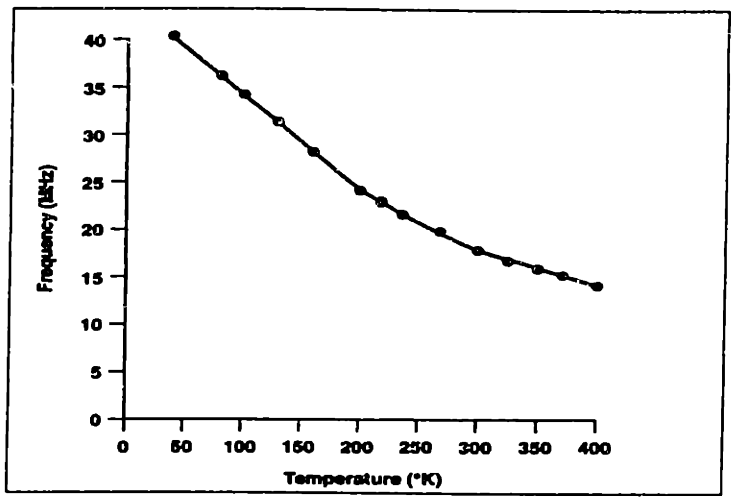


Figure 3.2: CMOS Speed –vs- Temperature

The second relationship between performance and temperature is a bit more subtle and relates back to the device reliability. It is common practice in the electronics industry to derate the performance of components based on the temperature delta above ambient. For example, the gain, input and bias currents of typical op-amps are rated at several different temperatures. If the maximum operating temperature cannot be well controlled, the instrument designer must derate the performance of the components based on the worst-case operating temperature resulting in a sub-optimal design.

In addition to reducing the maximum operating temperature. a second critical factor in the performance of test instrumentation is the minimization of thermal gradients across the PCB. These gradients can cause timing errors between instruments resulting in measurement

[17] Lienhard, 1986

errors or test failure, and at greater temperature extremes, thermal gradients can also cause assembly failures due to mechanical warpage. A common failure mode for surface mount devices, especially ball grid arrays, is mechanical failure of the soider joint due to thermally induced shear stresses. However, these mechanical-assembly failure modes are best dealt with locally, through improved package and interconnect design rather than a global cooling system redesign.

Temperature stability in some ways is also a performance issue. As mentioned above, there are well-defined relationships between temperature and device performance and therefore changes in temperature will affect the accuracy of the test results. For this reason, all ATE systems perform periodic calibrations to reduce this effect. Typically, this calibration is performed at system startup and after every 2 degree Celsius change in temperature. This recalibration process can take several minutes and often up to ½ hour. The 2 degree Celsius value is a compromise between accuracy and convenience; as the temperature stability decreases, this interval must also increase to avoid frequent recalibration, which reduces the system throughput.

3.3.1 Thermal Analysis

There are two independent aspects to thermal management: how much power is being generated, and how much can be dissipated by the cooling system. Generally, the former drives the design of the latter. The two sections that follow describe these two aspects of system design with the goal of developing a roadmap of cooling needs and capabilities. The section on instrument power trends attempts to predict cooling requirements of future test cells based upon the anticipated power levels and clock frequencies. This is followed by a section describing the capabilities of existing cooling technologies. These two sets of data are then combined to create a high-level picture of the development needs in area of ATE cooling systems.

3.3.1.1 Instrument Power Trends

Modeling power consumption and heat generation in ATE equipment is a complex problem because there are multiple subsystems and countless design trade-offs that must be considered. The goal of this analysis is not to develop a highly accurate tool that fully describes all of the issues. Rather, the goal is to develop a first-order understanding of the

phenomenon that drive power consumption in test instrumentation and use this to make an estimate of future needs.

There are several primary sources of power consumption and heat dissipation in a test cells, such as device drivers, processors and power supplies. However, the power consumption of these devices are all determined by the design implementation and have little to do with the device under test (DUT), and therefore it is difficult to relate test cell power dissipation and cooling requirements directly to the device parameters. Functionally, however, there are only three heat dissipating tasks that are fundamentally required to successfully test the device:

- driving signals to the DUT and measuring the response
- power supply for the DUT
- thermal conditioning of the DUT (hot chuck and thermal soak)

With the exception of some support systems, all components within a test cell should functionally fall into one of these three categories. Therefore, if the power consumption trends of these functions are understood, some insight can be gained into the power trends of the entire system.

The fundamental amount of power required for a CMOS transistor to drive a digital signal to a DUT can be calculated using the equation below:

$$Power = n \cdot C \cdot V^2_{dd} \cdot f$$

Where:

n	=	number of digital channels
C	=	capacitance of device and line
V_{dd}	=	supply voltage
f	=	frequency

All of these parameters except capacitance are tracked in the SIA roadmap. Chapters 4 and 5 discuss the trends and effects of device capacitance in detail. For the purposes of this analysis, it can be assumed that capacitance will fall from a current value of 100 picoFarads to approximately 25 pF in 2012. Using this data, instrument power consumption trends can be forecasted directly from SIA roadmapped device paramters.

The power consumed by the DUT is explicitly tracked in the roadmap and therefore no analysis of this aspect is necessary. Power consumption and heat generation of thermal conditioning devices should not change significantly for future devices and in addition to

this, these systems are generally insulated from the instrumentation in such a way as to not influence the thermal performance of the tester. For these reasons, it is valid to ignore their effect, and this analysis will focus on the power consumption of the DUT and drivers. Figure 3.3 shows the power consumption trends of these two parameters.

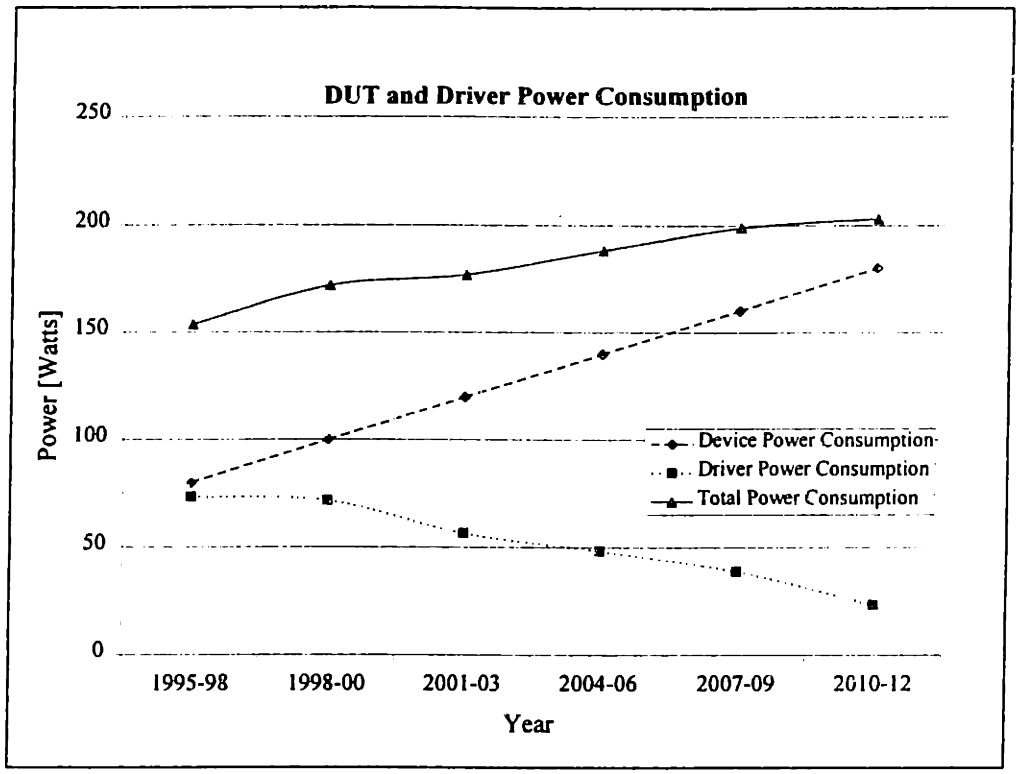


Figure 3.3: Predicted Power Consumption of the DUT and Instrument Driver

It is interesting to note that the theoretical power required to stimulate the DUT is forecasted to decrease over time. This is due to the V^2 term, which will decrease more rapidly than the linearly increasing clock speeds. At first, this seems counter-intuitive, but it does make sense when one considers that the same technology used to decrease DUT power consumption, i.e. smaller feature sizes and lower voltages, will also be used to decrease ATE instrumentation power consumption.

The graph above does not yet represent a completely accurate view of the test cell power consumption. In reality, the bulk of power in test systems is not consumed by the DUT and driver but instead by support electronics, such as formatters, comparators, pattern and signal generators, processors and power supplies. However, the power consumption and efficiency of these electronics should roughly scale with the DUT and driver instrumentation.

Another important consideration when evaluating power and cooling is the density of the electronics and heat generation. Obviously, dissipating 100 Watts in 0.1 cubic centimeter (cc) is considerably more difficult than in 1.0 cc. Despite the increasing pin count and functionality of test systems, there is continuous pressure to reduce the total system volume to minimize cleanroom floor space. One method of accomplishing this is to move instrumentation from the mainframe into the test head. This increases the thermal density of the head, but reduces the overall power consumption through the elimination the drivers required to send the signals over the length of the cable. The data clearly shows that test head volume is increasing while mainframe volume is decreasing. Overall, power density increased 4x over and 8x increase in pin count. Figure 3.4 shows some historical data that supports this trend.

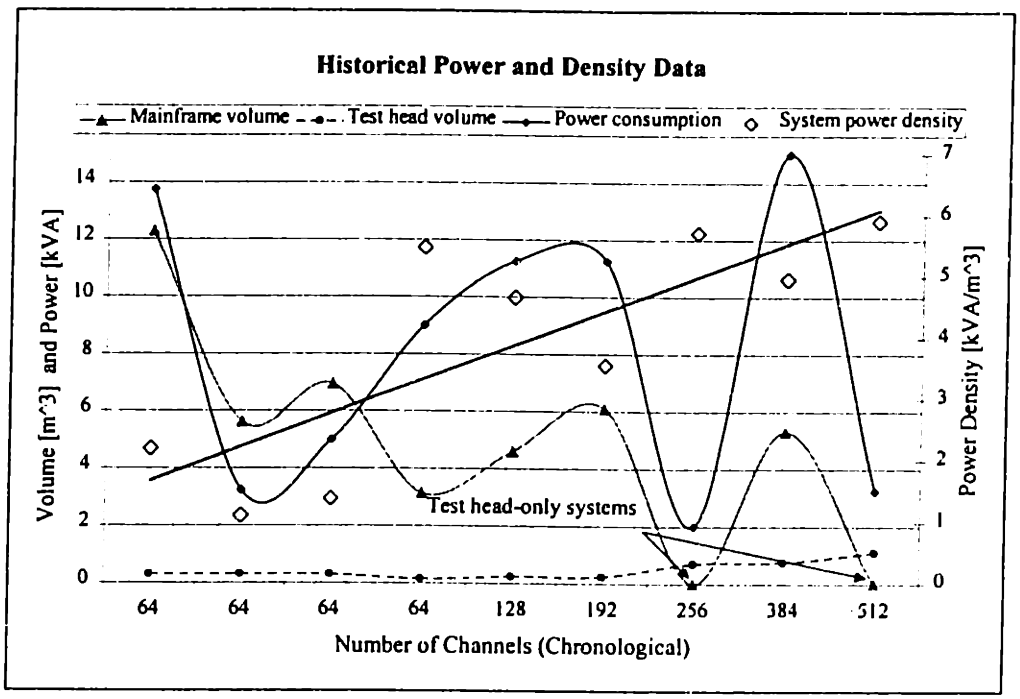


Figure 3.4: Historical Power Consumption and Density Trends

3.3.1.2 Cooling System Capabilities

There are a number of cooling technologies available and in use in ATE systems. Generally, cost and complexity scale with cooling capacity and therefore the simplest system appropriate for the task is chosen. Historically, natural and forced-air convective systems

have been sufficient to cool test electronics. However, denser and more powerful test systems are exceeding the capabilities of air cooling and several newer ATE systems have already switched to liquid-cooled test heads and mainframes.

Because cooling plays such a significant role in the performance and reliability of test cell design, it is important to understand the limits of current and future technologies. Table 3.1 below lists the most common cooling methods, their performance range in heat transfer coefficient and power density as well as an estimate of the relative costs of these systems.

Method	Heat-Transfer Coefficient [W/m^2]	Maximum Power Density [kW/m^3]	Relative Cost (Nat. Conv. = 1)
Natural convection (air)	1-10	0.3	1
Forced air convection	10-300	10	2
Conducting cold plate	100-1000	30	8
Circulating cold plate	500-5000	150	12
Immersion	10000	300	20
Boiling water, LN2	25000+	1000	?

Table 3.1: Common Cooling Technologies

3.3.2 Results: Thermal Roadmap

The roadmap and cooling capacity data discussed above can be combined to get a clearer picture of the future needs of ATE thermal systems design.

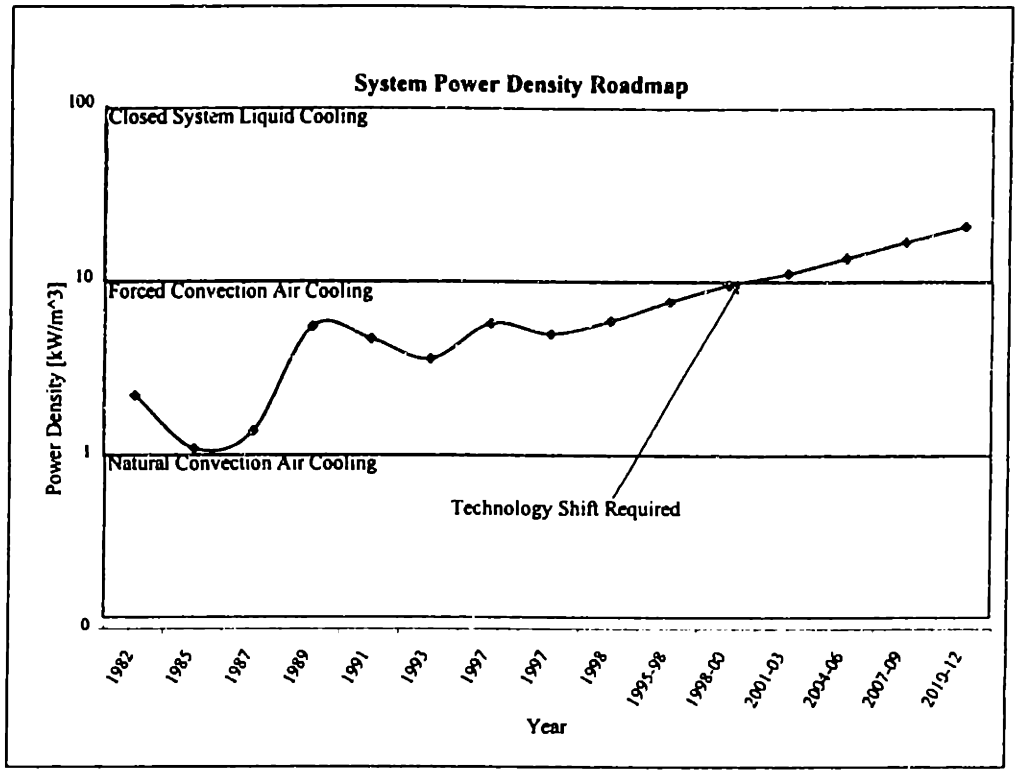


Figure 3.5: Cooling Roadmap

Figure 3.5 plots the roadmapped cooling requirements along with the ranges of cooling technologies. It is clear from this data that ATE systems have crossed the threshold between air and liquid cooling systems. If trends of increasing system power dissipation and decreasing system footprint continue, liquid cooling will be required on all future systems. However, this data also suggests that existing, commercially available convective and conductive liquid cooling systems have sufficient capacity to support ATE cooling needs well into the next century. While additional effort and focus is needed to understand the design, implementation and limitations of liquid cooling systems, research into new, uncommercialized cooling technologies probably does not present a significant opportunity for performance gains.

3.4 Mechanical Performance Roadmap

There are several mechanical subsystems within a test cell that are required to successfully test a device, but from the perspective of the DUT, there are two that are of primary importance. The first is responsible for delivering the DUT to the test electronics, and is generally the core motion system of the prober or handler. The second is responsible

for delivering the signal to the DUT and includes the interface, docking systems and electrical fixturing. This section analyzes the critical aspects of these two sets of mechanics and relates them to the device parameters tracked by the roadmap presented in Chapter 2. These relationships will then be used to predict the future mechanical requirements of these critical interface components.

3.4.1 DUT Delivery Mechanics

The sole task of DUT delivery mechanics is to present the device to the test fixturing. This needs to be done accurately so that good electrical connection is made between the test fixture on the tester and contacts on the device, and quickly so that the tester does not sit idle waiting for parts to test. There are two primary forms of these mechanics; wafer probe stages and package handlers. In general, the placement accuracy required for wafer probing is several times greater than that of package handling. This is because the package provides a mechanical fan-out of signals from the fine pitch of the wafer bond pads as well as tightly controlled reference surfaces on the package substrate to which the leads can be aligned to the tooling. These features, shown in Figure 3.6, require that the package only be nominally aligned to the fixturing before being blindly inserted.

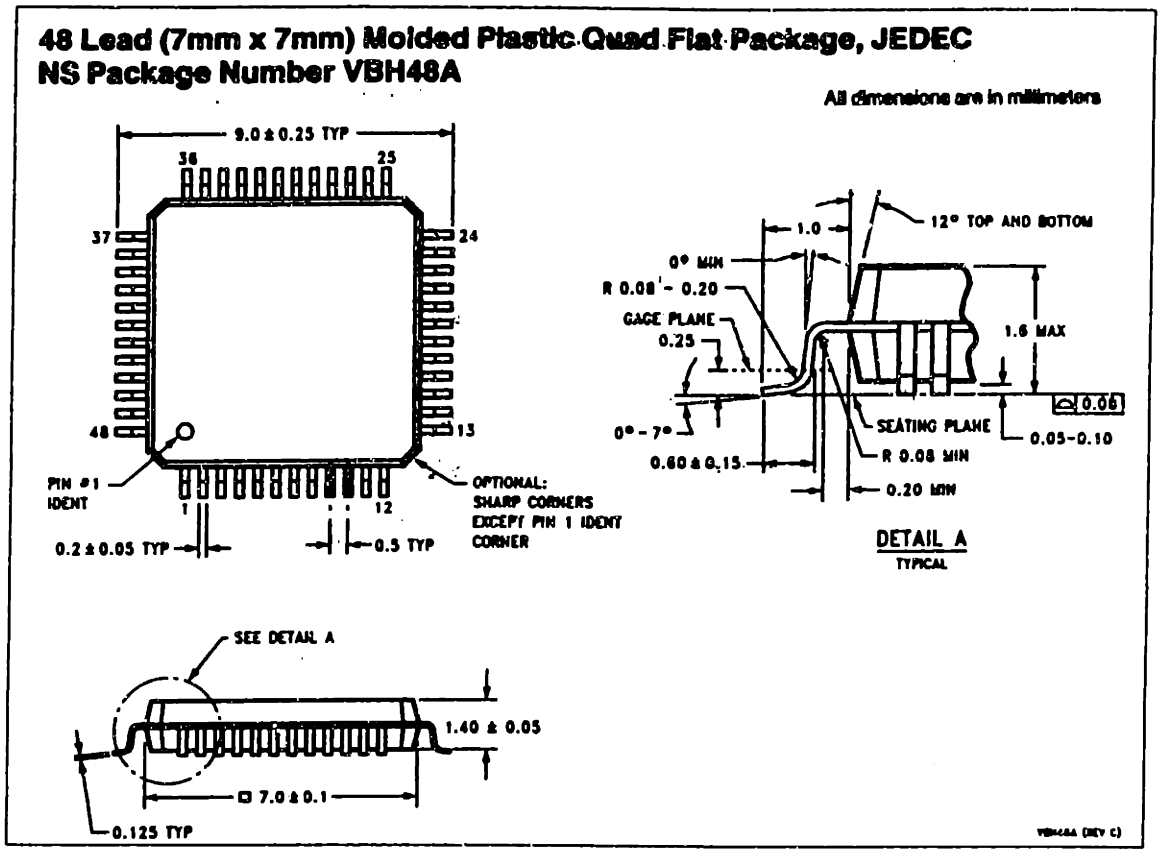


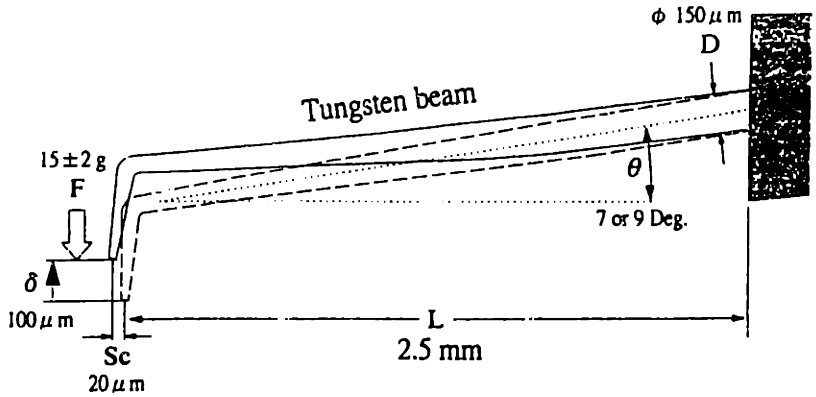
Figure 3.6: QFP Package Specification

There are no such mechanical alignment features on a bare die and therefore alignment between the bond pads and test fixture must be achieved through other means. Typically, this is achieved using optical alignment systems that are capable of resolving the position of the die and that of the test fixture. These operations are generally done off-line and a coordinate transformation is performed to calculate the optimal position of the die relative to the tooling.

The first-order analysis of the positional requirements of DUT delivery mechanics developed below focuses on the operational accuracy required align and contact a bare die within a wafer probe environment. This limited scope analysis is representative of future requirements of all DUT delivery systems for several reasons. First, as pin count and package density increase, the difficulty of testing packages will soon approach that of probing wafers, and therefore the accuracy requirements of device handling are a subset of those of wafer probing. This trend is reinforced by the growing demand for known good die (KGD) and chip scale packages (CSP) that require the positioning of singulated die with the

mechanics of a handler. Ultimately, the fundamental limit on packages will not be imposed by wafer or package manufacturers, instead it will be bound by the placement and soldering mask accuracy of PCBs and surface mount assembly machines. Second, the use of optics to close the positional loop virtually eliminates any static positional errors introduced into the system via interface misalignment. Therefore, the primary concern for mechanical interface development is the closed-loop accuracy of a wafer or package positioning stage.

The positional accuracy required to successfully probe a die is a function of the die geometry, pad pitch, test fixture accuracy and, for parallel test, the arrangement of die in the array. Figure 3.7 shows typical values for probe card tolerances, Figure 3.8 shows a photo of the probe needles and scrub marks on a die bond pad, and Figure 3.9 shows the geometry used to perform this analysis.



Standard Probe Diameter	0.010"
Standard Tip Diameter	0.0015"
Standard Tip Length	0.007"/0.015"
Standard Tip Tolerance	+ 0.003" / - 0.001"
Standard Planarity **- up to 1.0"** **- 1.0" to 2.0"** **- 2.0" +**	 ≤ 0.0005" ≤ 0.0007" ≤ 0.0010"
Standard Alignment	Scrub Center of Pad +/- 0.005"
Standard Probe Depth	0.080 + 0.010"/- 0.005"
Standard Contact Force	0.8-2.5 gm/mil
Scrub Action (nominal)	0.0001" – 0.00015" per mil of over-travel
Probe Materials	Tungsten, Rhenium Tungsten, Beryllium Copper, Palladium

Figure 3.7: Probe Needle Geometry and Specifications

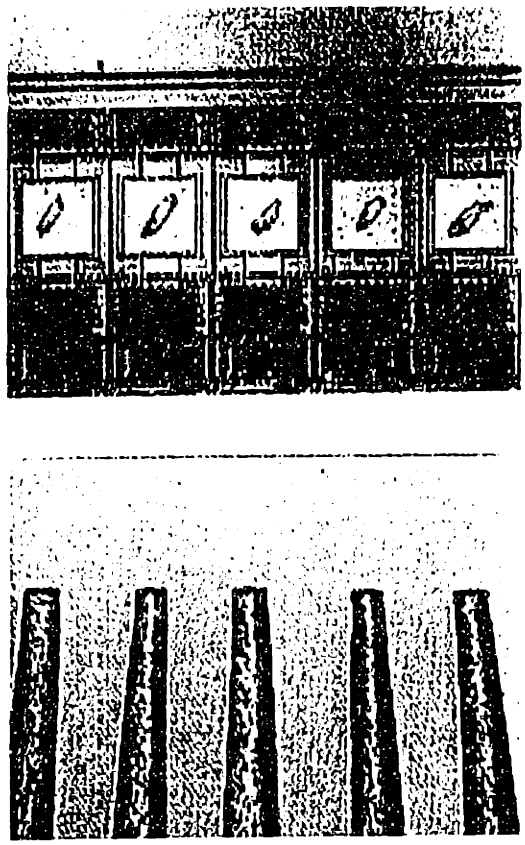


Figure 3.8: Probe Needles and Die Bond Pad Scrub Marks

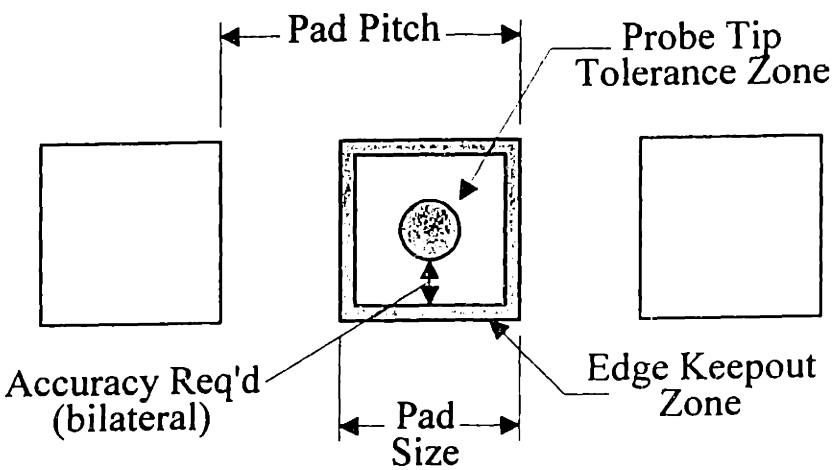


Figure 3.9: Probe Pad Geometry

The required accuracy can be approximated by simply summing the bond pad and probe needle tolerances. The future accuracy requirements of test systems can be assessed by inserting the device geometry, pin count and pad pitch data from the SIA roadmap, into this

model. The results, shown in Figure 3.10, suggest that future DUT delivery systems will require at least 5 microns of positioning resolution for single-site testing and better than 2 microns for highly parallel testing.

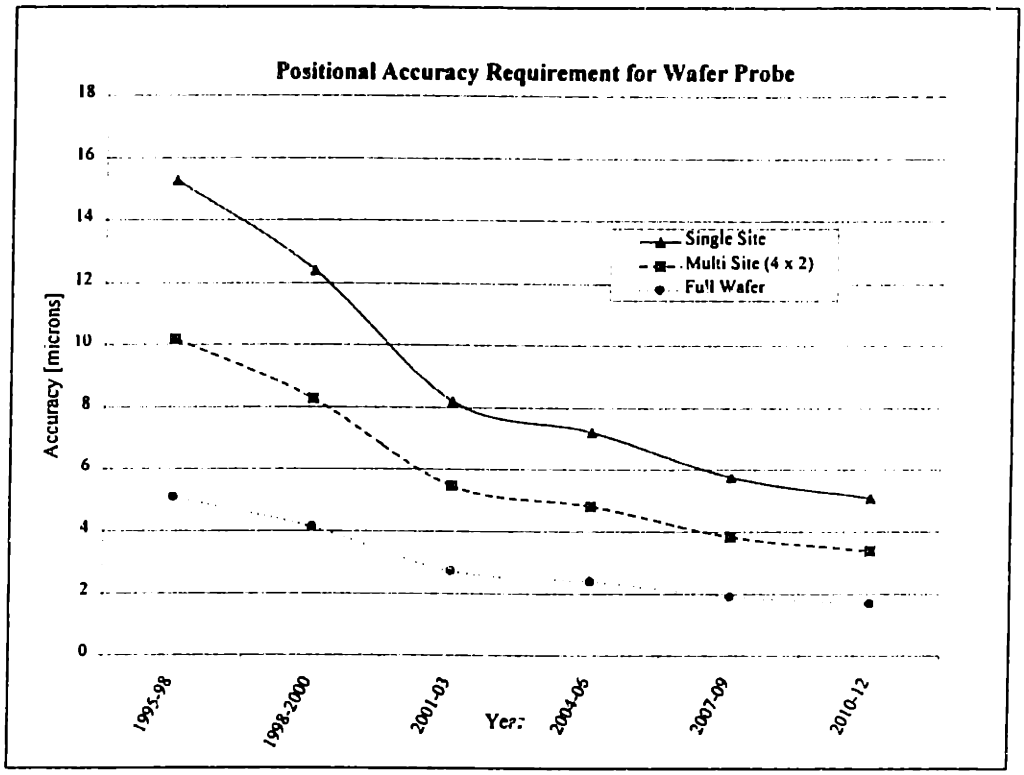


Figure 3.10: Device Positional Accuracy Requirements

Current package handling systems that rely on molded package features for alignment do not possess sufficient accuracy to meet these requirements. This is fundamentally a limitation in the package manufacturing process since it is unlikely that any molding or resin casting process will be capable of holding micron tolerances across a 2-5 cm package. A more likely solution for both package handling and wafer probing is the use of optics to measure the location of the bond pads, solder balls or leads and fine-align prior to contact. This method is currently used in most probe environments and is gaining acceptance in advanced handler designs.

Once the position of the device has been measured, delivering it to the contactor or probe needles with sufficient accuracy is not likely to be a technological hurdle. Current wafer probers have stage resolutions of 1-2 microns and wafer steppers routinely position to better than 0.1 micron. The greater challenges will be cost and throughput. Larger devices, larger wafer sizes and multiple packages tested in parallel require greater insertion forces and

larger handling stages. These larger, more massive stages take more power to move, require longer settling times and will result in lower throughput.

3.4.2 Signal Delivery Mechanics

The second half of the test cell mechanics are those components that deliver the test electronics and test signals to the DUT. There are two primary challenges in design of these systems; supporting the growing electrical contact forces and packaging and routing the high density of signals between the instrument and the DUT.

The contact force required to make an electrical connection varies significantly with the connector design and the current flowing through the connection. Connection forces for typical ATE signal transmission fixturing components range from 0.5 to 100 grams per pin and the force required to interconnect a typical high density fixture can often exceed 100 kgf. This interconnect force must be supported by the fixturing structural components and thus often competes with the electrical requirements for space between the pin electronics and the DUT.

There are three device trends that are driving increased interconnect forces; increasing pincounts, increasing power consumption and decreasing voltages. The relationship between pincounts and force is fairly simple, each additional device pin requires approximately 2-3 additional interface pins. The additional pins are usually required for shielding and differential pair measurements. Decreasing device voltages, both for power and signals, coupled with increasing power consumption, result in higher current carrying requirements. These high currents require more pins and lower interconnect resistance.

Theoretical work exists that describe the relationships between contact force, material properties and connection integrity[18]. For example, the contact resistance of an interconnect, and thus the current carrying capability, is directly proportional to the surface area of contact. This area is in turn related to the contact force as described by the Hertzian contact force equations[19]. While it is possible to use these relationships to calculate the minimal contact force per pin relationship, this is generally not practical because other factors such as manufacturing tolerances, contact cleanliness and oxide formation remain un-modeled. Most

[18] Clayton, 1982
[19] Timoshenko, 1959

contact designs use a spring element to compensate for manufacturing tolerances and a scrubbing action between the two contacts to compensate for dirt and oxide. These spring and scrub features, and not the minimum Hertzian contact force, are what generally dominate the interconnect mating force. Therefore, empirical data from existing industry solutions can instead be used to approximate the minimum obtainable interconnect compression force.

There are a number general types of interconnect designs, each has unique benefits and preferred applications. These most common of these are as follows:

- Pin and edge connectors: these are used primarily in cabled interfaces and backplanes. They are bulky and can have insertion forces in excess of 100 gm/pin.
- Spring pins: These are the standard interconnect method between the pin electronics and fixtures. Pin compression forces range from 20 to 60 gm per pin.
- Cantilevered beam: These are used to probe wafers and in package device test sockets but may be difficult to use in other areas of the interface. Bending forces can be as low as 1 gmf/pin.

A number of other interconnect technologies are available and are used in limited applications. These include designs that use flexible membranes, buckling beams or wire-bonded 'micro springs' to develop the contact force. All of these are either cost prohibitive or immature technologies and none currently offer a complete interconnect solution. There are also some innovative zero-insertion force interconnect designs, but because these require bulky mechanics, they are not capable of obtaining the pin density required for ATE fixturing.

Because the actual force generated by an individual contact is a function of the amount of spring compression, the contact force depends on the accuracy of the mating contact surfaces. However, at least one of the surfaces is generally a PCB material such as FR4, which generally has a very low elastic modulus (2.5 Mpsi). This creates an increasingly difficult problem; high pin counts generate high forces that deform the PCB, which in turn requires additional range of spring compression to compensate for these errors and thus even higher forces are generated. This cycle is generally broken by adding stiffening members to the interface circuit boards, but this also has limiting returns due to the volume of space they consume.

Current designs primarily use spring pins to interconnect fixture elements and more fragile, lower force, cantilever beams to make contact with the DUT. Therefore, the highest

compression forces and deflections are at the fixture interface. The primary impediment to decreasing the interface compression force much below the current 50 gm/pin is the fragility of the components and the positional accuracy of the interface.

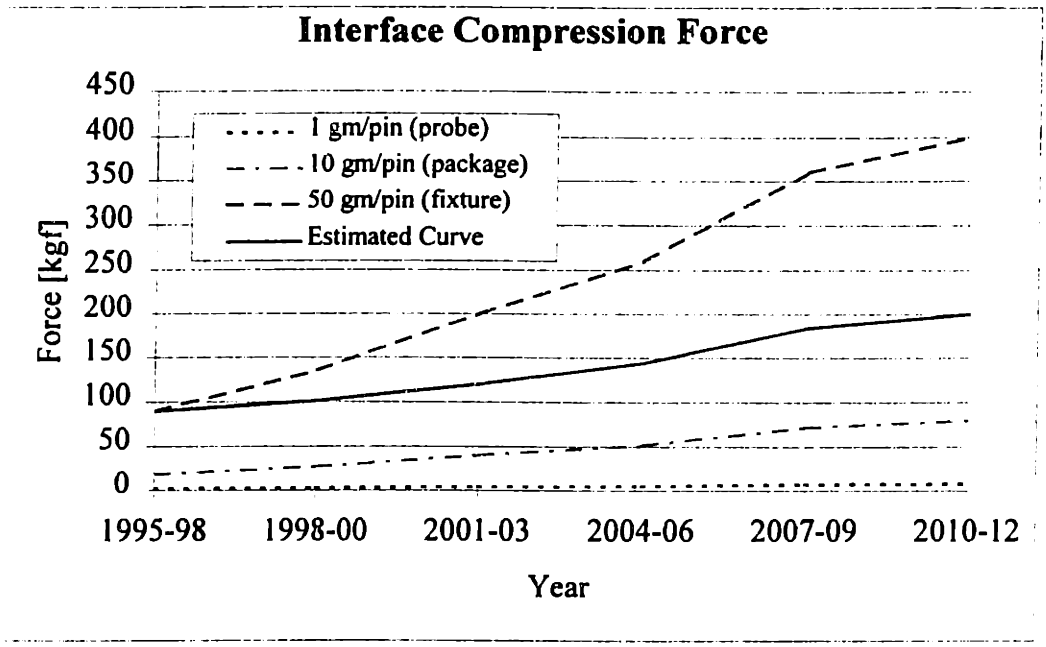


Figure 3.11: Interface Compression Force Predictions

Figure 3.11 shows the range of interface interconnection force that will be required for future interfaces based on the range of possible per/pin compression force. The worst-case is a continuation of existing spring pin designs, which will result in forces of up to 400 kgf. A more likely scenario is a gradual reduction in per/pin forces as pin counts increase, which will be achieved through the application of more fragile DUT interconnect technologies and at greater mechanical cost. Increased interface and PCB stiffening structure will be required as this force approaches its upper limit of 200-250 kgf. This is clearly approaching the limits of the existing interface designs, which are required to support this load over a 300mm to 500mm PCB, and is typically only edge supported.

A second critical area of signal deliver mechanics concerns the routing of signals from the array pin electronics to the contacts on the DUT. These signals are typically routed on a PCB traces and the drastic difference in signal densities presents a 'fan-out', or signal escape problem near the DUT, as shown in Figure 3.12.

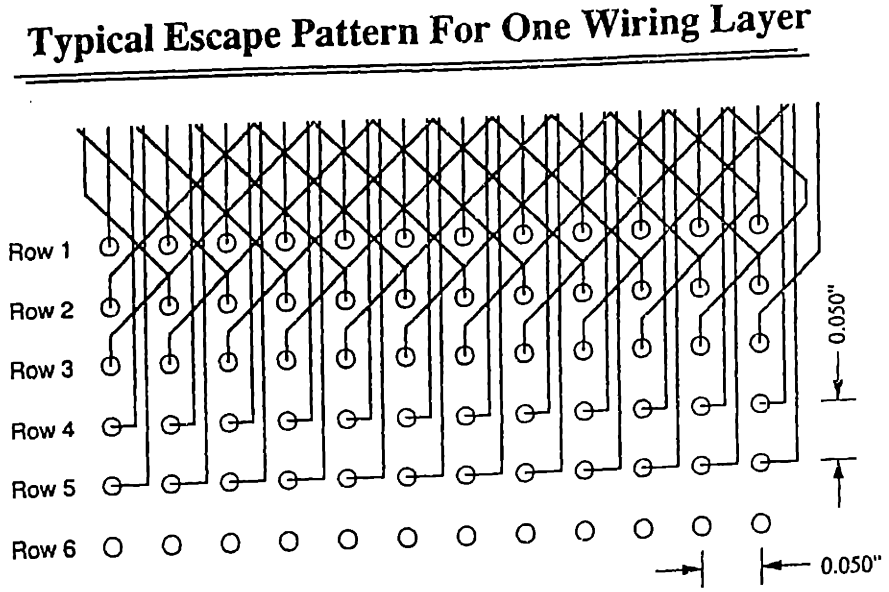


Figure 3.12: Signal Fan-Out

This problem is not unique to ATE fixturing and exists with any high density, grid-array packages or high density connector. However, there are some differences that make the ATE environment particularly challenging. ATE instrument pins are generally distributed in a fixed pattern about the device, regardless of the position of the DUT I/O's. For example, a mixed signal tester might have all digital channels on the right half of the tester and all analog on the left half. This requires additional routing length and signal layers to distribute the test resources to the proper position on the DUT, which may have the pins distributed in a different pattern. Also, some tests, such as high speed digital 'fly-by' testing, require that multiple test pins be connected directly to a single device pin, thus further increasing the density of pins near the DUT. Tighter PCB fabrication accuracy and higher number of board signal layers are currently the only ways to address this issue on current test applications, and both of these approaches are limited by PCB fabrication processes and both significantly increase fixture cost.

There are a number of approaches to analyzing the fan-out problem.[20,21] Figure 3.13 shows one method of relating the trace width and board layers to the number of device pins.

[20] Harper, 1991
[21] Buck, 1997

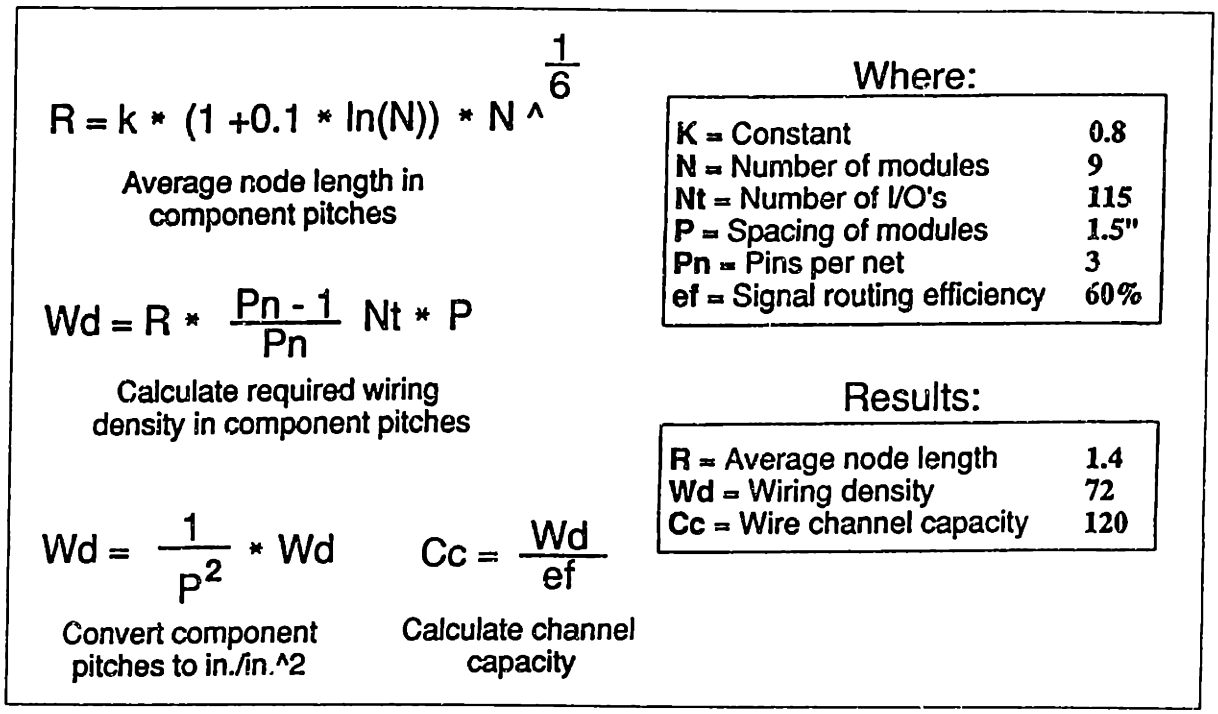


Figure 3.13: Method for Determining Fan-Out Density

A typical fixture may have a single or multiple modules (DUTs), but in the simplest case of a single DUT, the wiring density scales linearly with the number of device pins. This relationship can be combined with the SIA device I/O roadmap data to project the PCB fixturing needs for future test systems. Even for a single site test fixture, the above model shows that future PCB test fixture requirements predicted by the SIA will soon exceed existing PCB process capabilities. Since PCB process capability is driven by the same devices that are driving ATE design, it is likely that PCB technology will continue to improve. However, multi-layer PCB fabrication may be approaching the limit of the manufacturing tools, and barring any significant advances, may not be able to keep pace with the test fixture density requirements. There are a number of possible design countermeasures to this possibility. The expected solution offered by the SIA is an increased reliance on BIST techniques. Moving some test functionality on to the DUT will reduce the need for external test pins, but this approach also reduces the useful surface area of the device, adding cost with no added value. BIST is not an ideal solution and will only be adopted if no other alternative exists. Another approach may be to increase the layout flexibility of the tester pin electronics. For example, simply allowing the digital and analog instrument locations to be

interchanged will add a degree of freedom when optimizing signal routings. A third approach might be to blur the line between pin electronics and fixture by moving some of the test resources on to the fixture. For example, typical instruments have ASICs that are specifically designed for test applications. These devices contain the final stages of the DUT drivers and the first stages of the response measurement circuits. Multiple copies of these functions can contained in a high density package and placed on the fixture, very close to the DUT. Moving the first stages of measurement very close to the DUT reduces performance requirements on the remaining signal lines, allowing them to be packed more efficiently. Also, data leaving the ASICs can be multiplexed, further reducing the signal density. This 'tester on a chip' approach also changes the economics of the test industry. Moving the custom ASICs out of the tester and on to the fixture lowers the capitol cost of the system, but also adds intellectual property and consumables to the test cell.

3.5 Cost Roadmap

Cost of test can constitute a significant percentage of a semiconductor product's production and capital equipment cost. Test cell system cost is often the last consideration of the designer and usually one of the first concerns of the user. Understanding the primary drivers of system cost is important because it permits informed trade-offs between cost and performance, but modeling the costs of test cell mechanics is a difficult and imprecise task. This is because historically, the bulk of ATE cost has been driven by the cost of instrumentation, not the cost of mechanics. In addition to this, true cost data is typically proprietary and therefore only approximations based on selling price are available. However, this is still a worthwhile exercise because the increasing mechanical content of test cells and the decreasing per pin cost requirements present the opportunity to redefine the cost paradigms of test cells.

The SIA roadmap forecasts the need for a continuous reduction in tester costs over the next 15 years. This includes a reduction in per pin cost from $3.3K today to $0.4K and a system cost reduction from $1.8M to less than $1M by 2015. This forecast does not include the cost of the wafer or device handling mechanics and the shift from 200mm to 300mm and 400mm wafer sizes as well as the continuous reduction in package and bond pad feature sizes will drive up the content and cost of these mechanics. Therefore, it is not likely that the total cost of a test cell will follow the SIA predictions. Instead, mechanical cost growth will tend

to outpace instrumentation cost reductions, resulting in increasing system costs overall. Figure 3.14 shows historical cost data for wafer probe and device handling mechanics that supports this claim.

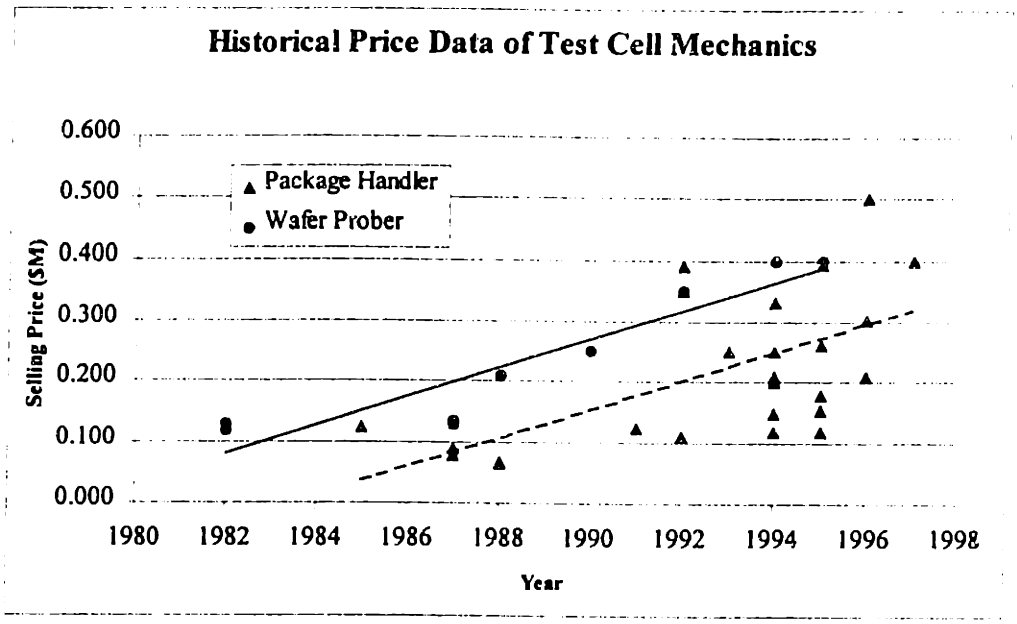


Figure 3.14: Historical Test Cell Mechanical Cost Data[22]

Historically, test cell costs have been dominated by the test electronics. In 1985, a tester cost approximately $1.5M and the wafer prober and additional docking mechanics totaled less than $200K. Between 1985 and 1995, semiconductor fabrication processes shifted from 2 micron to 0.5 micron and wafer sizes doubled from four to eight inches. While speed and pin counts increased and the performance of tester instrumentation incrementally improved, tester costs have remained relatively flat over this time frame. During this same time period, test cell mechanics have rapidly progressed, in part through the addition of new features that improve the test cell productivity such as multiple optics systems for automatic alignment, automated wafer and fixture handling, yield tracking software and environmental management systems. As a result, the cost of test cell mechanics has more than doubled and thus the cost of these mechanics as a percentage of total test cell costs has increased significantly.

[22] Source: Dataquest

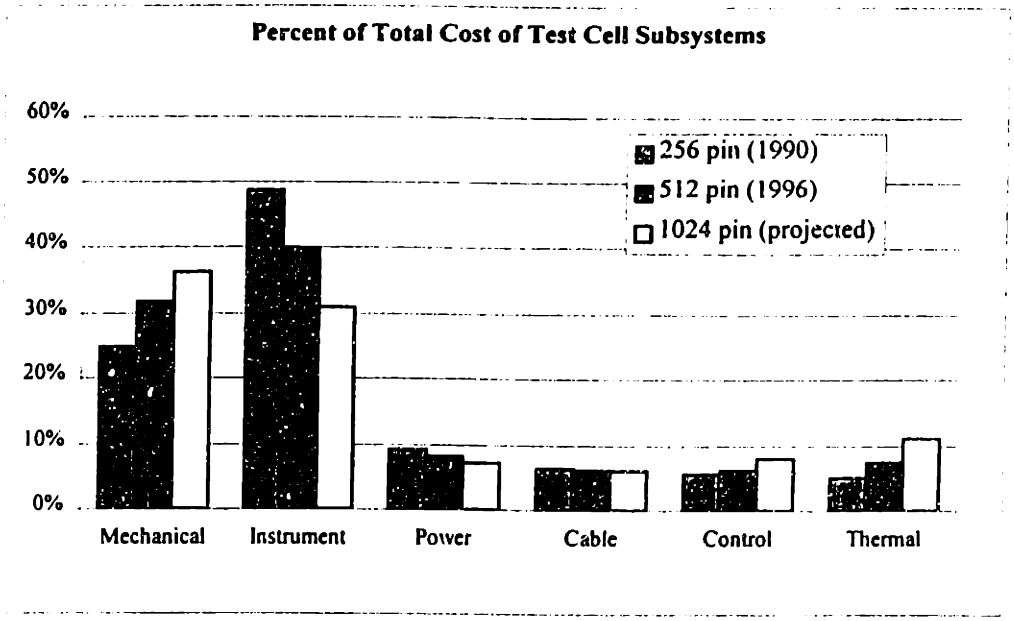


Figure 3.15: Test Cell Cost Distribution

Figure 3.15 shows relative cost data from two representative test cells as well as the predicted costs of a next generation system based on some of the data presented in the sections above. While this data supports the assertion than test cell mechanical costs are increasing, this trend cannot continue unchecked because despite the increasing need for mechanical performance, the core feature set of a test cell will continue to be the pin electronics. Therefore, strategies must be developed to curtail the expanding cost of mechanics and instrumentation without reducing the gains in cell performance or productivity.

Reducing costs while increasing functionality has been the credo of the semiconductor industry and is a direct result of Moore's Law. However, this is much more difficult for mechanical components because, unlike integrated circuits, the material and manufacturing costs of this mature industry tend to increase over time.

Some of the cost reduction in instrumentation may be enabled by higher device-level integration and through increased adoption of BIST techniques[23]. However, to achieve significant gains in cost performance, new architectural approaches are required that radically change the mechanical cost equation. Some lessons can be learned from the cost reduction progress of semiconductors, specifically, that higher levels of modularity and integration

result in lower system costs. While these two approaches, increased modularity and integration, appear to be in conflict, when properly implemented, they both have the potential to significantly reduce system costs.

3.5.1 Test Cell Integration

One approach to reducing test cell mechanical systems cost is to reduce or eliminate components by integrating the electronics and mechanical systems. The historical development of test cells that produced separate pin electronics and device handling mechanics has also built in many unnecessary mechanical systems redundancies and non-optimal cell configurations. A high level functional analysis of a test cell requirements can be used to separate the components that are fundamentally required to test a device from those that exist only to support the development legacy. This short list of functional components can then be re-combined into a single system that contains only necessary components, thus minimizing cost. For example, the separation of pin electronics and handling mechanics implicitly requires a cable bundle and manipulator. A system architecture that integrates the electronics and mechanics in a single system eliminates these components with no loss in functionality. The data presented in Figure 3.14 above shows that cabling costs constitute approximately 7% of system cost, and integrating the testhead with the mainframe would eliminate this cost, along with an additional 5-10% in structural and manipulator costs, with no resulting loss in test functionality. There are many other examples of redundant systems, including metrology frames, power distribution, thermal systems and machine controllers.

There are several drawback of an integrated systems architecture[24]. Integrated systems architecture assumes that a single design is sufficient for all applications and is difficult to scale and therefore tight integration and sharing of resources limits the flexibility of the system to adapt to future requirements. In addition, the current, separable architecture adds a level of risk abatement in that the entire cell is not at risk if a single component fails. If a $200K handler breaks, it can quickly be replaced and the $1.5M tester will not sit idle. An

[23] This will effectively lower the tester pincount and performance requirements
[24] Cutherell, 1996

integrated system must inherently be more reliable because of the increased capitol and throughput costs that are at risk.

3.5.2 Test Cell Modularity

A more modular systems architecture addresses some of the drawbacks of an integrated system and may offer another path to cost reduction. The trade-offs between modular and integrated systems are generally well understood and documented. Modular systems are generally easier to design, service and upgrade but may sacrifice performance and production cost when compared to integrated systems. However, the complexity of a test cell and its diverse range of applications make it impossible to cost optimize all aspects of the system simultaneously, and therefore a modular approach, pursued within the context of a well defined system architecture, allows individual modules to be cost optimized with minimal impact on the performance of the entire system. Modules can be developed internally or subcontracted out to companies with specific expertise, further expanding the range of potential configurations. The well defined boundaries between modules permits multiple sources and approaches to intra-modular design problems without risk to the system performance.

Optimization of individual modules is unlikely to yield cost reductions large enough to significantly impact the total system economics. However, a modular architecture does provide a means for cost reduction in that it enables test systems to be built that more closely meet the test requirements. Current test cell architectures are designed for worst-case test scenarios which results in significant infrastructure costs for lower performance derivative products. For example, the mainframe cabinet, cooling, power supply and computing capacity of a 1000 pin digital mixed signal system is significantly larger than that required for a low digital pin-count DSP device tester. An integrated approach to these two systems would result in either a high zero-pin cost or the development of two separate designs. The high zero pin cost is difficult to justify for a low pin-count system, and the development of separate systems results in higher production costs due to lower volume and fewer systems over which the engineering costs can be distributed. A modular architecture addresses this issue by discretizing the system into functional elements, or modules. A scalable system can then be built by combining these modules a'la carte, allowing a system to be more closely

optimized to a specific application. Achieving this level of modularity requires that the infrastructure, interfaces and capacities of all components be well-defined prior to the development of any of the modules.

3.5.3 Comparison of Modular and Integrated Architectures

In the end, some combination of modular and integrated architecture will be needed to address the performance cost challenges of future test cell design. Integration of major test cell systems should be used to reduce or eliminate unnecessary or redundant components such as machine frames, thermal systems, cables and manipulators. Modular will be used to increase system application flexibility and provide a path for derivative configurations. The first step in this effort is to tabulate all of the necessary and desired functions of a test cell, as shown in Figure 3.16.

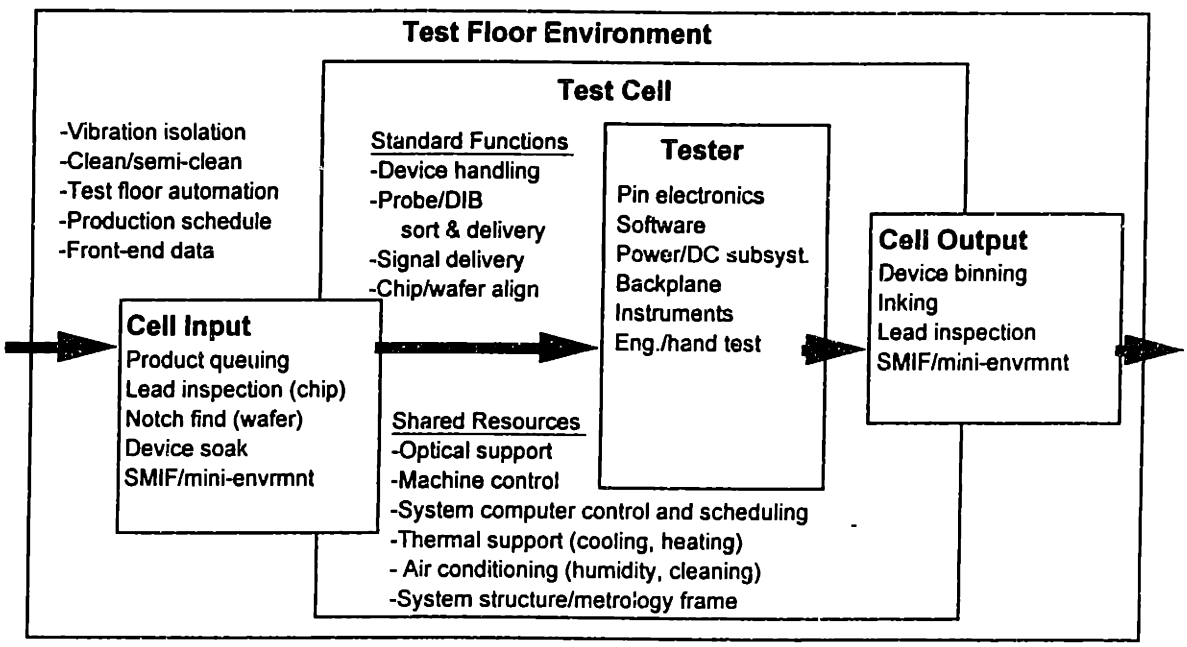


Figure 3.16: Test Cell Functions

Those functions that are required for all varieties of systems and can be shared are candidates for cost reduction through integration. Those that are optional, can be scaled, or have multiple design approaches should be considered as modular elements.

3.6 Summary of Chapter 3

This chapter discussed how the SIA roadmaps influence the design and evolution of test cell mechanics. These relationships were developed by correlating the device parameters to critical test cell attributes such as thermal performance, mechanical accuracy and stiffness and system cost. In particular, it finds that:

- Power densities of future test electronics will surpass the capability of forced ambient air cooling. More advanced and expensive technologies such as liquid and thermionic cooling need to be investigated and implemented.
- Mechanical positioning accuracy requirements of test cell DUT handling equipment and test fixturing will continue to increase. However, these requirements will never exceed those used in wafer fabrication and therefore this aspect of test cell development can follow the advances of front end equipment.
- The mechanics required to support fixture interconnect forces will continue to increase with pin counts. New interconnect technologies that are higher density and produce lower contact forces are required.
- Signal trace densities between the pin electronics and the DUT will soon outpace the process capabilities printed circuit boards. New approaches to the distribution of test resources and increased adaptation of BIST technologies will be required to test future, high pin-count devices.
- The mechanical functionality and content of test cells will continue to increase relative to that of the pin electronics, while the total test cell cost is forecasted to decrease. Test cells will need to change significantly to accommodate this cost pressure. New architectures that share common resources and reduce non-test related infrastructure costs need to be adopted.

Like the SIA roadmap, the analysis, forecasts and recommendations presented above represent a snapshot of a rapidly changing industry. This first attempt at roadmapping the mechanical needs of the ATE industry will require continuous review and revision by a much larger audience if it is to remain useful and up to date.

4 Signal Transmission Model Development

4.1 Introduction

The transmission of signals between the pin electronics and the device under test (DUT) is one of the crucial elements of a test cell and is becoming increasingly more difficult as device speeds increase. The design of these systems, however, is not a new science; the field of high frequency signal transmission has origins that predate the semiconductor industry. While Hertz and Marconi began experimentation with frequencies as high as one gigahertz in the late 1800's, the real advances came at the end of WWII with the intense work on radar and radio communication technology. Since this time, there has been continuous, incremental improvement in the design of high frequency communications as witnessed by the proliferation of pagers, cellular telephones and satellite TV. However, while many of these devices transmit digital information that is modulated in the gigahertz (GHz) range, they are fundamentally analog instruments that communicate with low frequency (<100 MHz) digital processors. Testing these devices is a relatively simple matter of supplying a few high frequency analog signals, which are then converted to a small number of low frequency digital signals. As cellular and wireless technology improve, analog speeds will increase and testing will be incrementally more difficult. However, these devices will remain low pin count, low speed digital device with a few analog lines and methods exist to test such devices.

A far greater test challenge will be imposed by the microprocessors and SOCs of the future. As digital clocks approach the GHz range and pin counts number in the several thousand, the difficulty of transmitting accurate, coherent signals between the device and test electronics will grow exponentially. This is a test challenge that will require significant advances in methods and technology to solve. The design of such a signal delivery system combines the fields of microwave signal transmission, electronic packaging, mechanical design and manufacturing. Because of the highly cross-functional nature of this field, it is not widely understood and few good design tools exist. The first step in addressing this void is to develop a model that capture the fundamental principles of these disciplines and the coupling between them. This chapter describes these principles, design goals, components and modeling methods. The chapter begins with a qualitative description of high-frequency

signal transmission theory and its application, including an overview of where it is used why it is important. Included in this section is an interpretation of the device roadmap that shows exactly why this area will become a major obstacle in the near future. This is followed by a discussion of the impact and constraints that transmission systems impose on ATE systems. The bulk of the chapter is devoted to the development of electro-mechanical signal transmission models. These models are capable of making fast, first-order approximations of the relationships between the mechanical design parameters of signal delivery systems and their corresponding performance. The assumptions and limitations of the model are also discussed and Chapter 5 uses these models to make some preliminary predictions on the future challenges of signal delivery systems based upon the trends of the SIA roadmap.

4.2 Signal Transmission Defined

Due to the variety of devices and applications, ATE instrumentation is designed to be generic and flexible in use and the application of this equipment to test a particular device often requires custom tooling, or fixturing. Because of this, the test environment is usually quite different from the design, or operating environment of the device. One of the most significant differences is in the details of how the signals are transmitted in to and out of the device. For example, in a real application such as a desktop computer, the microprocessor is in close proximity with the resources with which it must communicate. The external data bus (e.g. PCI, NuBus or EISA) is usually less than 8 inches away and memory modules are even closer, usually 1-2 inches. More often in high performance applications, some of the memory is cached on the same silicon as the microprocessor to further reduce this distance. Because of the flexible nature of tester design, this tight level of integration is nearly impossible to reproduce in the test environment using a generic tester. For example, to properly simulate the worst-case operating environment and make the required measurements, test instrumentation and fixturing is typically an order of magnitude larger in size than the final device application and the path lengths are generally an order of magnitude longer. Typical test and application environments are shown pictorially in Figure 4.1 below.

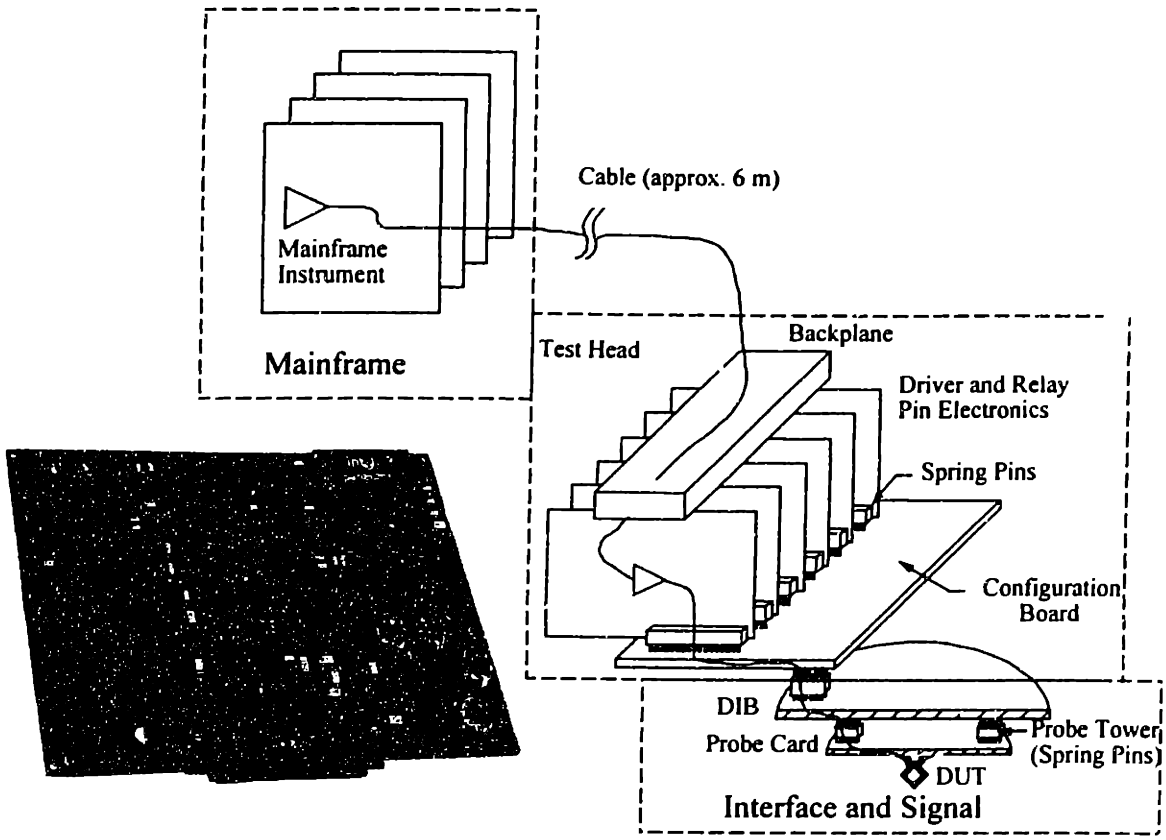


Figure 4.1: Application and Test Environments

In both of the above cases, the mechanical design of interconnects between devices is critical to the performance of the system. For the purposes of this analysis, a signal transmission system is therefore defined as the collection physical components (i.e. PCB traces, cables and connectors) used to deliver the signal from the tester to the device under test. Any difference between the test transmission system and that of the final application will introduce errors and uncertainty into the measurement and is therefore undesirable. At low frequencies, the physical implementation of this signal transmission line has little effect on the signal as long as the DC circuits are the same. However, as frequencies increase and the signal wavelength approaches the length of the line, and as device density increases and device power decreases, the design of the transmission system can have a first-order effect on the signal.

There are four primary causes of signal degradation in transmission lines as follows:

- Impedance mismatches at the interfaces of transmission components causes reflections at high frequencies.
- Normal printed circuit board and connector technologies do not have sufficient bandwidth to support high frequency signals.
- The DUT lacks sufficient power to drive signals over the lengths of these lines.
- At sufficiently high frequencies, the transport lag due to the line length causes timing and stability problems.

Each of the above phenomenon are described in greater detail below.

4.2.1 Impedance Matching of Transmission Lines

There is a fundamental aspect of a transmission line that distinguishes it from a simple DC circuit. Standard DC circuit analysis makes an implicit assumption that distances between nodes of a circuit are infinitesimal, or conversely, that the speed at which information is transmitted between nodes is infinite. This is equivalent to the 'quasi-static' assumption used when analyzing static mechanical systems and it is generally valid for low frequencies and short path lengths if one considers that signals propagate at nearly the speed of light. This assumption breaks down as the fundamental wavelength of the signal approaches the path length of the circuit. The exact wavelength of a signal is a function of both the frequency and the dielectric constant of the material that supports the electromagnetic fields.

$$\lambda = \frac{C}{f\sqrt{\varepsilon_r}}$$

λ = *wavelength*

C = *speed of light in vacuum (3 x 10^8 meters/second)*

f = *frequency[Hz]*

ε_r = *relative dielectric constant of medium*

Bus speeds on a typical PC run at about 60 MHz and common PCB materials such as FR4 have a relative dielectric coefficient of approximately 4.5. This yields a fundamental wavelength of approximately 2.3 meters. However, the signals on this bus are generally digital pulses, or square waves, which contain higher order harmonics. In digital testing, the edge rate, or rise time is the specification that determines the number of harmonics required to accurately transmit the digital pulse. Two representations of a square wave, shown in Figure 4.2 below, contain only odd harmonics whose significance decrease in a linear fashion.

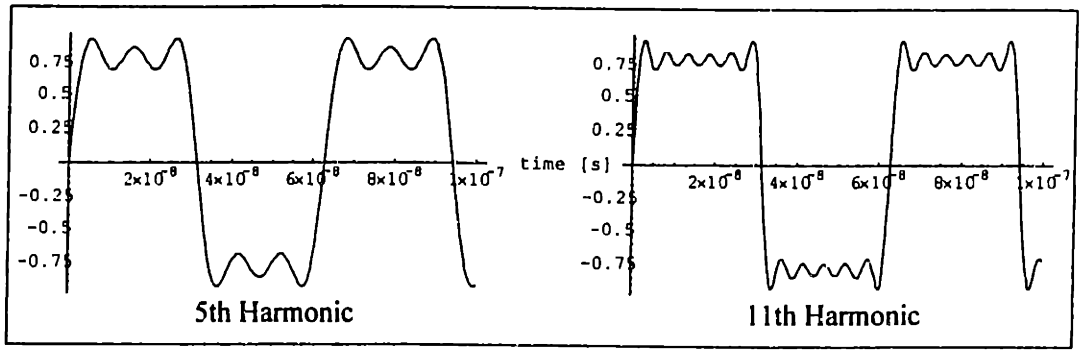


Figure 4.2: Square Wave Harmonics

Typically, low accuracy or slow edge rate tests require the 1st through 5th harmonics and high accuracy tests may up to the 11th harmonic. Therefore, if quasi-static circuit analysis is used, then the wavelength of the 5th or 11th harmonics must be much longer than the total path length. In the PC bus example above:

$\lambda [5^{th}] = 0.46$ *meters (2.3/5)*

$\lambda [11^{th}] = 0.21$ *meters (2.3/11)*

The distance from the processor to the slot on a PC bus is approximately 0.15 meters (6 inches). This suggests that 60 MegaHertz is close to the maximum frequency that will allow the quasi-static assumption. This is one reason that modern microprocessors operate at two different frequencies and use internal 'cache RAM'. The internal clock, currently running at approximately 250 MHz, is too fast to communicate with other devices on the long expansion bus. Therefore, a slower, external bus clock is run at an integer fraction of the internal clock. To speed access to memory, internal caches are built into the processor so that the path length to commonly used blocks of memory does not have to leave the microprocessor package.

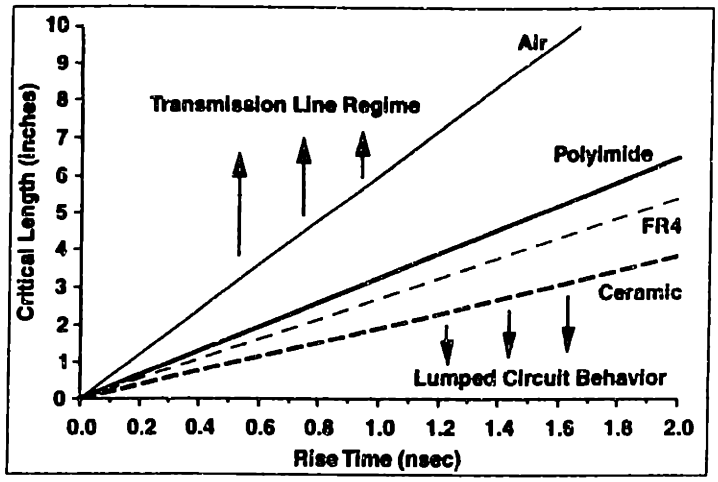


Figure 4.3: Transmission Line Defined

Figure 4.3 graphically shows the relationship between rise time, path length and dielectric constant. What happens in the frequency range beyond the 'quasi-static' region? The answer to this question lies in the finite speed of the waveform and the simple Ohm's Law relationship between voltage and current. To understand when a signal can be considered quasi-static, and when it must be modeled as dynamic signal, one must look at the inputs and outputs of the line on which it is travelling. Figure 4.4 below shows a 60 MHz (fundamental) and a 660 MHz (11^{th} harmonic) on a ½ meter line. For the lower frequency, it is clear that the line is short enough so that the input and output signals are always at essentially the digital same level. This is not the case for the higher frequency signal and the actual output is also a function of the line length and phase of the input signal. Therefore, a small change in length has a small effect on the 5^{th} harmonic wave, but may cause a swing from rail to rail in the 11^{th} harmonic wave.

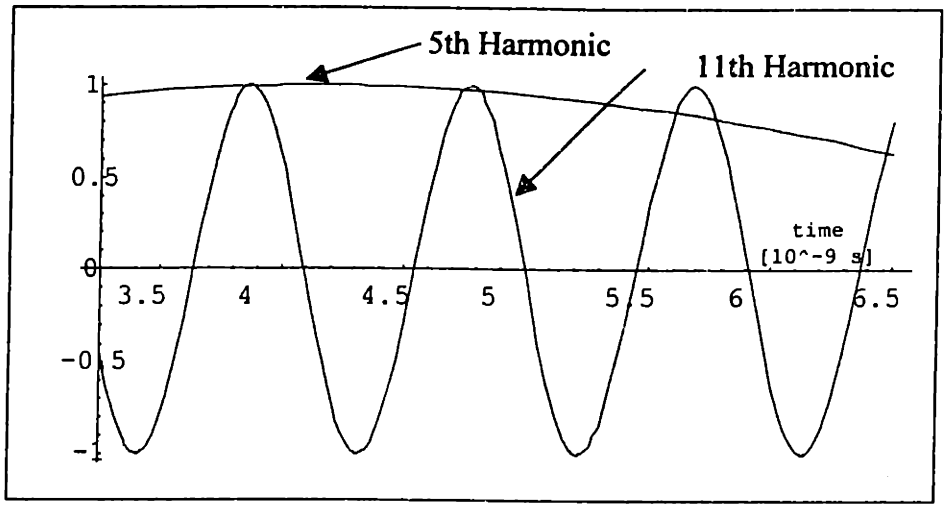


Figure 4.4: Relationship Between Length, Frequency and Phase

For PCB traces and coax cable, a common modelling assumption is that the line is lossless and has constant impedance along its length. Using Ohm's Law the current can be at any point can be calculated. For the low frequency signal, the current at any point along the line can be found by simply dividing the sum of the line and load impedances into the voltage. For the higher frequency case, the current varies with the voltage and therefore is a function of phase and length. Before the wavefront reaches the load resistance, it's current is dependent only on the line impedance. When the waveform reaches the load resistance, any mismatch between the line and load impedances will cause a discontinuity in current, resulting in a reflection of some of the wave. This is better illustrated by looking at the voltage and current propagation of a voltage step on an unterminated line as shown in Figure 4.5. In this case, the absence of a load resistance is the equivalent of an infinite impedance mismatch, resulting in a complete reflection of the pulse.

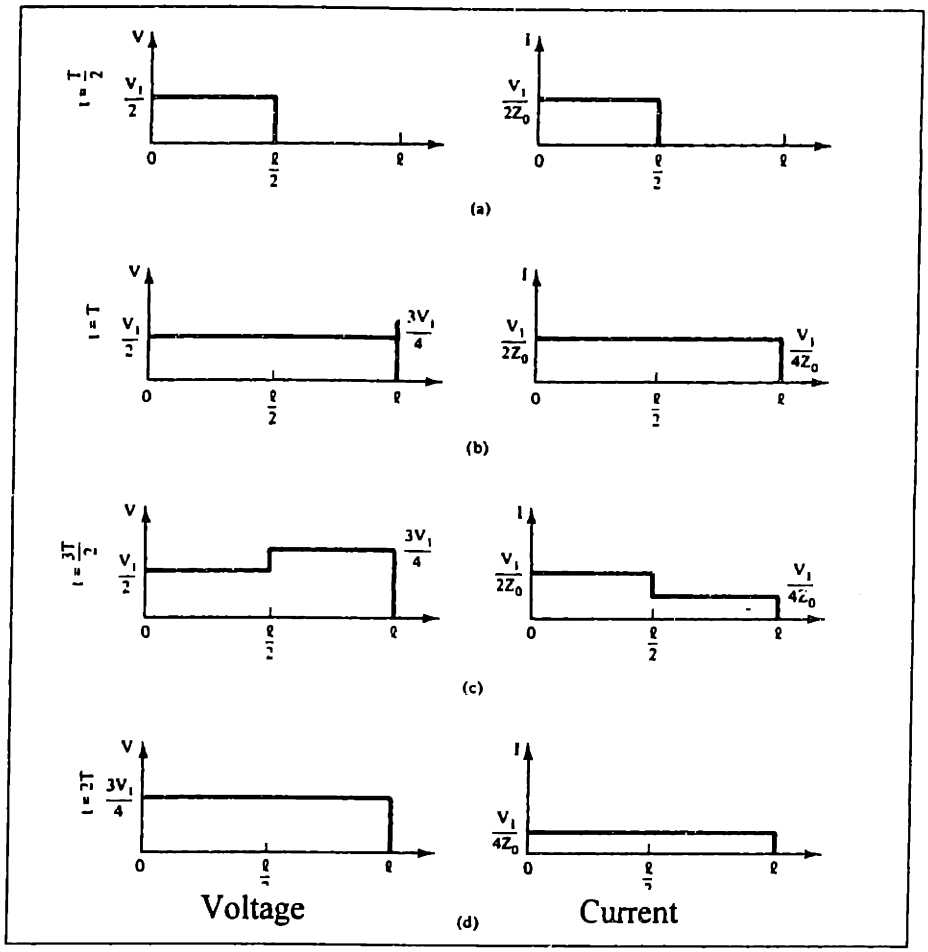


Figure 4.5: Propagation of a Pulse on an Unterminated Line

The effects of line length tuning are of critical importance in narrow-band frequency instrumentation such as microwave transmitters and the properties of ¼ and ½ wavelength lines are often used in controlling the behavior of these types of circuits. However, the high spectral density of digital signals reduces both the benefits and drawbacks of this phenomenon and in first-order analysis, this effect of line length can be ignored.

Therefore, at frequencies below the quasi-static threshold, phase information is unimportant and circuits can be modeled using standard analysis techniques. As frequencies approach and exceed this boundary, design and analysis of circuits requires the additional consideration of the impedance along the transmission line. A mismatch in impedance at a connector, for example, will result in the reflection of part of the signal. This reflection will result in a loss at the receiving end of the line and a spurious signal at the sending end. It

should be noted that reflections are a purely conservative phenomenon and, in first order analysis, no energy is dissipated.

4.2.2 Transmission Line Bandwidth

Other, non-conservative, higher-order effects serve to limit the bandwidth of transmission line components at high frequencies. They generally fall into one of the four following categories:

- Conductor loss
- Dielectric loss
- Leakage loss
- Radiation loss

The dominant mode of conductor loss is the pure ohmic resistance, or i^2R loss. Resistance of typical conductors such as copper, silver and gold range in the hundreds of nano-ohms and therefore this can generally be ignored for moderate length lines and moderate currents. However, at high frequencies, currents tend to flow near the surface of conductors, effectively reducing the cross-sectional area of the conductor and increasing the apparent resistance. This thinning effect is proportional to the square root of the frequency and becomes significant in signal level lines at gigahertz frequencies.

Dielectric and leakage losses cover the dissipation of energy in the dielectric material between the conductor and ground. At low frequencies or low voltages, the admittance of the material can be assumed to be zero. As voltages increase, the dielectric may breakdown and current flow between the conductor and ground shield will become significant, resulting in leakage. Dielectric loss, a more important phenomenon for signal level voltages, is due to dielectric hysteresis at high frequencies. This is essentially i^2R losses due to eddy currents in the dielectric material and becomes significant as the wavelength of the signal approaches the dielectric thickness. While the currents in both of these cases are always very small, the effect is cumulative along the length of the transmission line and can become significant for moderate lengths of line. In fact, this is the dominant loss in printed circuit board transmission lines and effectively limits bandwidth though conventional designs to approximately 3 gigahertz.

Radiation loss is another high frequency effect that is significant in poorly shielded transmission lines. Coaxial cable and rigid waveguides generally have insignificant radiation

loss due to their circumferential shielding, but other transmission structures, such as microstrip PCB traces and bare wires, can lose a significant amount of energy due to radiation. Generally, transmission lines are well shielded to prevent this phenomenon and to minimize the reverse effect of absorbing radiated noise from other lines.

4.2.3 Lumped and Distributed Capacitance

There are many types of semiconductor technologies used in making integrated circuits. One of the most common building blocks of conventional digital circuits is the bipolar transistor, which is characterized by slow edge rates and large quiescent power consumption. The growing popularity of portable, battery-powered devices has driven the demand for lower power technologies such as complimentary metal oxide semiconductors (CMOS) which is characterized by very low static power dissipation and 'stiff' (high impedance) outputs that swing the full supply range. One drawback of these devices is that they lack the ability to drive large currents at low voltage levels. This is generally not a requirement for a digital TTL circuit, but it becomes a problem in test fixturing due to the lumped capacitance of discrete elements such as contactors and vias and the inherent distributed capacitance of transmission lines. Therefore, the driver must be capable of providing sufficient current to charge the line capacitance at each clock cycle. If the path length is long, the capacitance may be sufficiently large so that there is no signal left at the end of the line. This fact is the primary explanation of the existence of test heads. Early ATE designs used to test bipolar devices used long cables between the instrument and DUT. The advent of CMOS required that the path between the CMOS DUT and the receive portion of the instrument be as short and as low in capacitance as possible. Instrumentation and additional line drivers were added to the fixturing near the DUT and this eventually grew into the testhead based architecture of today's ATE.

In general, the problems described above are true for both stray capacitance and inductance in the signal line. As frequency increases, inductive reactance ($X_L = j\omega L$) and capacitive admittance ($X_c = 1/j\omega C$) both increase linearly with frequency. Small lead inductances, which at lower frequencies can be neglected, become significant at microwave frequencies and any shunt capacitance can effectively short out signals.

4.2.4 Transport Delay

In addition to the impedance matching issue described above, transmission line path length has another impact on testing. This can be seen in the schematic of a typical digital channel, shown in Figure 4.6.

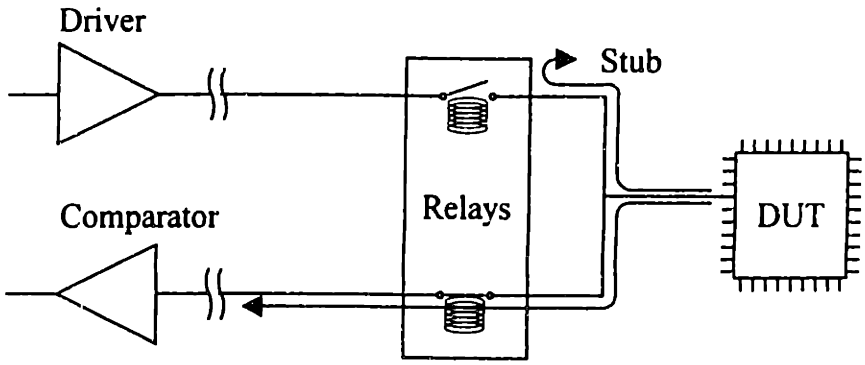


Figure 4.6: Schematic of Digital Pin Electronics

Instruments are designed in such a way that a DUT pin can be tested in both the send and receive modes. This can be accomplished using a driver and comparator separated by a relay that is placed near the DUT to switch between the two to eliminate the stub capacitance of the unused line. Often, the output of the instrument driver is determined by the response from the prior stimulus and the finite speed of signals travelling along these lines creates a finite delay between the input stimulus and the output response. At low clock frequencies over short lines, this delay can be ignored. However, as clock frequencies increase, this delay becomes significant and if the roundtrip delay becomes longer than a clock cycle, there is insufficient time to calculate the appropriate response. As a result, the tester must run at a lower speed, introduce delays in the test patterns, or greatly increase the complexity of the test program to accommodate for the disjunction of the input/output relationship. The first two options do not test the device under normal operating conditions and therefore reduce the accuracy of the test and the third option becomes increasingly difficult as the number of delayed clock cycles increase. The ideal solution is to maintain a path length that is short enough to avoid this issue altogether. A 100 MHz clock has a cycle time of 10 nanoseconds, corresponding to roughly 2 meters of line length, which is not a significant challenge. However, at 1 gigahertz, the round trip path length shrinks to 20 cm. It will be extremely difficult to place to place the required quantity of digital drivers, possibly more than 4000, all

within 10 centimeters of the DUT, and therefore it is important to understand and model this feature of signal transmission lines.

4.3 Physical Implementation of Signal Transmission Lines

The preceding discussion dealt with transmission lines as abstract entities with delay length, capacitance and impedance, but no mention has yet been made to the physical implementation of these elements. There are many techniques and components available to build transmission lines all with distinct benefits and drawbacks. In ATE signal transmission system design, there are a number of attributes that are considered. Of primary importance are:

- Signal quality/bandwidth: This is the sum of some of the effects discussed above.
- Ease of fabrication: Test systems may have several hundred to several thousand digital lines and therefore any component used must be simple and cost-effective.
- Configurability: The generic nature of ATE requires that it be configurable to meet a diverse range of needs. The signal transmission system is the most practical area for this configuration to occur and therefore this path cannot be completely hard tooled.
- Durability: Transmission systems are subject to both mechanical and electrical wear. As a component in a production fixture, it is likely to see significant mechanical abuse and when operated at high frequencies, may suffer from electrical wear phenomenon such as dendritic growth at the contacts and embrittlement due to ion migration. Generally, mechanical failure is much more common than electrical failure in these components.

The design of every signal transmission component and system is a trade-off between these attributes. At quasi-static frequencies, most transmission elements behave as simple DC circuits with the shield or ground plane as a signal return. As frequencies increase, the interactions of the electric and magnetic fields become significant. Specifically, if one inspects Maxwell's equations, the time rate of change of the electric and magnetic fields are interrelated. At higher frequencies, and therefore higher time derivatives, these electromagnetic effects begin to dominate the behavior of the transmission line.

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$$

$$\nabla \times \mathbf{H} = \frac{\partial \mathbf{D}}{\partial t} + \mathbf{J}$$

$$\nabla \bullet \mathbf{D} = \rho$$

$$\nabla \bullet \mathbf{B} = 0$$

Where:

E = *electric field*
B = *magnetic field*
D = *electric displacement field* = $\varepsilon_o E$
H = *magnetic intensity field* = B/μ_o
J = *conduction current density*
ρ = charge density

While the details of electromagnetic theory are beyond the scope of this work, there are several principles that must be understood to properly design transmission elements. The most important of these is the distributed nature of transmission line impedance. Much like continuous beam analysis, transmission lines cannot be modeled using lumped capacitive and inductive elements if its high frequency behavior needs to be captured. At any point along the length of the line, the total impedance per unit length must be constant or reflections will result. To avoid these reflections, a common impedance is typically selected and all transmission components are design to nominally match this value. In addition to this, the line must be terminated with the same impedance to avoid a reflection at the end of the line. There are several industry standards for nominal impedance, but the most common is 50 ohms.

Another important fact that results from Maxwell's equations is that high-frequency signals propagate as transverse electro-magnetic (TEM) fields, as can be found in the E and B fields of Maxwell's equations. Figure 4.7 below shows a few of the modes of energy propagation in a typical coaxial transmission line Because there are multiple, three dimensional modes, transmission lines must be carefully designed so that they can properly support these fields. This is usually accomplished by partially or completely surrounding the field with a conductor that is tied to ground. However, this is not always possible, especially at discontinuities such as connectors. These sorts of discontinuities require complex magnetic field solvers to accurately model and will not be discussed in this thesis.

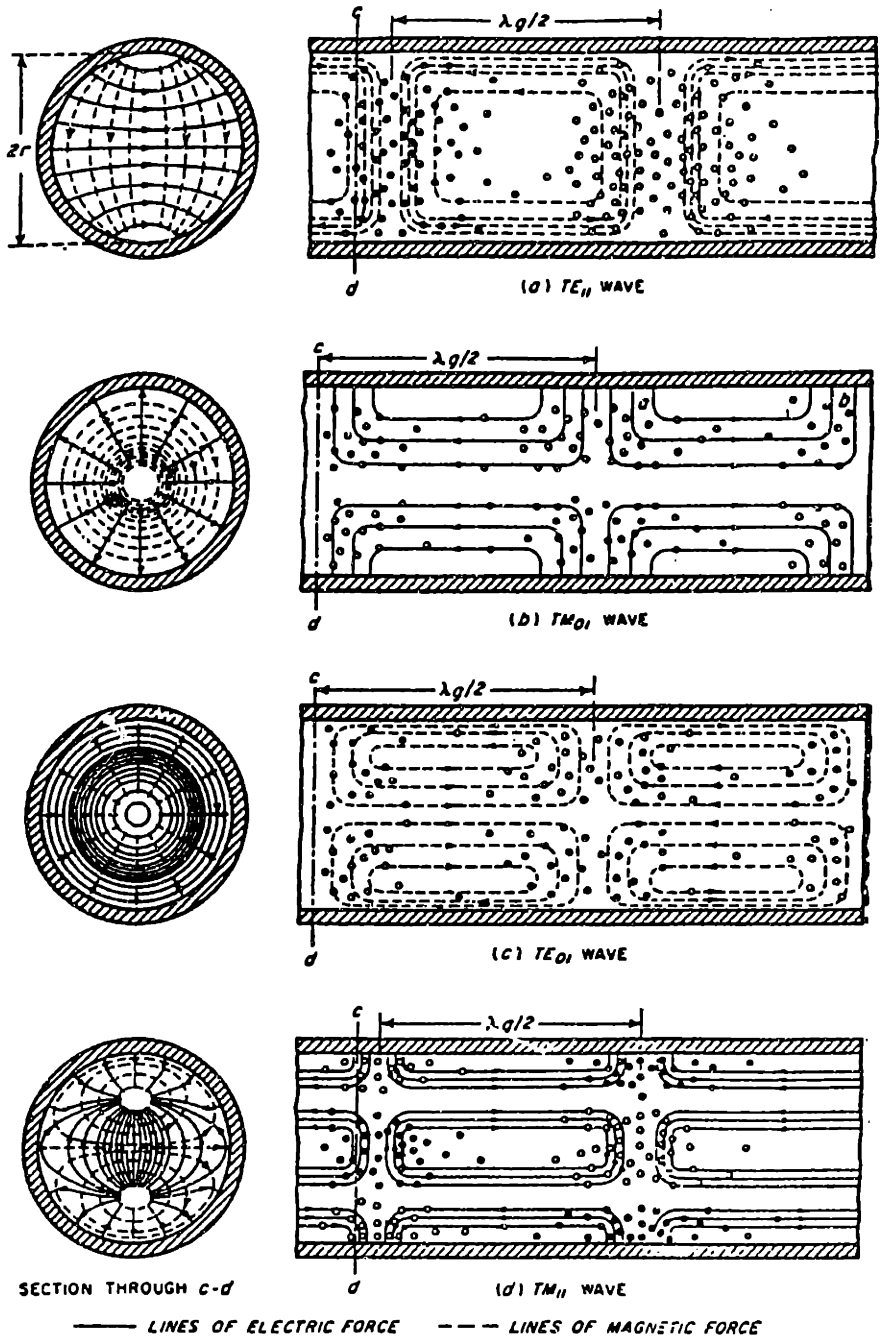


Figure 4.7: Primary (TE11) and Higher-Order Modes in a Circular Waveguide[25]

The following sections review the physical descriptions of common transmission line elements, along with how they perform against the above criteria. These elements include cables, printed circuit board structures, connectors and DUT fixturing.

[25] Terman, 1955

4.3.1 Cable and Waveguides

The most widely recognized transmission line element is the coaxial cable, which consists of a center conductor and a coaxial outer shield separated by a dielectric material, as shown in Figure 4.8.

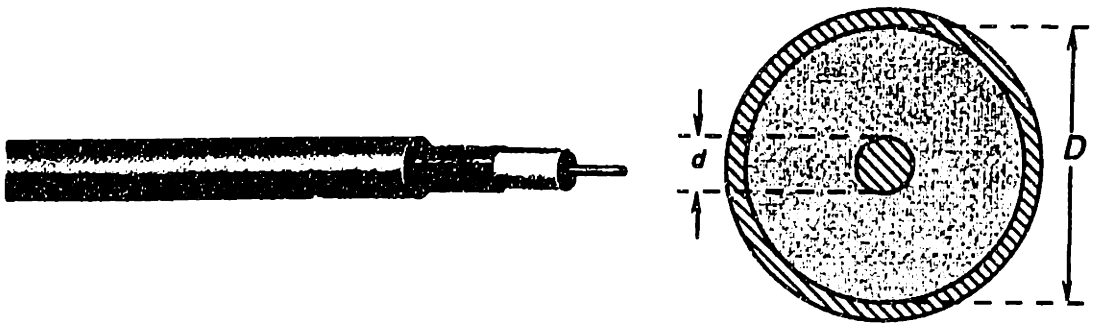


Figure 4.8: Coaxial Cable Construction

At quasi-static frequencies, the coax line acts as a simple circuit with the center conductor as the signal line and the shield as the ground return. At higher frequencies, the signal is carried primarily in the field between the conductors, supported by the dielectric material. Due to its symmetric design and complete encirclement of the center conductor, coaxial cables can properly support the EM fields and thus generally have very high bandwidth. Characteristic impedance of coaxial cables can be calculated using the following equations:

$$Z_o = \frac{\eta_o}{2\pi\sqrt{\varepsilon_r}} \ln(\frac{D}{d})$$

$$\alpha_c = (\frac{0.014272\sqrt{f}}{Z_o})(\frac{1}{d} + \frac{1}{D})$$

$$\alpha_d = 0.0912 f \sqrt{\varepsilon_r} \tan(\delta)$$

where:

η_o: dielectric constant/permittivity of free space (ϵ_o/μ_o)
ϵ_r: relative dielectric constant
Z_o: nominal design impedance
f: frequency [rads/sec]
D: outside diameter
d: inside diameter

Dominant modes of loss in coaxial cables are conductor (skin effect) and dielectric losses, represented by α_c and α_d above.

Small numbers of coaxial cables are generally easy to fabricate and are extremely flexible. However, they are not a good selection for high pin count applications because large bundles of cable tend to be difficult to manage and are prone to failure due to mechanical fatigue.

A similar cable transmission line is the twisted pair, which is common in telephone and Ethernet applications. This arrangement approximates coax where the ground line spirals about the signal and the two are separated by the dielectric insulating material. At lower frequencies where the wavelength is much longer than the spiral pitch, usually less than 100 megahertz, this is an acceptable substitute for coaxial cables. It has the same benefits and suffers the same drawbacks as standard coax cable.

A third transmission element that may be placed in the cable category is the waveguide. This is essentially a hollow tube or rectangular extrusion that is used to contain the EM fields. Because there is no center conductor and no dielectric material other than air, waveguides have very low loss. The lack of a center conductor means that waveguides cannot support DC signals. Due to the fact that microwaves have similar wavelengths to those of audio pressure waves, these waveguides often look very similar to acoustical components. However, except in tuned microwave components, these large mechanical structures are generally not used in testing. One variant of this design often used in narrow-band microwave testing is the semi-rigid waveguide, which is essentially a hollow conductive tube made of malleable metal with electrical connectors on the ends. While these components exhibit excellent high frequency performance, they are impractical for high speed digital (HSD) test fixturing due to their bulk and rigidity.

4.3.2 PCB transmission lines

Complex transmission line structures are typically implemented using printed circuit board (PCB) technology. They can generally be divided into two distinct categories, microstrip and stripline. Modern PCBs are built using multiple layers of signal and ground planes, with signals on the outer layers and power and ground planes on the inside layers and if more than two signal layers are required, they are generally placed between two ground planes. Signals that reside on the outer layers of the PCB are referred to as microstrip, and those on the inner layers are called stripline. The proximity of the signal to the ground plane, separated by a dielectric material, creates a controlled impedance environment similar to that of a coaxial cable. Figure 4.9 shows electrical and magnetic field lines of an asymmetrical stripline element.

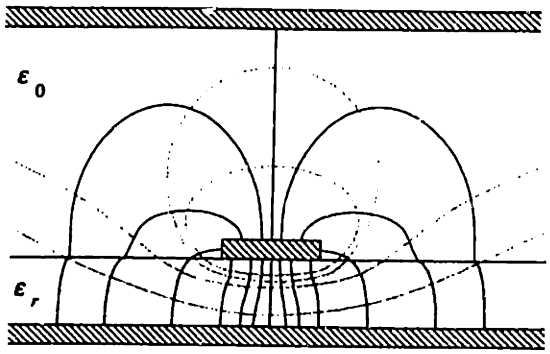
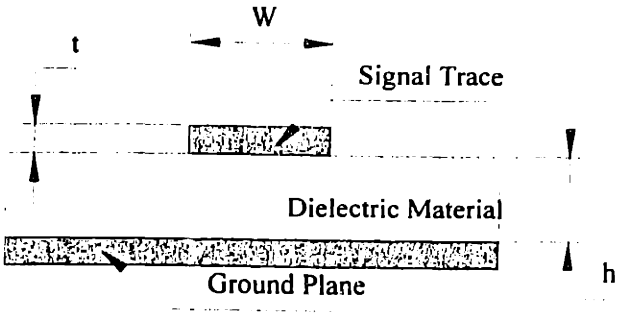


Figure 4.9: Fields in a Non-Symmetric Stripline Structure[26]

Microstrip structures are those that reside on the outer layers of a PCB, but in close proximity to a ground plane. The strip width, dielectric material and dielectric thickness determine the characteristic impedance of the line. The equation in Figure 4.10 below can be used to approximate the nominal impedance of a microstrip structure.

[26] Wadell, 1991

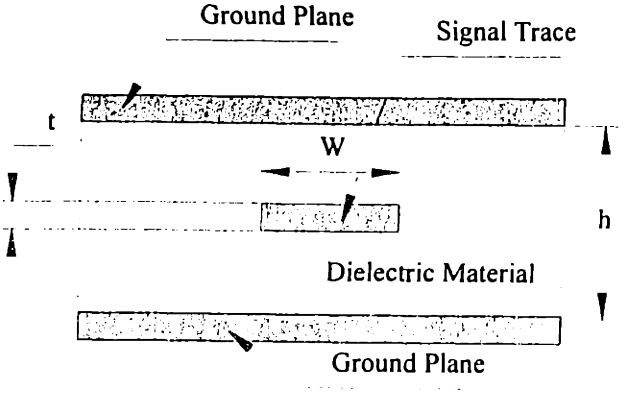


$$Z_0 = \frac{87}{\sqrt{\varepsilon_R + 1.41}} \ln\left(\frac{5.98h}{0.8w + t}\right)$$

Figure 4.10: Microstrip Geometry and Impedance Calculation

Unlike coaxial cable, microstrip does not have an axi-symmetric ground plane surrounding the signal trace. Because of this imbalance, microstrip lines cannot support higher order TEM modes which therefore limits its performance to the low gigahertz range for typical applications. Use of nonstandard geometries and exotic dielectric materials such as cyanate ester or glass reinforced Teflon can increase this bandwidth somewhat, however, these options have other drawbacks such as larger geometries, higher cost or poor mechanical properties.

The stripline structure consists of a PCB trace buried in dielectric material, usually equidistant from two ground planes. The equation in Figure 4.11 below can be used to approximate the nominal impedance of a stripline structure.



$$Z_0 = \frac{60}{\sqrt{\varepsilon_R}} \ln\left(\frac{4h}{0.67\pi w(0.8 + \frac{t}{w})}\right)$$

Figure 4.11: Stripline Geometry and Impedance Calculation

Because of it's symmetric geometry, stripline more closely approximates the performance of coaxial cable and therefore is somewhat more efficient that microstrip. Frequencies beyond 4-5 gigahertz are typical for short lengths of stripline structure.

While stripline has better performance, there are some significant disadvantages over microstrip circuits. Due to the encapsulated nature of stripline circuits, it is difficult to add series components, such as capacitors and resistors, to the line as is often done when tuning a test fixture. Stripline boards also tend to be thicker than microstrip for the same number of circuits due to the extra ground and dielectric layers required.

A third PCB transmission line structure that is somewhat of a compromise between advantages microstrip and stripline is the coplanar waveguide shown in Figure 4.12.

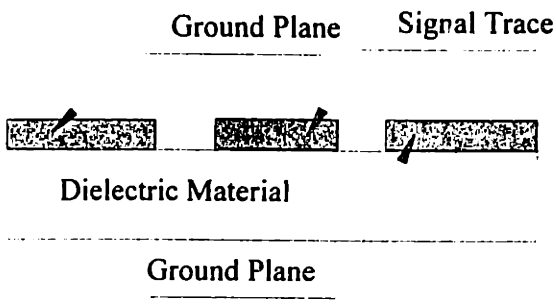


Figure 4.12: Coplanar Waveguide Geometry and Impedance Calculation

This structure has an efficiency somewhere between the previous two. It has the benefit of being on the accessible surface layer, but has a lower potential trace density. Linear arrays of waveguides with a single ground between 2 signal lines have about half the surface density, but if there is sufficient surface area, will result in approximately half the thickness of a stripline board. Also, for long runs, the two ground strips must be interconnected at regular intervals to prevent long ground loops. Impedance calculation for coplanar waveguides are somewhat more complex due to the variety of ground configurations and complexity of fields. Design equations can be found in Appendix 9.1 as well as several of the references.[27]

Despite their drawbacks, PCB transmission structures have many advantages that make them attractive for sub-gigahertz applications. Specifically, they are dense, durable and with

[27] Wadell, 1991

the exception of stripline, easy to repair and modify. They are also relatively inexpensive to manufacture using photolithography and many advanced electronic CAD tools are available to aid in their design.

4.3.3 Transmission Line Connections

It is unusual for only one type of the above transmission line component to be used in an ATE test application. More often, several different types are combined in series to build a fixture that not only has the electrical performance required, but also meets the mechanical packaging and application/configurability requirements. Thus, some means of connecting two dissimilar components is required and they can generally be categorized by the frequency of which they are mated, as follows:

- Permanent connection: mated once during manufacture/assembly.
  (e.g. solder, via, wire crimps)
- Configuration connection: mated only during assembly and reconfiguration.
  (e.g. backplane connector: cable connectors)
- Fixture connection, mated whenever type of DUT is changed.
  (e.g. spring pin, ZIF, cables connectors)
- Test connection: mated once per test.
  (e.g. probe needle, test socket, ZIF)

All connections between mating transmission elements present the potential for discontinuities in the impedance of the transmission line and thus introduce the potential for signal degradation and reflections. As one might expect, the connectors with lower mating cycles tend to have the best electrical performance. There are generally four issues of concern with most electrical high-speed connections:

- Mechanical reliability
- Contact resistance at the mating connector surfaces
- Unbalanced impedance
- Ability to cleanly transmit higher order TEM modes through connector transition.

Mechanical reliability is especially important for the latter two types of connectors, which are likely to undergo from hundreds to tens of thousands of mating cycles. Contact resistance introduced at the conductor boundary is generally due to the formation of metal oxides on the conductor surfaces. Gold plated contacts, combined with a wiping, scrubbing or piercing action between the two contacts used to break through this oxide layer are often

used. However, care must be exercised in the design of these contacts because these motions can also shorten the life of the contacts or cause irreparable damage to the DUT bond pad. The mechanical and packaging constraints placed on these connections often make it difficult to maintain a controlled impedance environment, but because the connector path lengths are generally short compared to the signal wavelengths, these impedance changes can be treated as lumped elements when modeling the transmission circuit. The final area of concern in connector design is their ability to smoothly transition the TEM modes from one transmission line type to another. Much like the way right angles in fluid pipe design and sharp corners in acoustics can cause losses, the transition between a planar microstrip trace to a perpendicular coax line can cause local disruption in the TEM fields. However, this is generally a high frequency, narrow-band, high order effect that requires magnetic field solving software to properly analyze and can be ignored for first order, broadband analysis.

In the sections that follow, a brief description of the functional requirements of each class of connector is given, along with relevant equations for analysis.

4.3.3.1 Permanent connections

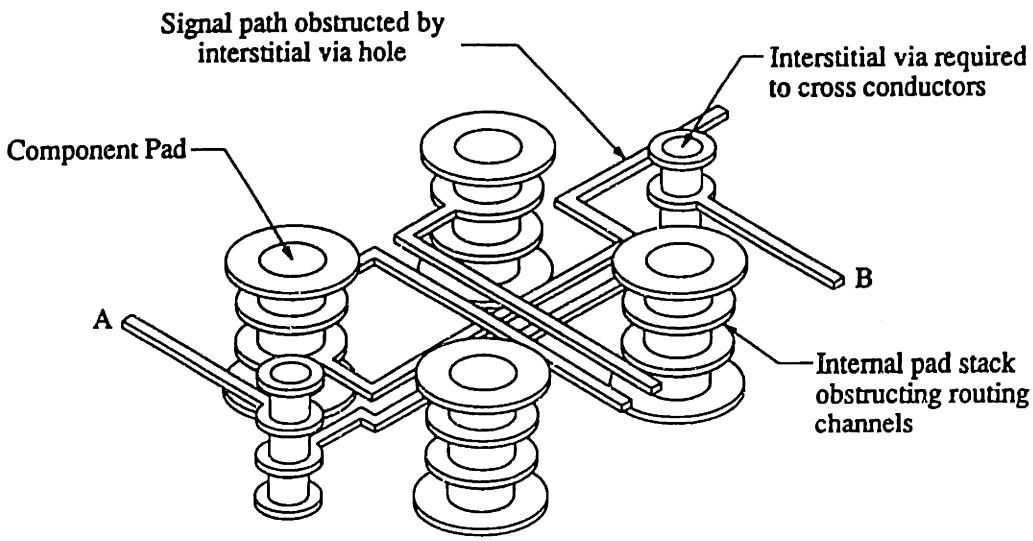


Figure 4.13: Structure of PCB Pads, Vias and Signal Traces[28]

Permanent connections are distributed throughout an ATE system, but are often overlooked as sources of transmission degradation. The two most common types of this

[28] Buck, 1997

class are PCB vias and solder pads as shown in Figure 4.13. Vias are used for inter-layer connection and for through-hole components and pads are used for surface mount devices. Vias are essentially vertical jumps between horizontal PCB layers and therefore do cause some TEM field disruption. However, the most significant impact of these structures is due to the fact that, depending on the shape of the pad and ground plane, they may have stray capacitance or inductance. For example, a standard 0.060" round pad can have several picoFarads (pF) of capacitance. This can often be reduced to a few tenths of a pF by reducing the pad diameter and relieving the ground plane underneath the pad, but as mentioned above, one of capacitances detrimental effects is cumulative and extremely dense boards may require several vias per trace. Generally, permanent type connections to not suffer from mechanical reliability issues, however solder interconnections may fail under shear if a PCB with an array of connections is subjected to bending or extreme temperature swings.

4.3.3.2 Configuration connections

Configuration type connections, such as backplanes and edge connectors generally are not a great source of problems due to their low number of mating cycles and the amount of engineering design effort that goes into this these types of components. A typical backplane connector is shown in Figure 4.14 below.

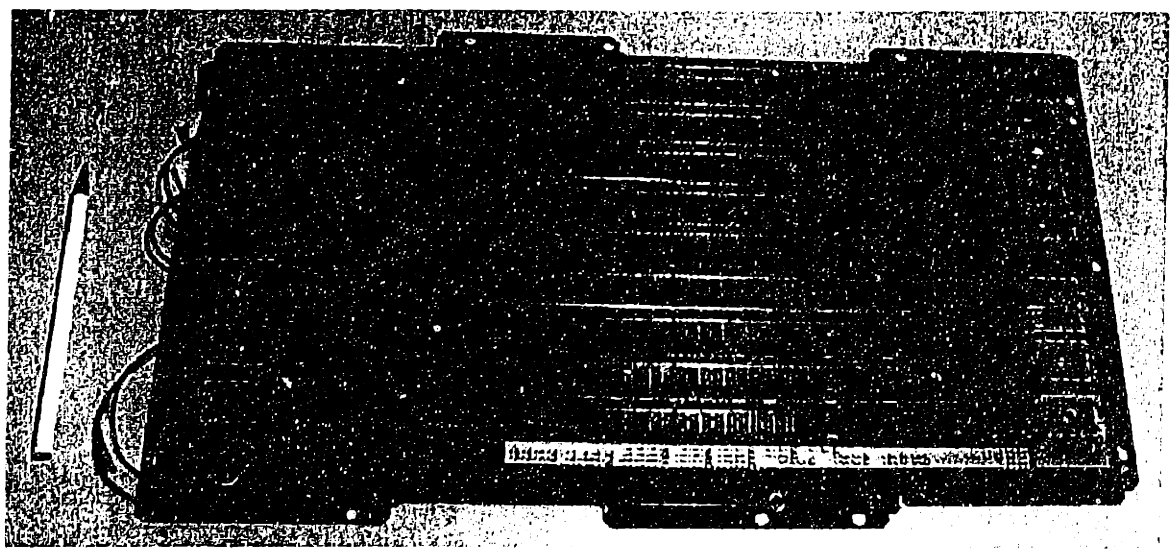


Figure 4.14: Backplane Connector Array

4.3.3.3 Fixture connections

Fixture connections are a common source of problems and compromise. One reason for this is because there are no stable, dominant technologies in this area and short device product cycles restrict engineering design effort in this area. Generally, the test fixture, referred to as a device interface board or DIB, is in the form of a round or square PCB, as shown in Figure 4.15.

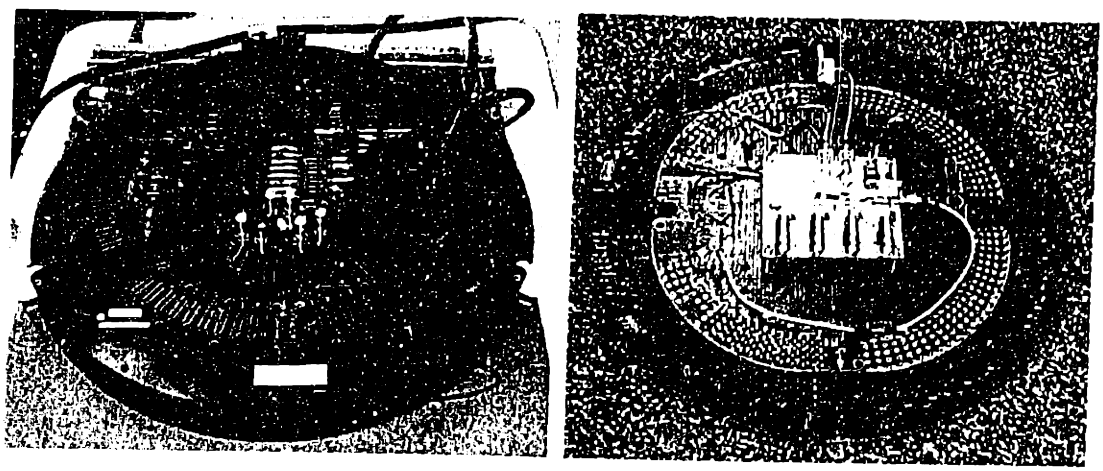


Figure 4.15: Device Interface Board (DIB) Test Fixtures

The most common method of connecting the test electronics to the fixture is through an array of spring pin contacts that connect through pads on the DIB. This arrangement is shown in Figure 4.16.

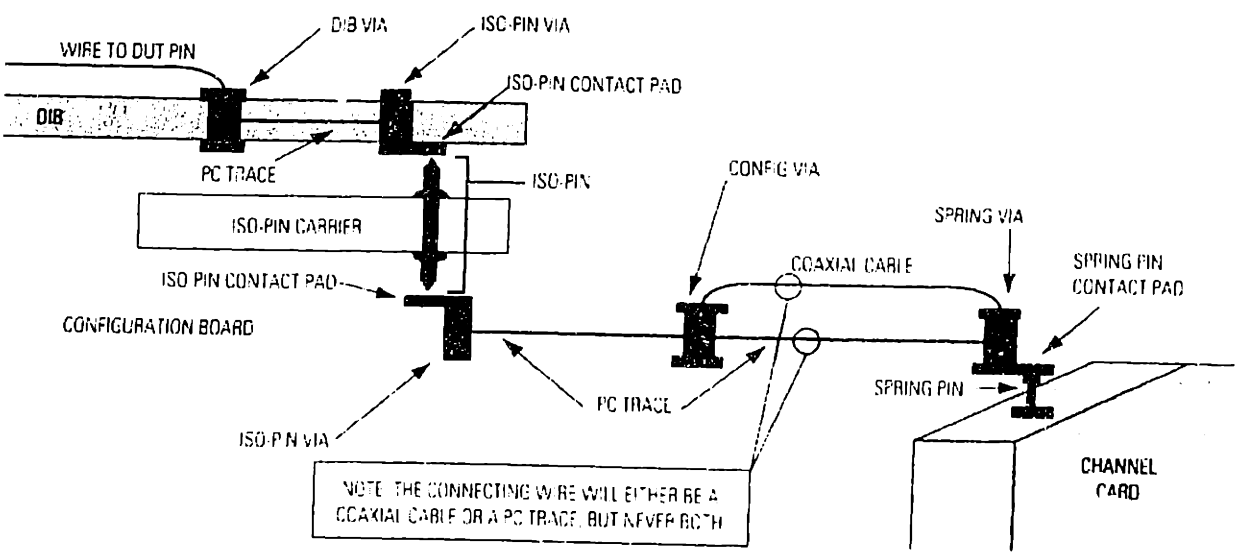


Figure 4.16: Schematic of Typical Spring Pin Interface

There are many ways to design a spring pin array, also known as a Pogo[29], or iso-pin tower, but all of them require separate ground and signal connections to control the impedance in the tower. Therefore, for each DUT channel, there will be at least two pins, one signal and one ground, required to transmit the signal. The signal and ground pins are then arranged in an array, separated by a dielectric material, such that the characteristic impedance of the line is controlled. Three common methods of arranging these arrays are shown in Figure 4.17.

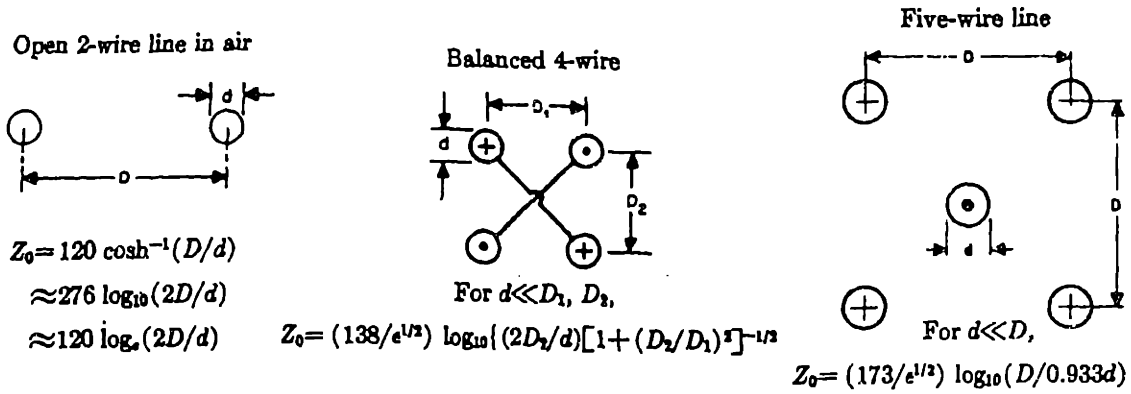


Figure 4.17: Controlled Impedance Spring Pin Configurations

Pogo towers are generally built from relatively close tolerance mechanical components and therefore have good transmission characteristics and are fairly robust. While signal integrity through a pogo tower is generally very good, this type of connection does have several major drawbacks:

- Cost: An assemble tower can cost about $5-$10 per pin and a 1000 line interface may have in excess of 3000 pins.
- Compression force: Force per pin ranges from 30 – 100 $grams_f$. Therefore a 3000 pin array may take several hundred kg_f to compress and may impart significant deflections to the fixture PCB. Often, deep walled stiffeners are attached t the fixtures to counteract this effect.
- Packing density: Because they are discrete mechanical structures, there are limits to the achievable signal density of this type of interconnect. Mechanical alignment tolerances require relatively large contact pads and signal path length increases with array area.
- Capacitance: Spring pin connections are generally designed to a nominal impedance and generally do not have any excess, or stray capacitance. However, the contact pads on which the pins touch do have some capacitance that is dependent on the pad diameter.
- Inductance: Most of the length of a spring pin array is contained within the dielectric material. However, the small, compressible length at the end of the pin is uncontrolled

[29] Pogo is a registered trade name of Pylon, Inc.

and appears as a series inductance whose value depends on the uncompressed length. This introduces a coupling between the electrical performance and the mechanical accuracy of the interface.

- Contamination: Spring pin contacts do not incorporate a scrubbing action during contact and therefore are susceptible to intermittent contacts due to dirt and dust. For this reason, the pins generally have sharp points to provide a piercing action.

The design of spring pin arrays is therefore an exercise in optimizing competing trade-offs of density, accuracy, capacitance and cost. Because of the design effort and cost that goes into these connectors, designs are generally standardized across a test floor or test system. A spring pin tower assembly is shown in Figure 4.18.

Figure 4.18: Spring Pin Tower Assembly

There are other types of fixture connections, though none as well accepted as the spring pin. Socket and edge connectors are occasionally used in lower pin count interfaces. Tight board layout and mechanical accuracy requirements and high insertion forces limit the size of these connectors. Zero insertion force (ZIF) connectors can be used for a small number of pins, but due to the complex mechanisms required to actuate a ZIF connector, a high density, high pin count ZIF interface has not yet been successfully implemented. Such a design might be a solution to many of the spring pin problems described above. Cables are often used for low to moderate pin count interfaces. However, they are not practical for high pin count or the same reasons described above, namely stiffness, cost, and reliability. Some new elastic

materials used to interconnect parallel PCBs are gaining acceptance. These composite materials mate dielectric elastomers with metal traces or wires to create a matrix that conducts only through it's thickness. This thin, slightly compliant material is placed between two PCBs and compressed to make a compact, high density connector and the relatively small distance between PCBs makes impedance matching less critical. The drawbacks of these materials include high cost, fragility and susceptibility to dirt and dust.

4.3.3.4 Test connections

The final class of connections includes those that make direct contact with the DUT. These fall into two general categories, probe cards and handler sockets.

Probe cards make direct contact with the DUT while it is still in wafer form. For this reason, it must be very accurate and very dense, yet the ends of the probe needles must have enough compliance to accommodate for misalignment and so that they do not damage the wafer if over-driven. The simplest method for accomplishing this is to use a long, cantilevered beam with a tip at the end, as shown in Figure 4.19.

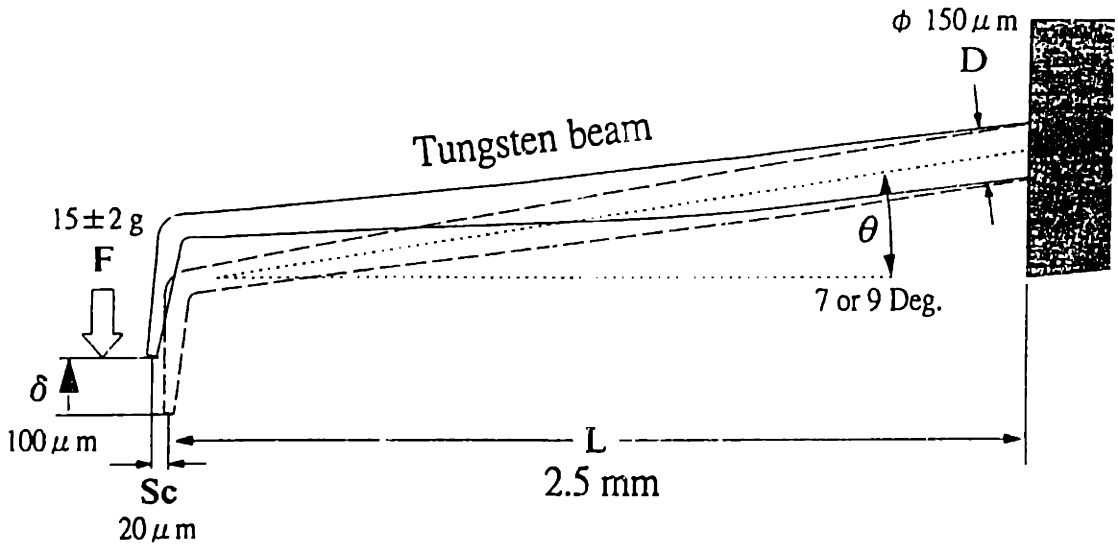


Figure 4.19: Cantilevered Beam Probe Needle

Due to the unsupported length, simple cantilevered beam probe cards suffer from Abbe errors and thermal drift and are adequate only for small, perimeter array devices. Higher density probing can be achieved by embedding the probe wires in epoxy to give it strength and thermal stability or a vertical blade structure can be added to the tops of the probes.

Simultaneous probing of multiple die can be achieved using high density, multi-site versions of the cantilevered technology. As can be seen from the simple epoxy ring and a multi-site probe cards are shown in Figure 4.20 and Figure 4.21, this technology is highly complex, fragile and is likely near the limit of its capability.

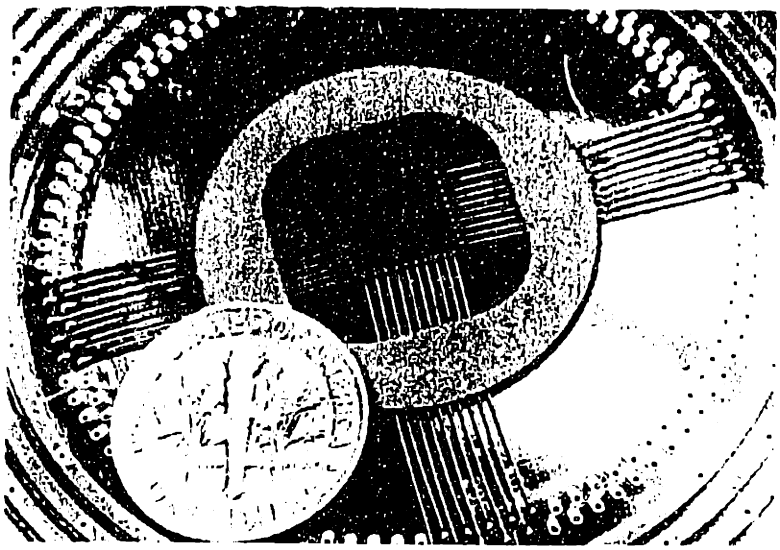


Figure 4.20: Epoxy Ring Probe Card

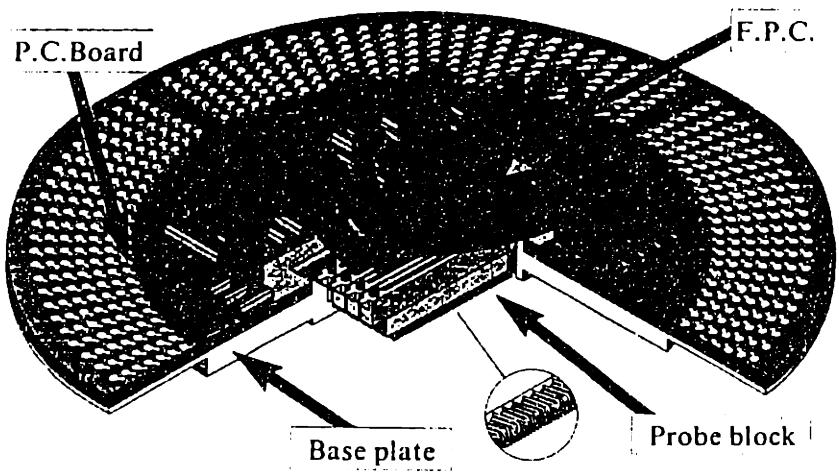


Figure 4.21: Multi-Site Cantilevered Beam Probe Card

Cantilevered beam probe cards are limited in that they are only capable of probing perimeter padded die. Some cards are capable of probing two or three rows of staggered pads, but a different technology is required for pads arranged in a 2D array. Vertical probe cards, capable of array probing, are realized using several different designs. The most common is the 'Buckling Beam' probe shown schematically in Figure 4.22 which takes

advantage of the non-linear force deflection curve of a beam under compressive buckling to give the probe tips the necessary compliance to probe the pads.

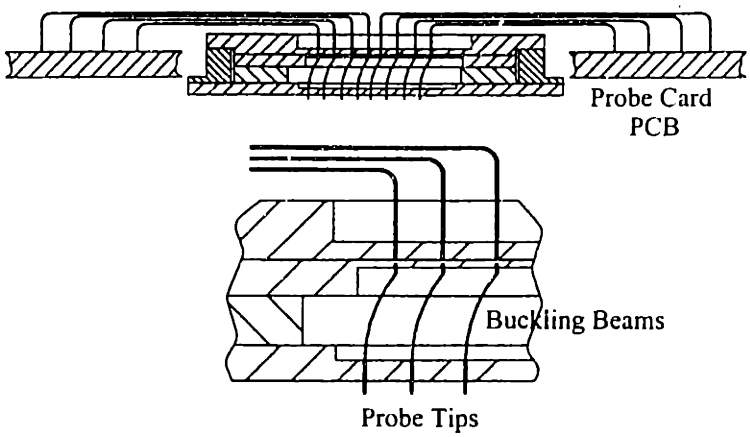


Figure 4.22: Buckling Beam Area Array Probe Card

All of the above technologies rely on a length of wire, either cantilevered or buckling, to provide the compliance required to probe the delicate wafer. It is often difficult to control the impedance of this length of the signal path and therefore, these unshielded probe needles can be highly inductive. Often, discrete capacitors are used at the base of the cantilever, which reduces ground bounce, but increases the return capacitance of the line.

There are a few emerging technologies that may address the probe needle inductance. The simplest is the coaxial probe needle, constructed by coating the needle wire with Teflon, or some other dielectric and then plating with electroless nickel, or by using a sandwich of dielectric between two conductive layers.

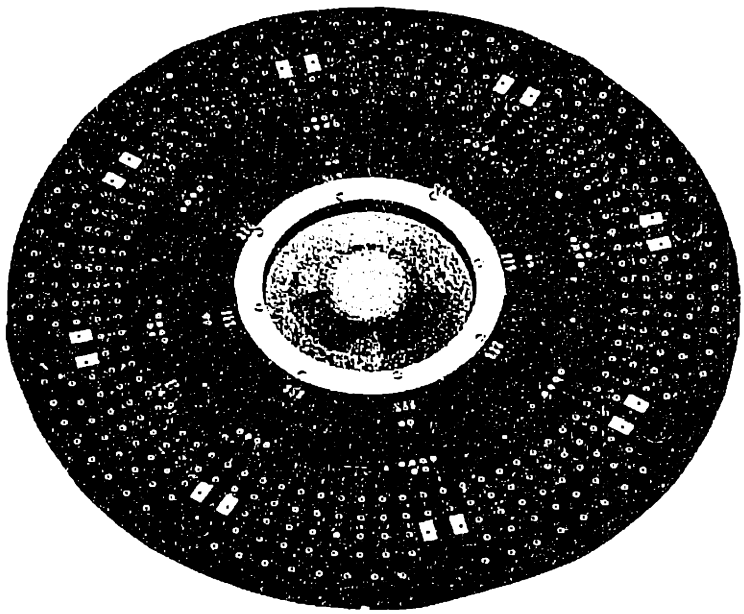


Figure 4.23: Membrane Probe Card

Figure 4.23 shows a membrane probe card which uses a flexible, polyamide-based material as the probe card substrate. Traces are run on the card directly to the device bond pad and gold bumps are used as probe tips. Air pressure or a spring is used for compliance and since the card is built using standard layered PCB methods, impedance can be tightly controlled. The newest form of probe card uses lithography and wire bonding technologies to build 'micro-springs' with plated contact points. This results in extremely short (less than 0.050") probe needles and therefore very low inductances. This design also has an advantage in that its manufacturing methods will automatically advance with the DUT fab technology. It is likely that this, or similar lithography based probe cards will become the dominant technology as device feature size continues to shrink. Eventually, it may be possible and necessary to add integrated test circuitry directly to the probe wafer, thus keeping the signal path short and low in capacitance.

Handler sockets, the second type of test connection, are used when testing devices after they have been packaged. Because most packages already fan out the device pins to a larger area, test sockets are generally easier to design and build. However they are still subject to the same constraints as probe cards, namely controlled impedance and compliant contact with the device. There are three basic types of test sockets:

- Cantilevered beam contacts
- Short contact
- Pin contact

Cantilevered beam contacts work on a principle similar to a probe needle. An assembly of beams are arranged around the perimeter of the socket. The length of the cantilevered beam limits the package size and contributes significant, uncontrolled impedance to the signal path. This type of socket is adequate for low pin count, low speed, perimeter leaded devices.

The package size and impedance issues are addressed by a number of short contact test sockets. These are assembled using a metal-on-elastomer material, flexible PCB or a novel S-beam approach. The S-beam floats on two elastic beams between the device and the PCB fixture. The shape of the beam is designed such that compression motion from the device rotates the beam, causing a scrubbing motion at both the device and the PCB contact, thus eliminating any oxide layer. Both the cantilever and S-beam type contactors are shown in Figure 4.24.

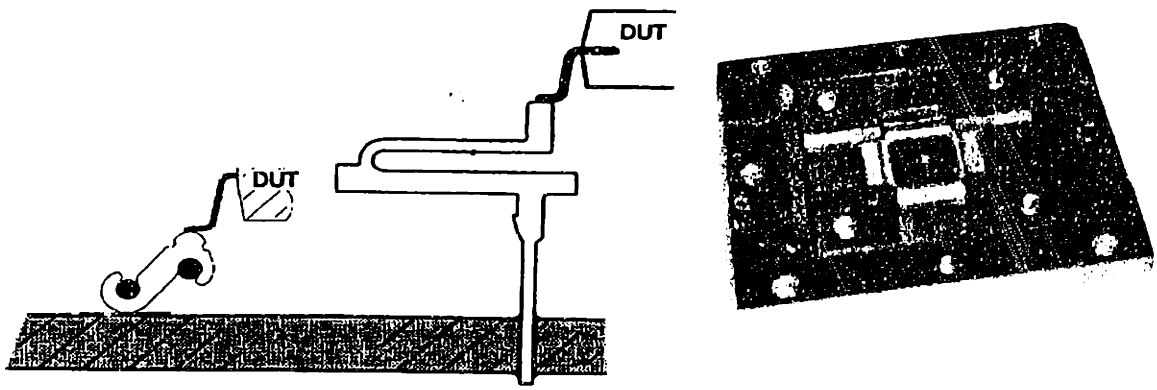


Figure 4.24: Package Contactor Designs and Contactor Assembly

Neither of the two test socket technologies described above is capable of testing large array devices, such as ball grid arrays (BGA) and pin grid arrays (PGA). Currently, the most common method for contacting these types of devices is an array of spring pins. Generally, the pitch of the device pads is sufficiently tight such that there is no means for controlling the impedance of these pins beyond the point from which they launch from the board.

4.4 Signal Transmission Roadmap

At this point, it is instructive to step back and see where the technology described above fits within the context of the overall roadmaps described in chapters two and three The question to answer is: What signal transmission performance will be needed to meet the demands of future devices? There are two primary trends predicted by the roadmaps that will drive signal transmission design: Higher digital pin counts and higher I/O speeds.

Pin counts, expected to exceed 4000 for high performance ASICs by 2010, are primarily a packaging concern, but have the potential to severely impact electrical performance. At issue is the ability to fan out or distribute such a high number of pins from the small area of the device(1-4 cm^2) to the pin electronics, while maintaining a controlled impedance environment. There are also concerns that fixture mechanical accuracy will not be sufficient due to the large number of pins, large array sizes and high probing forces. Chapter 3 investigates both of these issues in greater detail.

Device I/O speeds will likely be an even bigger challenge for ATE signal delivery. While internal clock speeds are predicted to exceed 2 gigahertz, it is unlikely that more than a few lines between the test equipment and the DUT will be required to run at this frequency. Of greater concern are the wide data buses that will run at speeds of up to 600 MHz. The difficulty of delivering several thousand controlled impedance lines to the DUT is exasperated by the decreasing signal voltage levels. Faster edge rates will also require smaller line capacitances, and round trip transmission delay will limit the total signal path length to less than a few inches. These trends are captured in Figure 4.25 below.

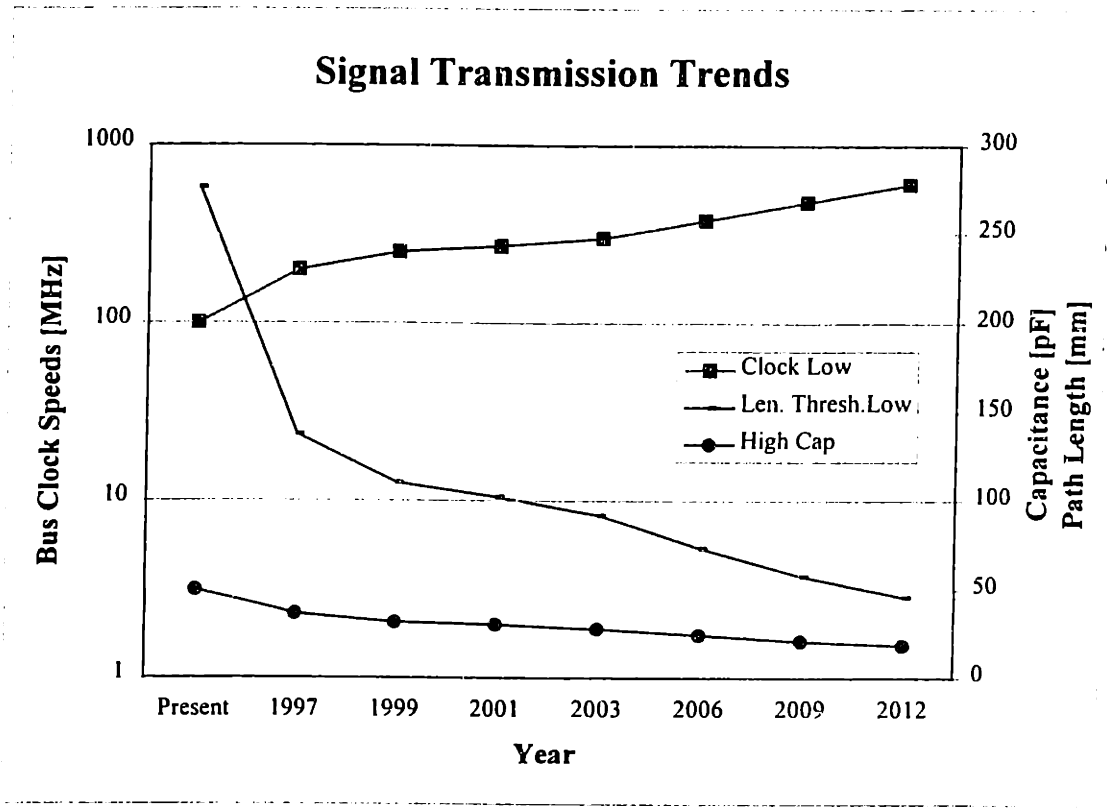


Figure 4.25: Signal Transmission Trends

Most of these problems are not unique to ATE systems. In fact, it is unlikely that the SIA roadmap predictions will be realized unless new technologies are developed that address the above issues in both the application and test of these devices. However, the majority of the lead users developing applications for these devices are not burdened by the additional mechanical constraints of ATE systems. Future PC's and portable electronics will remain fundamentally electronic devices with mechanical content limited to packaging, thermal and user interface concerns. ATE manufacturers will need to go beyond the purely electronic realm to address the complex electro-mechanical interactions involved in signal transmission.

Some of the theory behind this field was discussed in the sections above, but Maxwell's equations and complex impedance calculations are generally of value only to the researcher or experienced engineer. What tools are available for the engineer who will be designing these future test systems that require tight integration of the mechanical and electrical design? Currently, most design tools are focused on either one of these disciplines, but there are few, if any, that are capable of modeling both simultaneously at the systems level. The

next sections will briefly describe the features and limitations of some of the modeling tools that are currently available. This is followed by a description of a new design analysis tool developed as part of this thesis work. This tool, developed in the Mathematica programming environment, allows a designer to understand the effect of mechanical design trade-offs on electrical performance.

4.5 Signal Analysis Modeling

The design of ATE signal transmission systems is challenging because it combines microwave expertise from the field of Electrical Engineering and precision mechanics and manufacturing methods from the field of Mechanical Engineering. Research and development in either of these areas are somewhat of a black art, there is a small population of experts in each field, and an even smaller population that understands both of these areas. There are several commercially available modeling tools that can be used for analysis in this area; they can generally be divided into the following categories:

- Magnetic field solvers. (Ansoft)
- Electrical systems modeling. (SPICE)
- Microwave PCB design software. (EEsof)
- Mechanical systems modeling, FEA. (Mechanica, Algor)

Field solvers and FEA are high accuracy tools used to model the performance of individual components of a system. For example, Ansoft can be used to calculate the flux lines and impedance of a microstrip line and Algor might be used to determine the stiffness and deflection of a PCB fixture. With the exception of some simple optimization tools, both of these packages are limited to design analysis and are poor synthesis tools. There are a number of electrical systems modeling tools available, most of them are a variation of SPICE (Simulated Program with Integrated Circuit Emphasis) developed at Berkeley in the 1960's. A sample SPICE analysis is shown in Figure 4.26.

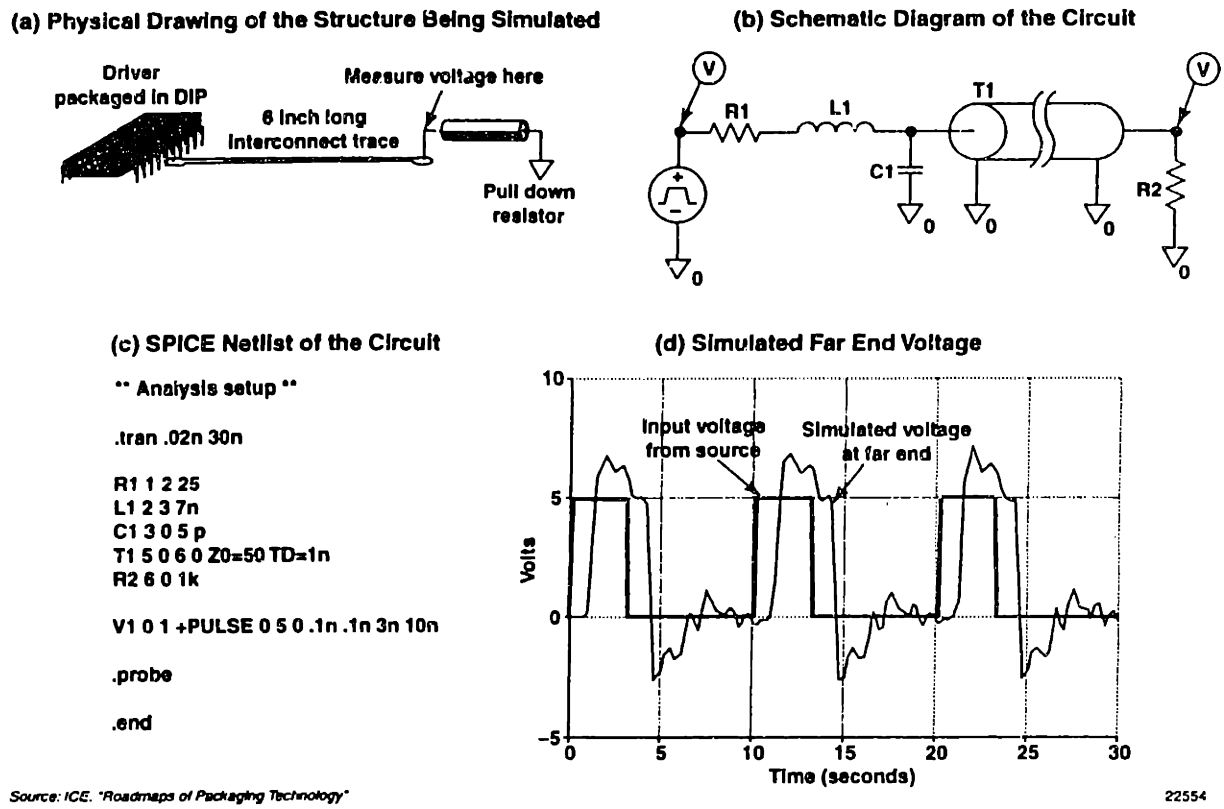


Figure 4.26: Simulation of Driver and Circuit Interconnect with SPICE[30]

These tools are capable of simulating very complex electrical circuit designs, from DC to microwave and from silicon transistors to power distribution networks. SPICE modelers also have some powerful tools for component tolerance analysis and links to PC board layout tools, however, all of this simulation remains strictly in the electrical domain. There are more specialized packages, such as EEsof, that bridge the gap between electrical component models and mechanical PCB implementation. These packages capable of modeling collections PCB structures, such as stripline, vias, sharp corners and stubs and analyzing them at the PCB systems level at microwave frequencies.

All of the above tools are capable of extremely accurate models of some part of a transmission system, however, none are usable for the type of system level modeling required for design concept synthesis and analysis. For this reason, a new tool was developed that encompasses all of the system components from the instrument to the DUT. The goal of this modeling effort is to build design tool that captures the mechanical design

[30] Source: ICE

parameters of ATE signal transmission systems and relates them to the relevant electro-magnetic principles. This is accomplished using the following steps:

1. A model of each of the components described above is developed that includes all relevant physical parameters (geometry and material properties) as input and return electrical properties (S-parameters, capacitance, and transmission delay) as output.
2. The above components are assembled within a larger framework that calculates system performance by cascading the component models above.
3. A transmission system is evaluated by defining the component models, assembling the components and then running the system model.

The models are programmed symbolically in Mathematica, and thus the output from step 3 is a function of all of the parameters input in step 1. Therefore, once a transmission system is defined in general terms (i.e. which components are used and in what order), the relationships between the electrical and physical properties can be explored. This enables he designer to quickly and graphically understand the effects of mechanical tolerance, path length and even temperature without having to build prototypes or running experiments.

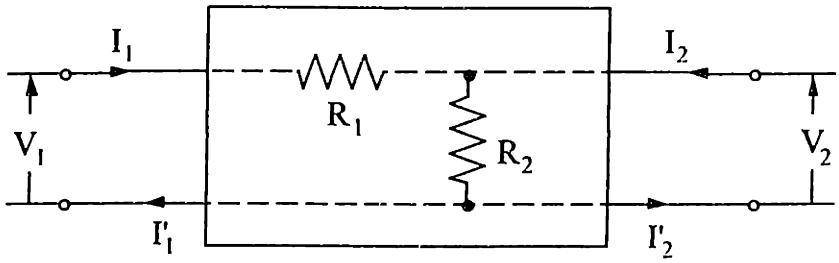
4.5.1 Analysis of Broadband Signal Transmission Using S-Parameters

The goal of this work is to develop a simple, first-order model of signal transmission systems that will give a designer basic insight into the performance and drivers of a system. As was mentioned above, quasi-static circuit analysis methods do not apply to these systems due to back loading of cascaded elements and reflections due to impedance mismatches. Ideally, this analysis could be accomplished using a spreadsheet and weighted, empirical data for each component in the transmission line. This method will not work however, because it is not the performance of the individual elements that governs the system. Rather, it is the interactions between the elements and a simple weighted model is incapable of capturing this interaction.

Several methods exist for modeling narrow band microwave transmission systems. These methods generally describe transmission elements as multi-port blocks with defined relationships between the inputs and outputs at every port. These methods are similar to cascaded filter analysis and Laplacian control system analysis methods, but they differ in one significant aspect. Typical control system analysis assumes that blocks have separate input and output ports and that signals and power flow along a closed loop. This uni-directionality limits the models ability to capture the effects of impedance mismatching, such as reflections

and back loading. What is needed is a method of capturing bi-directional flow within a network, rather than a single loop.

The simplest network model that captures all the relevant phenomenon is the Z-parameter, or impedance matrix. Figure 4.27 shows a generic 2 port impedance element. Unlike a Laplacian block, this element supports input and output at each port.



$$V_1 = z_{11}I_1 + z_{12}I_2 \qquad z_{11} = \left.\frac{V_1}{I_1}\right|_{I_2=0} \qquad z_{21} = \left.\frac{V_2}{I_1}\right|_{I_2=0}$$

$$V_2 = z_{21}I_1 + z_{22}I_2$$

in matrix form:

$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \qquad z_{12} = \left.\frac{V_1}{I_2}\right|_{I_1=0} \qquad z_{22} = \left.\frac{V_2}{I_2}\right|_{I_1=0}$$

Figure 4.27: Z-matrix Representation of a Transmission Element

Z-parameters of various elements can be easily calculated or measured by connecting it to a known voltage at port 1 and leaving port 2 open. The current flowing into port 1 and the open circuit voltage at port 2 can then be measured or calculated. Z_{11} and Z_{12} can be computed using the equations above, and Z_{21} and Z_{22} can be found by repeating this process at port 2. Z-parameters are useful for capturing input and output impedance relationships (Z_{11} and Z_{22}), but do not provide much physical insight into the behavior of a network. This is because the Z-parameter representation essentially models all components as voltage dividers and only models the behavior of the block when operating in an open-circuit environment. Transmission lines normally operate with matched impedances at the boundary and therefore these models do not capture the correct phenomenon. In addition to this, impedance matrices cannot model elements that have no open circuit current flow, such as transformers.

There are a number of other model definitions such as Y-parameters (admittance), or H and G parameters, all of which use combinations of voltage and current as dependent variables. All of these suffer from the fact that current and voltage are difficult to measure at microwave frequencies due to the interplay between input and output waves. A more intuitive, and easier to measure, variable is the power available at each port. Scattering, or S-parameters are defined in terms of incident and reflected wave energy and are more easily measured at high frequencies. S-parameters do not require open or short termination for measurement and exist for all common, linear circuit elements. The most important feature of S-parameters is the fact that their interpretation is intuitive and widely understood.

Figure 4.28 shows a more general two-port network element with a source impedance Z1 and a load impedance Z2.

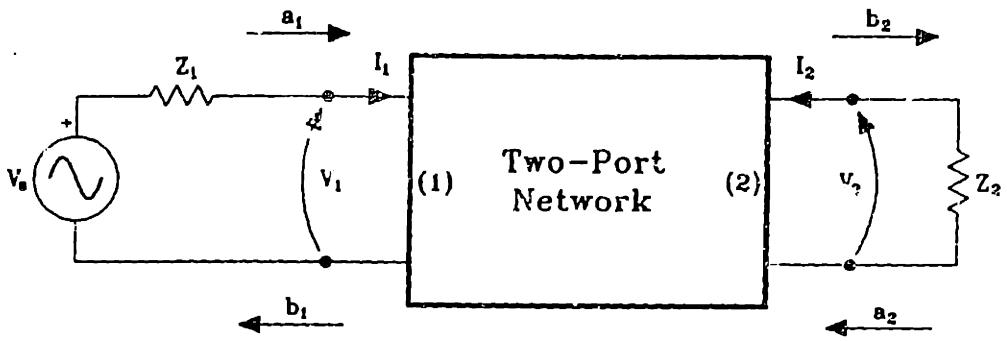


Figure 4.28: Two Port Network Element

Z1 and Z2 may not be discrete resistances, rather, they may represent the impedance of the rest of the transmission line as seen looking out of the two port network and is usually chosen to be a standard value, such as 50 ohms. The parameters a_i and b_i are measures of the incident and available power at the input and output ports of the element. There are a number of different definitions for these parameters, all of them essentially represent the RMS power at the port.

$$a_i = \frac{V_i + Z_i I_i}{2\sqrt{|\,\mathrm{Re}\,Z_i\,|}} \quad \textit{simplified:} \quad a_i = \frac{V_{inc}}{\sqrt{Z_0}}$$

$$b_i = \frac{V_i - Z_i^* I_i}{2\sqrt{|\,\mathrm{Re}\,Z_i\,|}} \quad \textit{simplified:} \quad b_i = \frac{V_{out}}{\sqrt{Z_0}}$$

The scattering matrices are then formed as follows:

$$b_1 = s_{11}a_1 + s_{12}a_2$$
$$b2 = s_{21}a_1 + s_{22}a_2$$

Because the nominal impedance, Z_0, is part of the definition of the S-matrix, it must be determined prior to the characterization of the element.

Scattering parameters can generally be measured and calculated for any discrete or continuous circuit or transmission element. The physical interpretation of the S-matrix coefficients can be calculated as follows:

$$s_{11} = \frac{b_1}{a_1} = \frac{V_1 - Z_1^* I_1}{V_1 + Z_1 I_1} = \frac{Z_1 I_1 - Z_c^* I_1}{Z_1 I_1 + Z_c I_1} = \frac{Z_1 - Z_c^*}{Z_1 + Z_c} = \Gamma_1 = \textit{Reflection coefficient}$$

Where Z_1 is the impedance looking into port 1 and Z_c is the nominal impedance and Z_c^* is the complex conjugate of the nominal impedance. S_{11} is often referred to as the reflection coeffiecent Γ, and is a measure of how much of the input signal is reflected out of the input port. In general:

$$|S_{ii}|^2 = \frac{|b_i|^2}{|a_1|^2} = \frac{\textit{Power reflected from the input}}{\textit{Power available from the generator}} = \textit{Return loss at port i.}$$

(reflection)

$$|S_{ij}|^2 = \frac{|b_j|^2}{|a_1|^2} = \frac{\textit{Power delivered to the load}}{\textit{Power available from the generator}} = \textit{Insertion loss at port i.}$$

(attenuation)

For passive transmission systems, conservation of energy dictates that both of these values are between 0 and 1 and the sum $s_{11} + s_{12}$ must be less than or equal to 1.

The scattering matrix representation may be used to model any linear transmission element and these matrices can be chained together to form a complete signal transmission model. Input impedance to a particular element is determined by the output impedance of the entire transmission system that preceeds that element, and output impedence is determined by the input impedance of the remaining transmission line.

Although the scattering models can be conceptually chained together, the S-matrices cannot be directly cascaded due to the way that the inputs and outputs are defined. S-matrices must be transformed to secondary matrix form, such as the scattering transfer (T-matrix) or the ABCD matrix, whose outputs are defined in such a way that they can be directly cascaded. Unfortunately, the parameters of the T and ABCD matrix forms are not commonly used or intuitively understood and therefore, once a complete model is built by cascading these matrices, it must be transformed back to a S-matrix for evaluation. These transformations are presented in section 4.6.2.3 below.

4.5.2 Application of Analysis Methods to Other Fields

Before discussing the details of how the above analysis methods are used to model real transmission systems, it is instructive to examine if these methods have applications beyond high frequency electrical signal transmission. Network wave analysis is by no means limited to the RF and microwave communications domain. These methods are applicable in any dynamic system whose components dimensions are on the same order of the wavelength of the stimulating frequencies. Optics and acoustics are two fields that use similar methodologies and the term "scattering" is originally derived from the reflection properties of optical elements. The physical implementation of microwave components such as waveguides and junctions are remarkably similar to those of acoustical components such as tuned ports and horns. This is due to the fact that a wavelength of a microwave frequency traveling at the speed of light is of the same order as common acoustical frequencies traveling at the speed of sound. An even simpler example is that of wave propagation on a continuous beam or string. Figure 4.29 is an illustration from an introductory physics text showing the effects of termination on wave reflection.

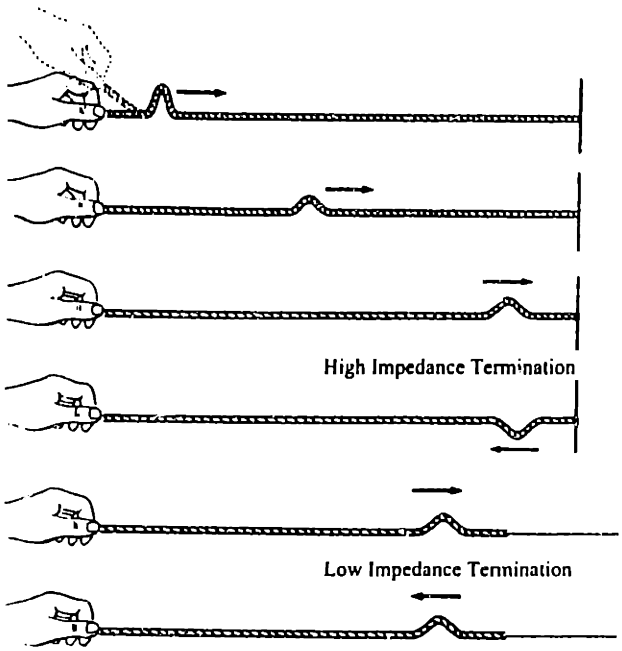


Figure 4.29: Wave on String Mechanical Transmission Line Analogy[31]

In both of the cases shown above, the impedance mismatch at the string boundary caused a reflection and resulted in no energy transfer from the string to the next element, therefore, $S_{11} = 1$ and $S_{12} = 0$. The mechanical equivalent of a typical signal transmission system would consist of a multitude of these strings in series with the addition of springs, masses and dampers at the string boundaries. As in electrical modeling, it may be possible to analyze these types of problems using high end tools such as finite element modelers, however, these tools are generally limited to a small number of transmission elements and are difficult to use for design synthesis. In addition to this, all of the available modeling tools are incapable of capturing the cross-domain coupling between the electrical and mechanical performance.

4.6 Model development

The basic model used for this analysis is a real representation of the S-matrix with additional elements to account for line length and capacitance. The model is implemented using *Mathematica*, a symbolic mathematical engine that is capable of solving large problems symbolically. The ability to retain the symbolic relationships between all the

[31] Tipler, 1982

variables is very powerful because systems of equation can be symbolically inverted, which permits the investigation of the relationships between any two variables. This is very appealing because systems can be optimized by varying both the global parameters, such as temperature, mechanical accuracy and manufacturing process capability, as well as variables local to elements, such as trace width, line length and dielectric constant. However, analysis of large systems is limited by the computing power, error propagation and machine accuracy. For this reason, some assumptions must be made to simplify and speed up the model analysis.

4.6.1 Model Simplification

With the exception of high order TEM losses at discontinuities, the scattering matrix definition described above is capable of exactly describing all of the relevant transmission phenomenon required for microwave signal analysis. This includes complex impedances, frequency dependant elements, internal losses, impedance mismatches at element boundaries and back loading of elements. While this is a very powerful methodology, it is also far more detail than is required for basic design analysis and for even very simple models, the complete scattering model becomes very difficult to analyze and solve. To simplify analysis, two basic assumptions are made to reduce the complexity of the model, as follows:

- All impedances are real: This eliminates all phase information, including phase delay, and ¼ and ½ wave reflections.
- The input and output impedances at element boundaries are independent of the neighboring elements.

The first assumption is not valid for narrow band microwave analysis, but is reasonable for a broadband digital signal due to the wide frequency content. Some important information is lost, such as total line length (in wavelengths) and line capacitance. This information, however, is linear and can be retained by summing the individual contributions of each element and can be tracked independent of the matrix data. The second assumption is valid for transmission systems that are designed around a nominal impedance, such as 50 ohms. With this assumption, a variation from 50 ohms in component impedance will cause a local reflection, but this variation will not load neighboring components. This unmodeled error will propagate throughout the system, but it's effect can be approximated by setting the

global nominal impedance to be slightly different from the design impedance, thus causing a similar reflection at component boundaries.

The goal of any modeling effort is to build a model that is sufficiently complex that it accurately captures the phenomenon of interest, yet simple enough so that it can be used easily. The difficulty in modeling electro-mechanical signal transmission lines is that the parameters of interest, i.e. geometry, temperature and tolerances, are driven by second and third order electro-magnetic phenomenon. For this reason, the scope of the initial model was limited to accurately capturing the cause-effect relationships between the parameters of interest, but at the expense of overall accuracy. However, the basic modeling framework developed is capable of more complex analysis if the above assumptions are tightened.

4.6.2 Model Structure

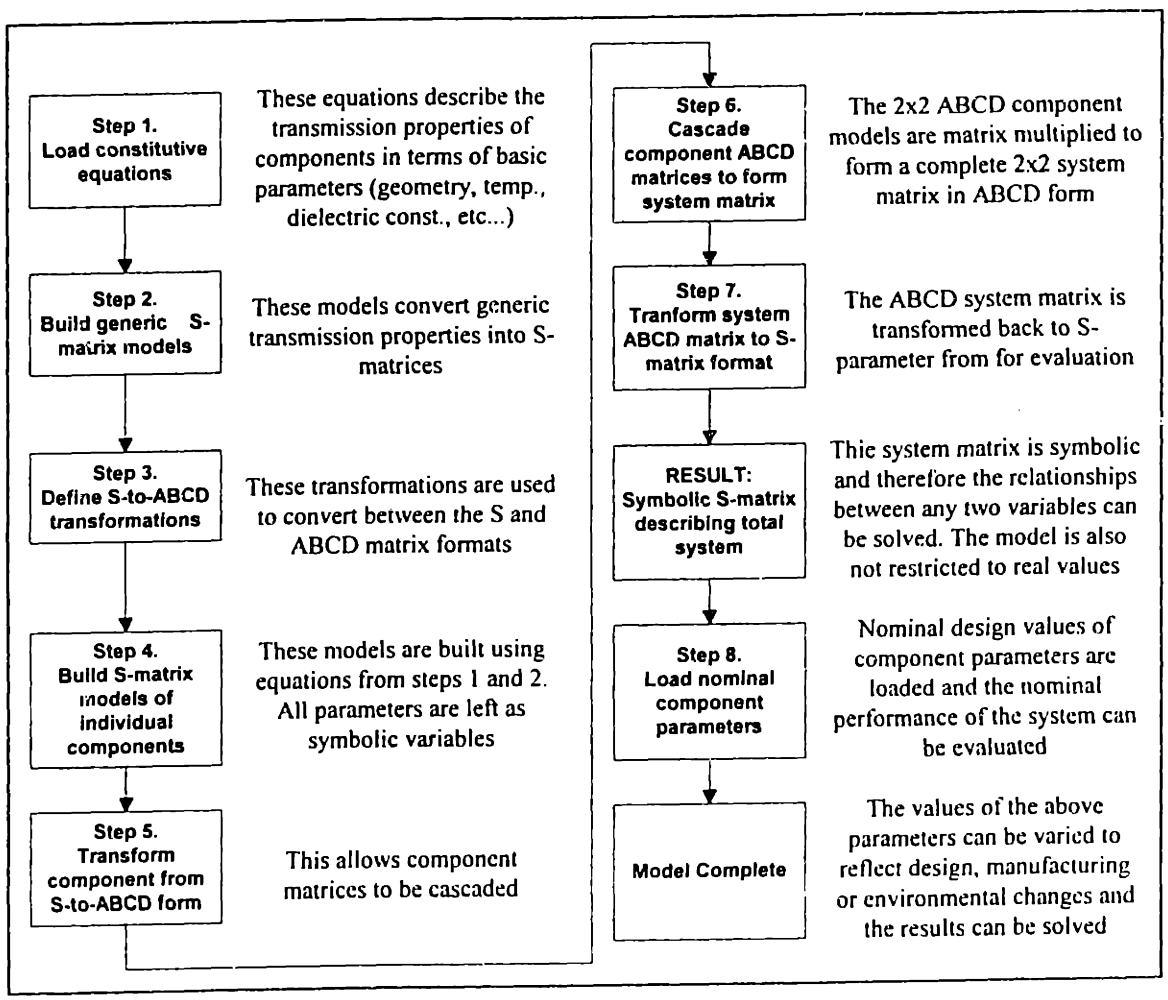


Figure 4.30: Flowchart of Analysis Methodology

The structure of the analysis process, shown in Figure 4.30 above, has the following core analysis components:

- Model definition. (Step 1)
- Constitutive relationships. (Step 2)
- Scattering matrix generation. (Step 3)
- Matrix transformations and cascading. (Steps 5,6 & 7)
- Component parameter definition. (Step 8)

The output of these eight steps is a complete, parametrically defined model of a signal transmission system design. The details of each of these steps are described in the sections below.

4.6.2.1 Constitutive Relationships

The constitutive equations used in this model primarily describe the relationships between physical dimensions and electrical performance. Physical dimensions include geometry, temperature, and manufacturing tolerances, as well as fundamental properties such as dielectric constants, conductivity and resistivity. Electrical performance parameters include resistance, capacitance, inductance, internal attenuation and phase lag. These parameters are used to generate a complex, frequency dependent impedance value for each of the components in the signal path. The simplified version of the model ignores the phase component of the impedance. The sections below briefly describe the primary constitutive relationships and code used in this model.

4.6.2.1.1 Lumped Elements

The simplest transmission line component is that which can be described as a lumped element. These generally include discrete components, such as bypass capacitors and chokes (inductors), as well as stray impedances due to component limitations. These include the PCB pad and via capacitances and inductances due to uncontrolled signal paths, such as a probe needle or contactor pin. The impedance values of discrete elements are taken at their nominal values, with an accommodation for the tolerance on the part. This is a reasonable assumption for surface mount components, but not for leaded discretes due to the inductance of the leads at very high frequencies. High frequency performance of leaded components can be modeled by cascading mulitple lumped elements. However, leaded components cannot generally be used in high frequency design for exactly this reason.

4.6.2.1.1.1 Spring Pin Contact Pads and PCB Vias

The excess lumped capacitance of PCB structures, such as Pogo pads and interconnect vias can be approximated as a simple, parallel plate capacitor using the following equation:

$$C = \frac{\varepsilon_r A}{d} \approx \frac{\varepsilon_r \pi r_{pad}^2}{t_{lam}}$$

In practice, it is impractical to use this relationship because the pad radius, r_{pad}, is generally controlled by the interface accuracy and density and the lamination thickness, t_{lam}, is usually controlled by the design of the microstrip transmission traces. In addition to this, good PCB design practice suggests that the ground plane be relieved around pads and vias, effectively limiting the capacitance to edge effects only. This results in an effective capacitance of approximately 0.2 to 0.5 picoFarads.

■ Channel Card Pads (ccpad)

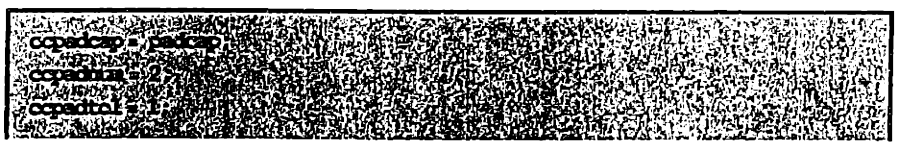


Figure 4.31: Capacitor Parameter Definition

Figure 4.31 shows a Mathematica code fragment of a capacitor model definition. This model includes 3 fields, a nominal pad capacitance (`padcap`) that can be defined globally for the model. `capnum` is a multiplier that allows for the lumping of several pads and vias and `padtol` is a local variable that can be used to account for manufacturing tolerances and variations.

4.6.2.1.1.2 Probe Needles and Contactors

Probe needles and socket contacts are similar in that they generally consist of a short length of uncontrolled signal path, which tends to be inductive. Both of these elements are modeled using the self-inductance of a simple round wire. An empirical relationship is shown below[32]:

[32] Grover, 1946

$$L = \frac{len}{500}[\ln(\frac{4len}{dia}) - 1 + \frac{dia}{2len} + \frac{\mu_r T(f)}{4}] \text{ [Farads]}$$

where:

len	=	wire length [cm]
dia	=	wire diameter [cm]
T(*f*)	=	empirical correction for ac effects

For most cases, the third term can be dropped and the last term approaches 0.25 and the equation can be simplified to:

$$L = \frac{len}{500}[\ln(\frac{4len}{dia}) - 1] \text{ [Farads]}$$

■ **Probe Card Probe Needle (pcneedle)**

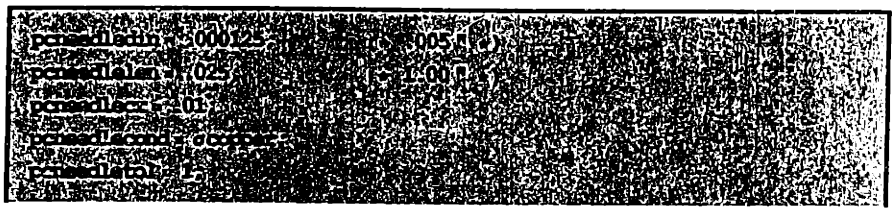


Figure 4.32: Probe Needle Parameter Definition

Figure 4.32 shows a code fragment of a probe needle model definition, which includes 5 fields. `din` and `len` are the geometric parameters of diameter and length. `cr` accounts for any contact resistance at the needle-pad interface, which is generally less than .1 ohm, but can be much greater due to dirt or improper scrub due to mechanical misalignment. `cond` is the material conductivity, which can be specified globally, and `tol` is a local variable that can be used to account for any manufacturing tolerances or other unmodeled variations.

4.6.2.1.2 Continuous Elemcnts

The distinction between discrete and continuous elements is a function of the operating frequency and modeling method. As discussed above, the physical length of any element should be short relative to maximum wavelength of interest, and if so, then it can be modeled as a lumped or discrete element. This is generally not the case in high speed digital signal delivery systems and often the signal path is several times the length of the fundamental signal. These transmission elements can be modeled by breaking them into pieces that are

small relative to the wavelength of interest and then treating them as lumped elements, as shown in Figure 4.33

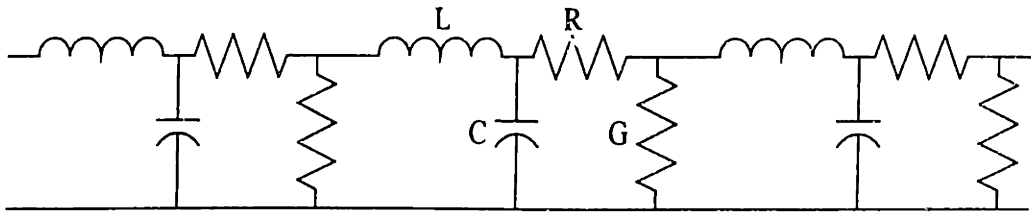


Figure 4.33: Lumped Element Transmission Line Model

The closed form solution of the impedance is:

$$Z = \sqrt{\frac{Z}{Y}} = \sqrt{\frac{R + j\omega L}{G + j\omega C}}$$

$$\gamma = \sqrt{ZY} = \sqrt{(R + j\omega L)(G + j\omega C)}$$

where:

L = Inductance per unit length
C = Shunt capacitance per unit length
R = Resistance per unit length
G = Admittance per unit length
Z = Characteristic impedance of transmission line
γ = Propagation constant

These two equations result in complex expressions for both characteristic impedance and propagation constant. Propagation constant is often separated into its real and complex components as follows:

$$\gamma = \alpha + j\beta$$

where:

α = Attenuation constant [dB per unit length]
β = Phase constant [radians per unit length]

A common simplification is to assume that the series resistance and shunt admittance of the line are small compared to the inductive and capacitive reactances. This results in real expressions for impedance and the loss propagation information A further simplification is to describe the attenuation constant in terms of the physical loss phenomenon as follows:

$\alpha = \alpha_c + \alpha_d + \alpha_r + \alpha_l$, where
α_c = conductor losses
α_d = dielectric losses
α_r = radiation losses
α_l = leakage losses

In practice, a well designed transmission line will have negligible radiation and leakage losses. However, at high frequencies, conductor and dielectric losses can dominate the performance of a transmission line.

At high frequencies, current tends to flow at the surface of a conductor rather than in the bulk material. This skin effect reduces the effective cross sectional area of the conductor and thus raises it's overall resistance.

$$R_{skin} \propto \sqrt{\frac{\omega}{cond}} SA$$

where:

cond = conductivity of the material
SA = perimeter enclosing the conductor cross section
ω = frequency

This frequency dependent change in resistance will affect both the attenuation and the characteristic impedance.

Dielectric losses are due to eddy currents in the dielectric material and occur as the wavelength of the signal approaches the dielectric thickness. Therefore, smaller diameter cable is capable of propagating higher frequency signals. The cutoff frequency, defined in terms of wavelength, can be approximated using:

$$\lambda_c = \frac{\pi(D+d)}{2}$$

For the purposes of this model, the dielectric loss is curve-fit using the cutoff frequency and empirical data. This is accomplished by assuming that it behaved as a first order low pass filter with the -3dB point at the cutoff frequency. This response cascaded with the

insertion loss (S_{12}, S_{21}) during the S-matrix generation described below. The combination of conductor and dielectric losses bound the performance and geometry of most transmission lines.

4.6.2.1.2.1 Coaxial Transmission Elements

Coaxial elements are modeled from first principles by calculating the coaxial inductance and capacitance per unit length as follows:

$$L_{coax} = \frac{\mu_o}{2\pi}\log(\frac{Dout}{Din})$$

$$C_{coax} = \frac{2\pi\varepsilon_o\varepsilon_r}{\log(\frac{Dout}{Din})}$$

where

ε_o, ε_r = *Permittivity (dielectric constant), relative dielectric constant*

μ_o = *Permeability*

Higher performance coaxial cables often use expanded PFTE (teflon) for a dielectric. This combines the low relative dielectric constant of teflon (2.1) with a partial ratio of air dielectric, resulting in a net constant between 1 and 2.1. Since this value is often not published, a nominal impedance of 50 ohms is assumed and the dielectric constant is backed out of the standard impedance calculation:

$$Z_o = \sqrt{\frac{L_{Coax}}{C_{coax}}} = \sqrt{\frac{\frac{\mu_o}{2\pi}\log(\frac{Dout}{Din})}{\frac{2\pi\varepsilon_o\varepsilon_r}{\log(\frac{Dout}{Din})}}} = \sqrt{\frac{\mu_o}{4\pi^2\varepsilon_o\varepsilon_r}(\log(\frac{Dout}{Din}))^2}$$

■ Channel Card Coax Cable RG-178/U (cccoax)

```
cccoaxdout = .00137;          (* .054 " *)
cccoaxdin = .00051;           (* .020 " *)
cccoaxlen = 1;                (* 4.00 " *)
cccoaxdiel = 1.406;
cccoaxcond = osilver;
cccoaxtol = 1;
```

Figure 4.34: Coaxial Cable Model Parameter Definition

Figure 4.34 shows a code fragment of a coaxial element model definition, which includes 6 fields. `dout`, `din` and `len` are the geometric parameters of outer and inner diameters and length. `diel` is the relative dielectric constant calculated above, cond is the material conductivity. As shown above, silver is often plated over copper to decrease the skin effect. `tol` is a local variable that can be used to account for any manufacturing tolerances or other unmodeled variations. The cutoff frequency is calculated from the difference in inner and outer diameters.

4.6.2.1.2.2 Microstrip and Stripline Models

Microstrip and stripline models are similar to that of coaxial cable. These structures also exhibit the same high frequency conductor and dielectric losses. The only significant difference between these structures and a coaxial cables is the geometry and the cutoff frequency. The empirical equations given in section 4.3.2 above were used to calculate the characteristic impedances of the microstrip and stripline structures.

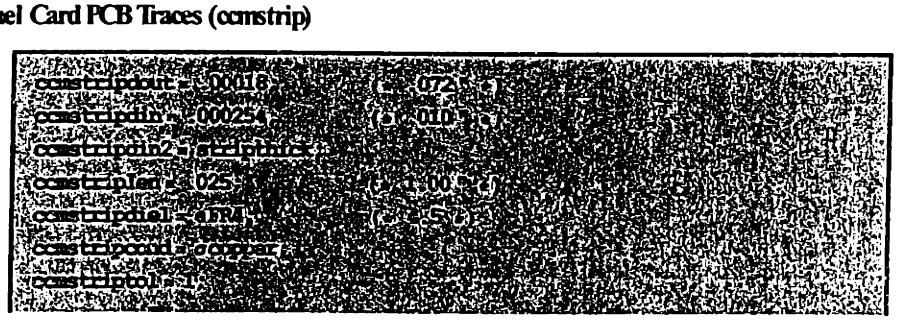


Figure 4.35: Microstrip Model Parameter Definition

Figure 4.35 shows a code fragment of a microstrip model definition, which includes 7 fields. `dout`, `din`, `din2` and `len` are the width, dielectric thickness, trace thickness and length of the microstrip trace. `diel` is the relative dielectric constant of the PCB laminate and `cond` is the material conductivity. `tol` is a local variable that can be used to account for any manufacturing tolerances or other unmodeled variations. The cutoff frequency is calculated using `din`, the dielectric thickness, as the characteristic dimension.

4.6.2.1.3 Spring Pin Models

Pogo pin arrays can be designed in a number of different configurations. The three most common designs are coaxial, five wire and planar pair. All three of these are similar in that there is a single signal pin in close proximity to a single or multiple ground pins, all of which are separated by a dielectric material. These three configurations are shown in Figure 4.17 above.

Each of these configurations has a different constitutive relationship between the pin geometry, layout and dielectric material and that they have a short section of uncontrolled pin near the tip which tends to be inductive. All of the pogo pin models are therefore broken into two sections; the body, which is controlled impedance and the pin, which is not. In the case of the planar pair and 5 wire pogo arrays, both of these sections use the same impedance model, but the body is designed such that its impedance is nominally 50 ohms based on the supporting dielectric material. The exposed pins have and air dielectric and therefore will tend to be inductive. The coaxial pogo array uses the coaxial element model described in 4.6.2.1.2.1 above for the pogo body and a simple 2-wire inductor model for the exposed pins.

■ Channel Card Pogo Pin, Coax Barrel (ccpogo)

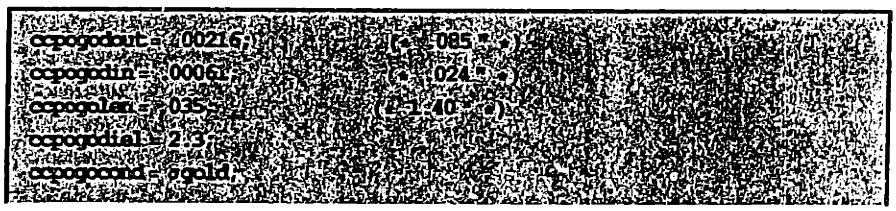


Figure 4.36: Coaxial Pogo Body Model Parameter Definition

Figure 4.36 shows a code fragment of a coaxial pogo body model definition. This is essentially the same coaxial model discussed above, with different geometry and dielectric constant. In this case, the cutoff frequency, calculated using the difference in inner and outer radii, is somewhat higher than standard coaxial cable due to the superior dielectric performance of air.

■ Channel Card Pogo Pin, Exposed Needle (ccpin)

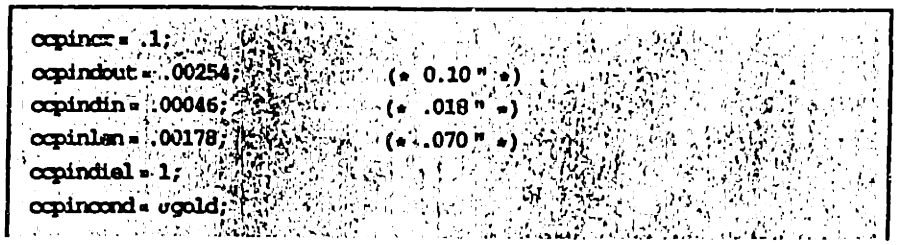


Figure 4.37: Exposed Pogo Pin Model Parameter Definition

Figure 4.37 shows a code fragment of an exposed pogo pin model definition, which includes 6 fields. `dout`, `din` and `len` are the separation, diameter and length of the pogo pins. `diel` is the dielectric constant of the surrounding air, which is usually equal to 1.0, but can be varied to account for humidity temperature changes or the presence of purge gasses.. `cond` is the material conductivity and `tol` is a local variable that can be used to account for any manufacturing tolerances or other unmodeled variations.

4.6.2.2 S-Matrix Models

The set of constitutive equations and parameters described above can be used to build a complete model of any general signal transmission system. To do this, they need to be transformed from scalar (real or complex) impedances to the scattering, or other matrix form. There are two basic cases of this tranformation; lumped elements and continuous transmission elements. In addition to the element impedances calculated above and the length in cases of continuous elements, the nominal or design impedance of the transmission system must be specified.

Figure 4.38 shows the scattering matrix formulations for lumped shunt and series impedances.

ELEMENT	S	
Z	$\frac{Z}{Z+2Z_0}$	$\frac{2Z}{Z+2Z_0}$
	$\frac{2Z}{Z+2Z_0}$	$\frac{Z}{Z+2Z_0}$
Y	$\frac{-Z_0Y}{2+Z_0Y}$	$\frac{2}{2+Z_0Y}$
	$\frac{2}{2+Z_0Y}$	$\frac{-Z_0Y}{2+Z_0Y}$

Figure 4.38: Lumped Impedance Scattering Matrix Formulation

Where

Z: Modeled impedance of component

Zo: Design nominal impedance of line (50 ohms)

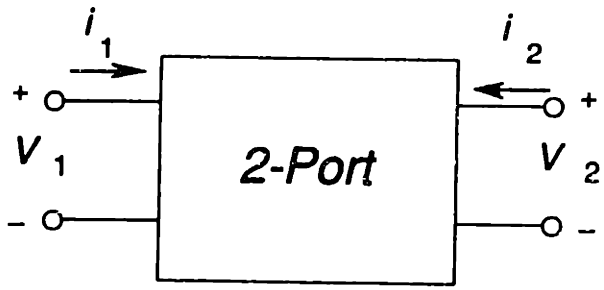
Figure 4.39 shows the same two cases for a continuous element transmission line. The general case includes a complex propagation constant, γ and includes both attenuation and phase losses. As described above, only the real portion of the impedance is considered and the attenuation and transmission delay are approximated through other means. Because these losses are tracked elsewhere, the simpler, lossless transmission line case can be used.

Z_0, l, General	$\frac{\Gamma(1-e^{-2\gamma l})}{1-\Gamma^2 e^{-2\gamma l}}$	$\frac{(1-\Gamma^2)e^{-\gamma l}}{1-\Gamma^2 e^{-2\gamma l}}$
	$\frac{(1-\Gamma^2)e^{-\gamma l}}{1-\Gamma^2 e^{-2\gamma l}}$	$\frac{\Gamma(1-e^{-2\gamma l})}{1-\Gamma^2 e^{-2\gamma l}}$
Z_0, l, Lossless	$\frac{\Gamma(1-e^{-j2\beta l})}{1-\Gamma^2 e^{-j2\beta l}}$	$\frac{(1-\Gamma^2)e^{-j\beta l}}{1-\Gamma^2 e^{-j2\beta l}}$
	$\frac{(1-\Gamma^2)e^{-j\beta l}}{1-\Gamma^2 e^{-j2\beta l}}$	$\frac{\Gamma(1-e^{-j2\beta l})}{1-\Gamma^2 e^{-j2\beta l}}$

Figure 4.39: Continuous Transmission Line Scattering Matrix Formulation

4.6.2.3 Transformations and Multiplication

Although the scattering parameter representation is capable of capturing all of the relevant phenomenon of signal transmission systems, it suffers from one minor fault. Due to the definition of input and output, S-matrices cannot be simply cascaded using matrix multiplication. Several formulations can be used to circumvent this, the most common are the use of intermediate matrix definiations such as the scattering transfer matrix (T-matrix) or the ABCD matrix. These two matrix types have different definitions, but functionally are similar and are equally valid as the scattering representation. The ABCD forumulation is used in this model and is shown in Figure 4.40 below.



$$v_1 = Av_2 - Bi_2$$

$$i_1 = Cv_2 - Di_2$$

Figure 4.40: ABCD Matrix Definition

In addition to the definition of the ABCD matrix, a transformation between the ABCD and S formats needs to be defined. The equations for this transformation are shown below.

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} \frac{(1+s_{11})(1-s_{22})+s_{12}s_{21}}{2s_{21}} & \frac{(1+s_{11})(1+s_{22})+s_{12}s_{21}}{2s_{21}} \\ \frac{(1-s_{11})(1-s_{22})+s_{12}s_{21}}{2s_{21}} & \frac{(1-s_{11})(1+s_{22})+s_{12}s_{21}}{2s_{21}} \end{bmatrix}$$

$$\begin{bmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{bmatrix} = \begin{bmatrix} \frac{A+B-C-D}{A+B+C+D} & \frac{2.0\,(A\,D-B\,C)}{A+B+C+D} \\ \frac{2.0}{A+B+C+D} & \frac{-A+B-C+D}{A+B+C+D} \end{bmatrix}$$

4.6.2.4 Parameter and Model Definition

The model elements defined above are sufficient to create a complete, symbolic description of a signal transmission system. The final modeling step in making a useful model involves defining the transmission path and elements. This includes the global variables, model layout and the local element variables.

Any variable in the model can be defined either locally or globally. Constants and some material properties such as the speed of light and permitivitty and permeability of free space are best defined globally. In addition to this, it is convenient to globally define material properties such as relative dielectric constants and conductivities, design parameters such as pad and via capacitances and manufacturing tolerance values. This global definition enables these values to be manipulated for the purpose of sensitivity studies. For example, the permitivitty of some dielectric materials is sensitive to temperature and the relationship between system performance and temperature can be evaluated by globally adjusting this variable. Also, the effect of PCB manufacturing tolerances can be evaluated by globally varying the trace or dielectric thickness. Examples of these studies are shown in Chapter 5.

The second step in building a model is to define the logical layout of the transmission elements. Figure 4.41shows a generic transmission path that might represent the delivery of a signal from the instrument driver to the DUT.

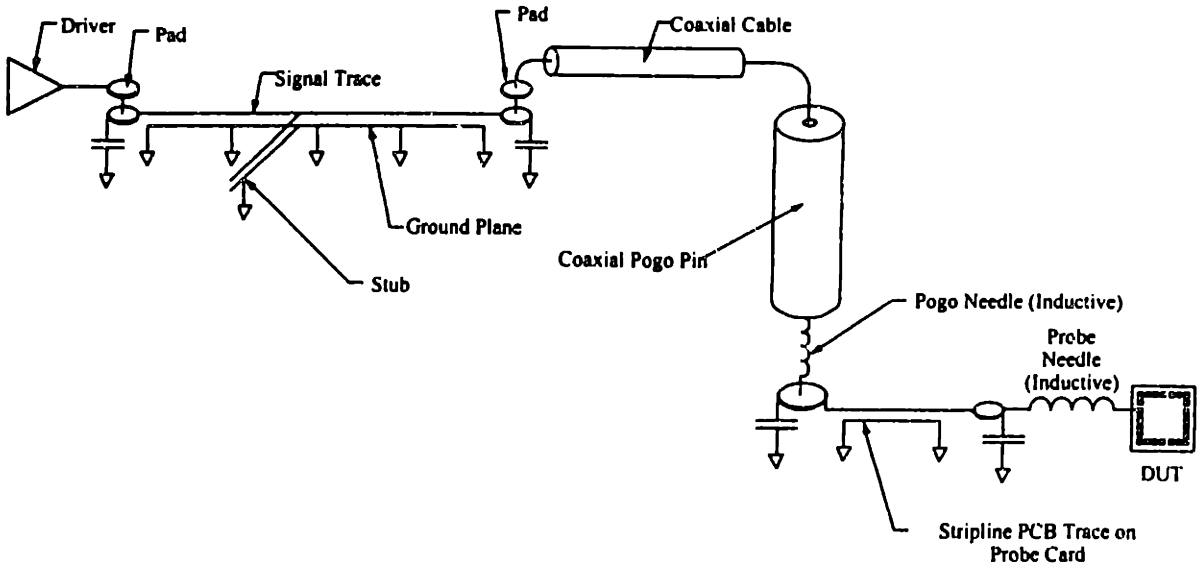


Figure 4.41: Schematic of Example Model

Figure 4.42 shows the same system in block diagram form.

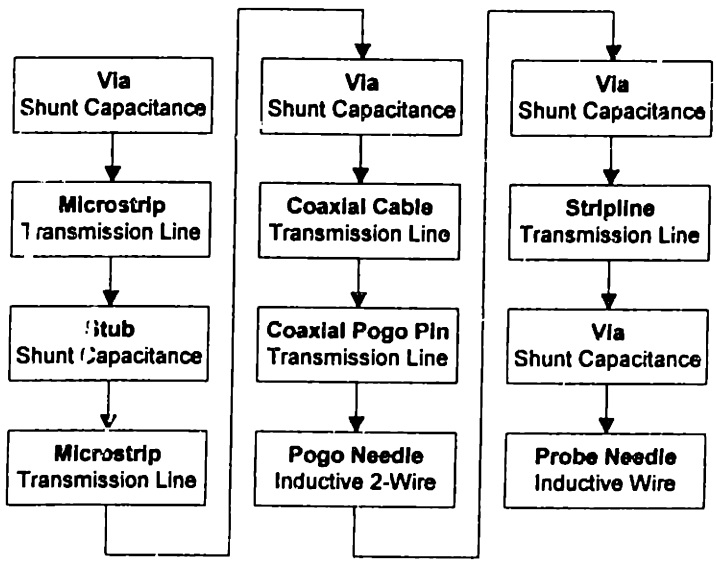


Figure 4.42: Example Transmission Path Model

This model can be simplified by combining some of the neighboring capacative elements, such as the stub and vias near the driver and the pair of pads on the probe card. A linear addition of these capacitances will produce a conservative bandwidth error. Figure 4.43 shows this reduced model.

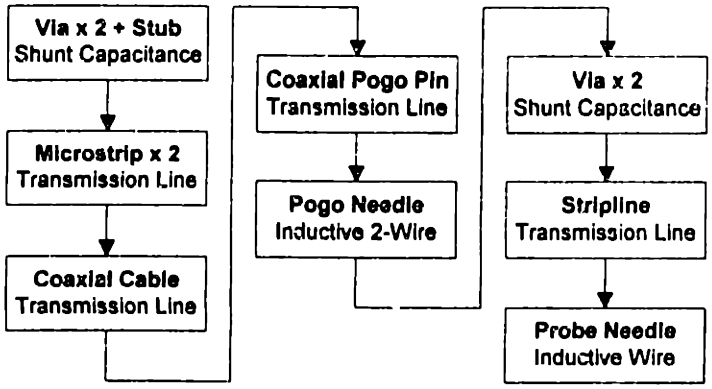


Figure 4.43: Reduced Example Transmission Path Model

This simplification also allows some insight into the effects of board complexity. The number of via interconnects is directly related to the PCB signal density due to routing problems. The relationship between this signal density and signal performance can be explored by scaling the number of vias per length of signal trace. Also, crosstalk effects can be considered by including the mutual inductance of signal lines.

The third step in building a model is the entry of the local element variables. Code fragments with these parameters were shown in Section 4.6.2.1 above. These blocks of local variables allow the modeling of almost any standard signal transmission configuration.

4.6.3 Application of the Model

The preceding sections describe all of the major elements of the transmission line analysis required for ATE systems. This framework of tools can be used to construct a model and predict the performance of typical signal delivery systems and test fixtures. The assumptions used to simplify the analysis limit the overall accuracy and usefulness of this model as an electrical analysis tool to approximately +/- 25%. However, the primary purpose of the model is not to analyze the exact performance of a specific design. Rather, it is intended to help gain insight into the relationships between the mechanical design and electrical performance of these systems. To this extent, the model is accurate and can be used to quickly perform a variety of analysis that formerly required much more complex modelling tools.

While the models and examples presented in this thesis do not capture all of the relevant phenomenon to completely model high frequency signal transmission, this tool is powerful

nonetheless. There are several types of useful analysis that can be performed well with theses tools. Preliminary comparisons of design concepts, design trade-off exercises and sensitivity studies are particularly well suited for this type of tool. Some examples of these analyses are presented in Chapter 5 below.

4.7 Summary of Chapter 4

This chapter discussed the importance and scope of high frequency signal delivery systems. This included the definition of transmission lines, an introduction to basic high frequency modeling theory and some background information on current design components. The balance of the chapter described the implementation of a model that is capable of evaluating the performance of a signal transmission system design. This model is intended to be used as a design tool in the development of high speed test fixtures. While the framework of the model is capable producing highly accurate results, simplifications in the current implementation limit the accuracy of this model. It is primarily useful as a tool to investigate design trade-offs and interactions between the mechanical design and electrical performance of signal delivery systems.

5 Results of Transmission Modeling

5.1 Introduction

The model framework developed in Chapter 4 is intended to be used as a tool in the design and development of signal delivery mechanics. To this end, the model is complete and is a useful tool. In addition to design oriented tasks, the model can also be used as a predictive tool to help gain insight into the current and future signal transmission needs as well as to explore relationships between mechanical design and electrical performance parameters.

This chapter discusses the results of several analyses using the above model that were formed within the context of the device roadmaps. These analyses were chosen to explore the limits of current technologies to understand where future development efforts should be focused to address these issues. The chapter begins with a brief discussion of current and future ATE signal transmission requirements that are based on the needs of the system on chip (SOC) market. This is followed by a detailed description of a baseline model that is derived from an existing state-of-the-art ATE design. Experimental and model results are compared to verify the accuracy of the model. Finally, some thoughts on future enhancements to the models are given.

5.2 Signal Transmission requirements

The device roadmap discussed in Chapter 2 describes the predicted needs of the semiconductor industry over the next 15 years. For this data to be useful for ATE architectural development, it needs to be reinterpreted within the context of a specific test application. From a mechanical systems and signal transmission viewpoint, the device that poses the greatest test challenge is that which requires both high performance and a large number of resources. The most challenging application from this perspective is currently the multimedia processor, which is used in computers and video equipment, that combine digital signal processing and a wide digital bus with analog inputs and outputs to allow direct communication with the outside world. These devices are the precursors to the true systems

on a chip (SOC) which are only now in their infancy[33]. SOC's will combine the analog functionality of DSP and cellular devices with the power of modern microprocessors and deep arrays of memory. One of the primary reasons for the combination of these functions on a single device is the need to reduce signal transmission problems. To the extent that this integration helps device performance, it also increases the difficulty of designing a test system.

To capture the scope of current and future test system architectural requirements it is not necessary to model the signal performance of every device I/O. Rather, it is sufficient to focus the analysis on the vital few test challenges, in particular, the growing number of digital pins that are running at the external bus speed. Ignoring the balance of analog, clock, power and ground pins is a valid assumption for several reasons.

Analog I/O can be ignored because their number is usually small in proportion to the digital I/O pins. For example, a typical 24 bit DAC will have a 24 to 48 line digital input bus for every analog output line. This ratio may vary based on accuracy, multiplexing and whether single or dual ended I/O's are used, but in general the ratio is at least 10:1. Clock lines are much more difficult to design and control because they can run at speeds up to 10 times faster than the data bus. However, because there is generally only a single clock line, it represents a minor challenge relative to the thousands of digital pins. Finally, power and ground supply pins may potentially outnumber the digital pins on future, low voltage devices. Some estimates predict that future devices will require a ground reference for every digital pin, essentially making all lines differential pairs. While this represents a packaging and layout challenge, the DC nature of these lines does not pose a great difficulty to resolve. One way to include their impact would be to double the mechanical volume of signal lines when modeling a digital bus.

The above assumptions can be combined with the SIA roadmap to create a model application scenario that describes how the critical device parameters will change over the next 15 years (1997-2012) Within the context of current multimedia and future SOC devices, the parameters of concern for ATE system design are as follows[34]:

---

[33] Walsh, 1997

[34] Source: preliminary 1997 SIA Roadmap

- The number of digital pins. This will increase from a current value of 200 to a maximum of 2000.
- The external bus speed. This will increase from the current value of 100MHz to a maximum of 600 MHz
- The signal accuracy requirement. It is assumed that, as a function of frequency, this requirement will remain unchanged.

The growing number of digital pins can be directly related to the growing bit depth of microprocessors and number of I/O busses. Because of mechanical packaging and PCB routing constraints, it is unlikely that this number will increase significantly beyond the predicted value of 2000. There are some emerging memory technologies, such as Rambus and Synclinc that utilize high speed serial busses between processors and memory which are currently running a 400 MHz bus and expect to be at 800 MHz by 1998. These technologies do use lower pincounts that conventional wide memory busses, but despite this, they may represent a push towards increased bus speeds and thus make testing more difficult that this model predicts. There are no known drivers to increase the accuracy of digital bus signals. There are many ways to specify and measure accuracy, including edge rate or rise time and insertion loss. For the purposes of this model, the –3dB bandwidth of the 7th harmonic will be considered.

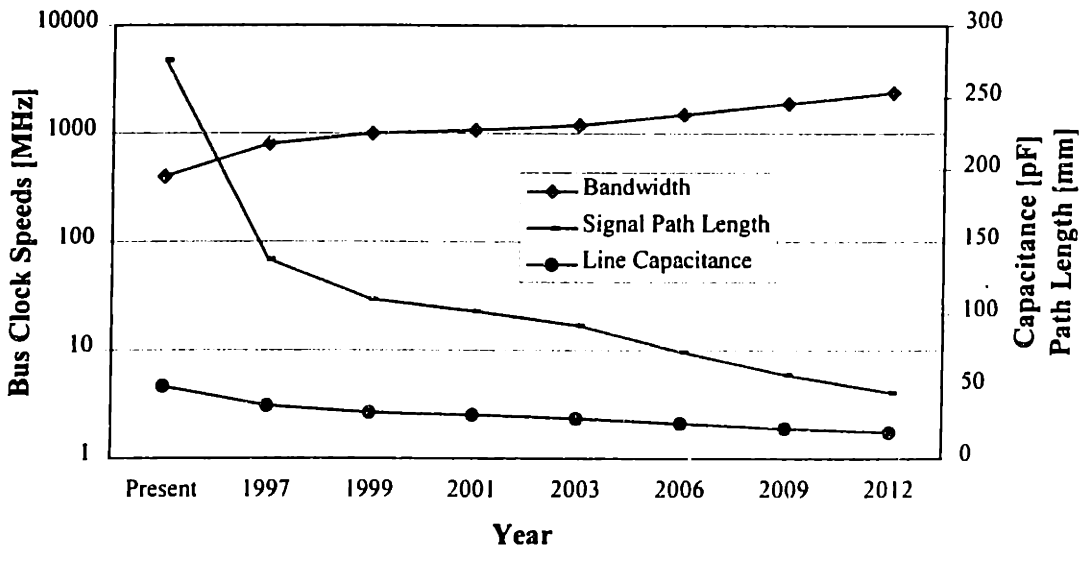


Figure 5.1: Transmission Line Performance Need Roadmap

Figure 5.1 translates these roadmapped device parameters into the signal transmission performance requirements of propagation delay, line capacitance and bandwidth. These are the three measures of transmission line performance that are tracked by the modeling tools developed in Chapter 4. The goal of this analysis is to determine whether existing technology is capable of meeting these roadmapped needs and if not, to understand which areas offer the greatest challenge to continued progress.

5.3 Baseline System Model

Using the framework described in Chapter 4, a model representing a test fixture signal transmission system was developed. This baseline system is derived from a typical 100 MHz digital instrument and available test fixturing elements. Performance of this system is explored using the model and compared to measurements of physical instruments to verify the models veracity. Once a valid system model has been developed, various scenarios can be run to explore the effects of the mechanical parameters, including the effects of mechanical docking repeatability, manufacturing tolerances and temperature changes. The signal transmission system described using existing components will likely not be sufficient to meet the needs of future technologies. Therefore, an additional set of design experiments explore the effects of changing and eliminating various components within the transmission path to see if a practical scenario exists, using current transmission components, that will meet the needs of future devices.

The baseline model, shown schematically in Figure 5.2, represents a typical probe application for a 100 MHz instrument.

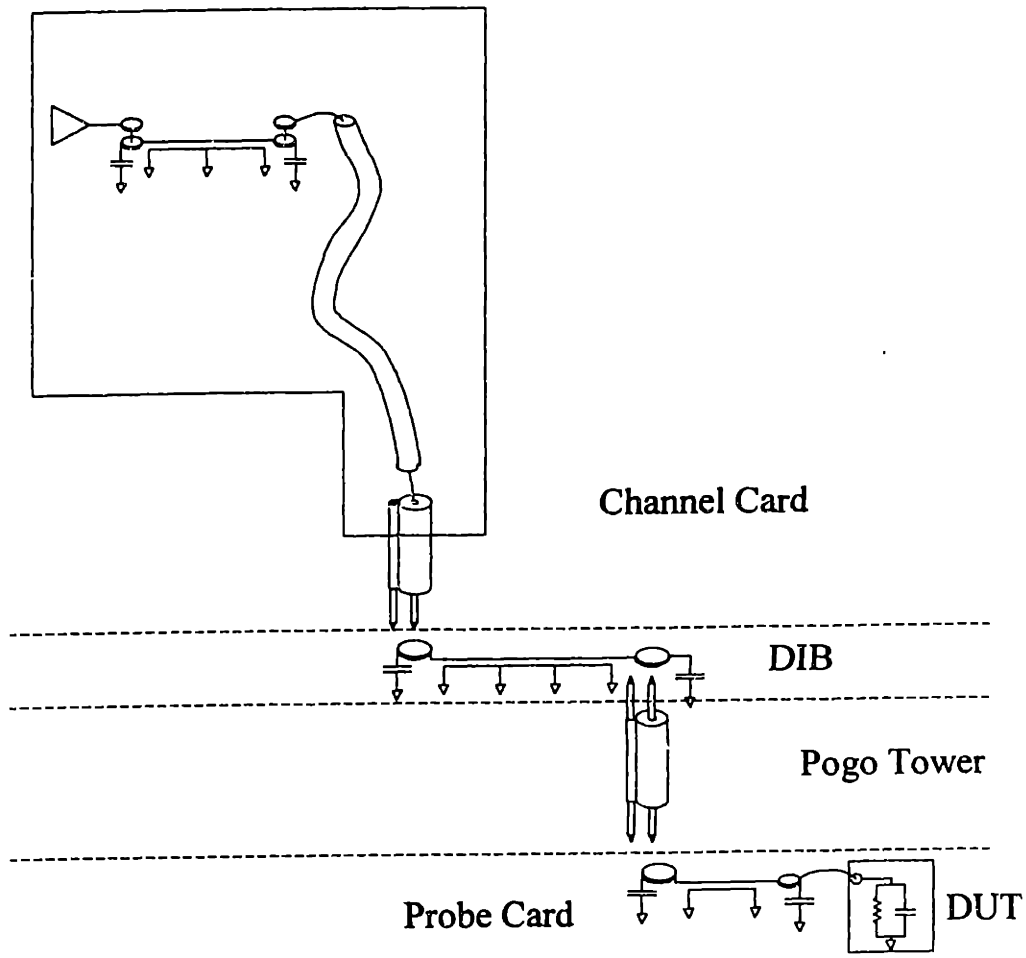


Figure 5.2: Schematic of Baseline System Model

This model is divided into the following four subsystems, which are discussed below:

- Channel card: PCB that contains the instrument drivers
- Device Instrument Board (DIB): PCB fixture that is used to configure tester
- Pogo Tower: Spring pin connection between DIB and probe card
- Probe Card: Fixture that contains the needles that probe the wafer

5.3.1 Channel Card

The channel card contains the drive and receive circuits and is the last stage of the digital instrument. Unless additional drive electronics are placed on the DIB or probe card, the channel card is the last opportunity to amplify and condition the output signal and it is the nearest point to which the DUT must drive its output.

Figure 5.3 shows a schematic of the signal delivery portion of a channel card, which contains the following transmission elements that must be modeled:

- Microstrip PCB trace from the driver package to the coaxial cable launch including capacitate vias and device solder pads
- A length of coax cable
- A spring pin connection block including a coaxial body and a 2 wire inductive contact pin.

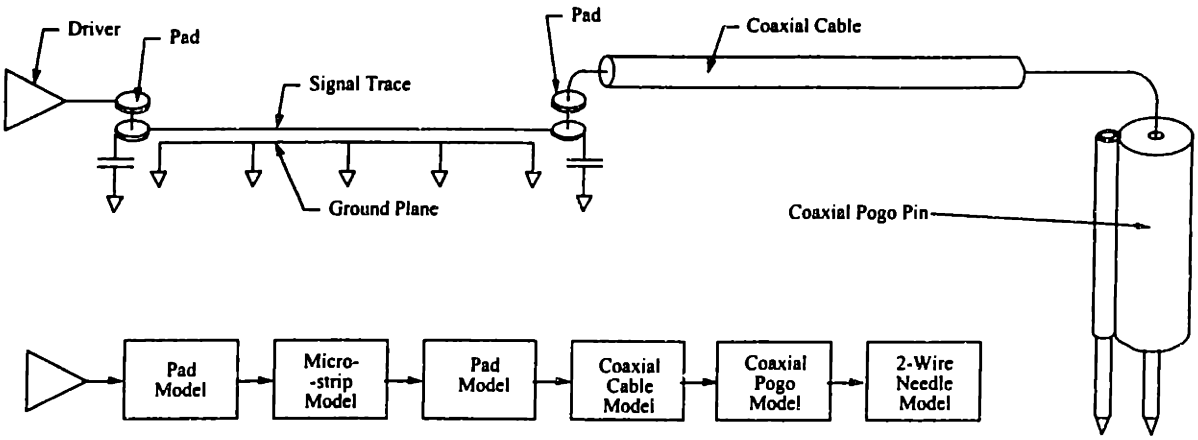
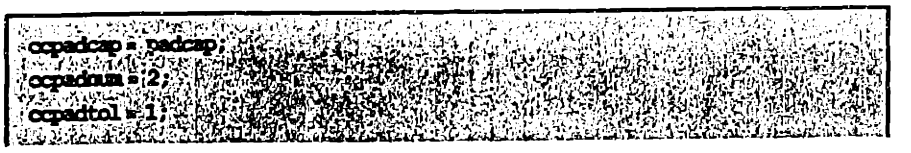


Figure 5.3: Channel Card Model

Some assumptions were made about the above components to simplify analysis. The microstrip line was assumed to stay on the surface layer of the PCB and no stubs were included for the relays between the driver and cable. Microstrip geometry was approximated so that the characteristic impedance was nominally 50 ohms. The solder pads of the driver and of the cable launch were included, but because the simple routing on the board, vias were ignored. Parameters for the channel card PCB and vias are shown in Figure 5.4 below.

■ Channel Card Pads (ccpad)



■ Channel Card PCB Traces (ccmstrip)

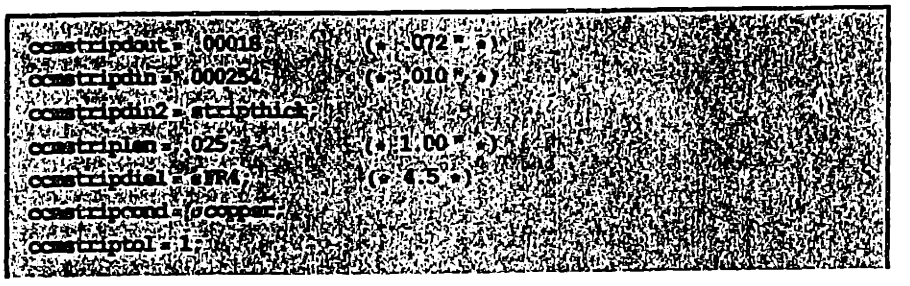


Figure 5.4: Channel Card PCB Parameters

Figure 5.4Data for the RG-178/U coaxial cable used in this design were taken from a manufacturers data sheet. The dielectric used in this cable is an expanded PFTE and therefore is between that of Teflon and air. This value was calculated so that the characteristic impedance of the cable was nominally 50 ohms. These parameters are shown in Figure 5.5.

■ Channel Card Coax Cable RG-178/U (cccoax)

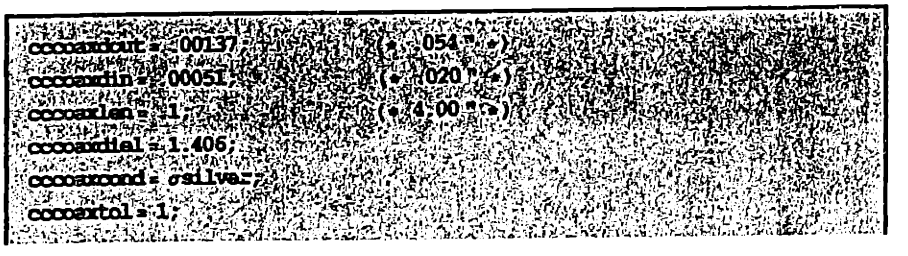


Figure 5.5: Channel Card Coaxial Cable Parameters

The coaxial cable is soldered directly to channel card spring pin connector assembly. This assembly is modeled with two sets of parameters; the coaxial spring pin that carries the signal is modeled using the standard coax cable equations. The exposed pair of signal and

ground pins are modeled using the inductive wire model. Parameters for these components are shown in Figure 5.6.

■ Channel Card Pogo Pin, Coax Barrel (ccpogo)

```
ccpogodout = .00216;          (* .085" *)
ccpogodin = .00061;           (* .024" *)
ccpogolen = .035;             (* 1.40" *)
ccpogodiel = 2.3;
ccpogocond = dgold;
```

■ Channel Card Pogo Pin, Exposed Needle (ccpin)

```
ccpincr = 1;
ccpindout = .00254;           (* 0.10" *)
ccpindin = .00046;            (* .018" *)
ccpinlen = .00178;            (* .070" *)
ccpindiel = 1.0;
ccpincond = dgold;
```

Figure 5.6: Channel Card Spring Pin Assembly Parameters

5.3.2 Device Interface Board

The primary purpose of the DIB is to provide some area between the channel card instrumentation and the DUT to configure and customize the transmission path. For linear and mixed signal devices, this is an important feature and DIBs for these devices are generally complex and difficult to model. However, with the exception of decoupling capacitors, digital DUTs generally do not require a significant number of devices in the transmission path and the DIB is used primarily as a spatial transformer between the geometry of the channel card pins and that of the pogo tower. As digital pin counts increase, the signal routing density will approach the point that even discrete capacitors will be too large to fit on the DIB and routing optimization will be required to eliminate excessive jumps between signal planes. For the purposes of the baseline model, a simple microstrip trace was used whose length was determined by the geometry of the channel card, pogo tower. The microstrip was initiated and terminated with pads for spring pin contact and a number of vias and additional path length were added to accommodate for routing difficulties. This model is

shown in Figure 5.7, the parameters of this and the following models can be found in Appendix 9.1.

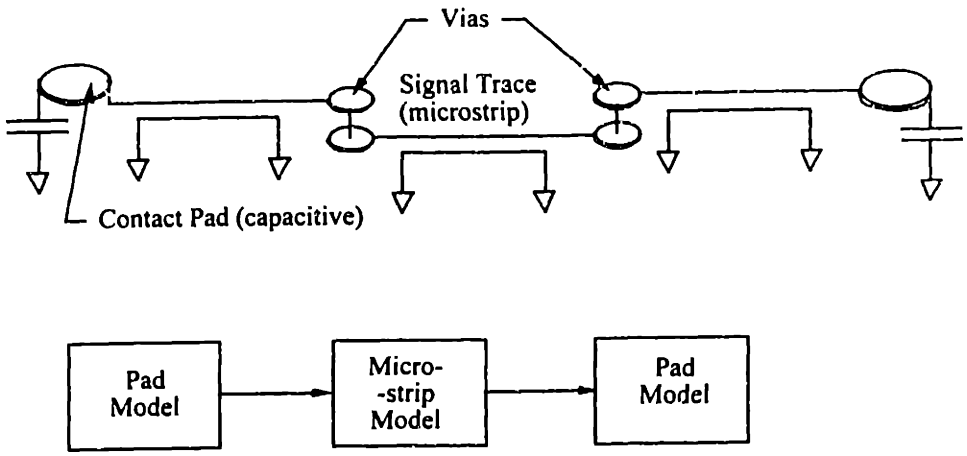


Figure 5.7: DIB Model Schematic

5.3.3 Pogo Tower

The pogo tower used for the baseline model is shown in Figure 5.8. The model consists of three blocks; a pair of signal/ground pins, followed by a coaxial body and another pair of signal/ground pins.

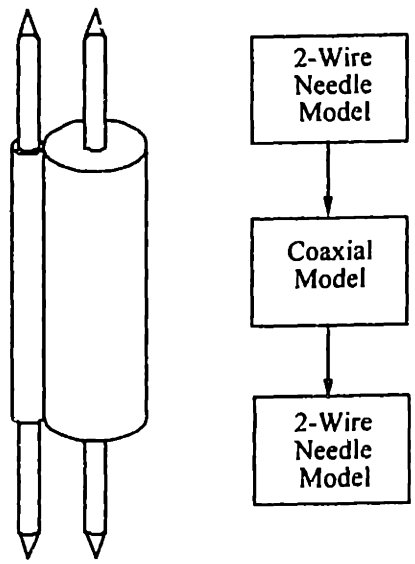


Figure 5.8: Coaxial Pogo Tower Model

5.3.4 Probe Card

The probe card is similar to the DIB in that it is simply a PCB with microstrip, contact pads and vias. The only difference is that the probe card has an additional probe needle. The needle is modeled as a bare wire and therefore is a pure inductance. Generally, this length of wire can cause significant ground bounce and therefore a matching bypass capacitor is often added at the base of the needle to help normalize the line. This capacitance is added to the total line capacitance. The probe card model is shown in Figure 5.9.

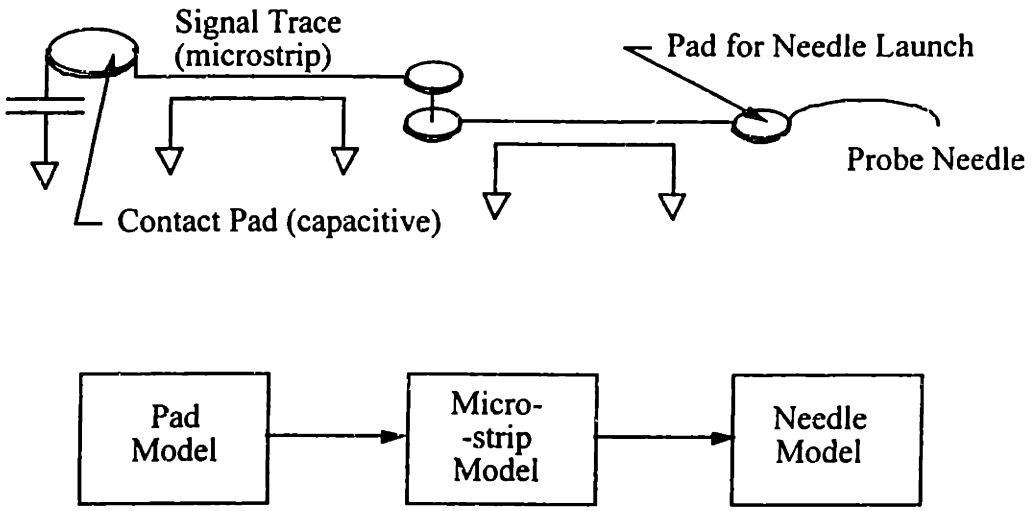


Figure 5.9: Probe Card Model

5.3.5 Line Termination

Some assumption must be made about the termination of the line at the DUT. If the line is left open, the impedance is essentially infinite and the input signal will be completely reflected. The model assumes a nominal 50 ohm load for the impedance of the DUT. It is rare that an actual device will have a true 50 ohm impedance, generally input impedances are much higher to limit current draw and are somewhat capacitive. In extreme cases, additional discrete components will be added at the base of the probe needle or on the needle itself to compensate for this mismatch. This value can be changed in subsequent models to simulate the testing of devices with unmatched impedance.

5.4 Analysis Performed on Baseline Model

The above baseline model components constitute a complete signal delivery system from the instrument driver to the DUT. These components can be cascaded as described in

Chapter 4 and the performance of the system and the effects of individual elements can be evaluated. Of primary concern are the following specifications:

- Overall propagation delay
- Input capacitance as viewed from the DUT
- Bandwidth as a function of insertion and return loss
- Higher order phenomenon
-

The model output of these parameters is discussed in greater detail below.

5.4.1 Propagation Delay

Propagation delay is tracked on a per element basis and summed as the elements are cascaded. Discrete elements such as pads and decoupling capacitors have no effective transmission delay and the delay of continuous elements is a function of the length and dielectric constant. The model outputs a scalar value for propagation delay. There is no specific target value for propagation delay, but smaller is always better. Current systems generally have approximately 3 nanoseconds of propagation delay, which approximately corresponds to a 300 MHz clock rate. As frequencies go beyond 300MHz, line lengths must decrease or data must be pipelined, further increasing test complexity.

5.4.2 Line Capacitance

Line capacitance is also tracked on a per element basis and summed as the elements are cascaded. With the exception of uncontrolled lengths of signal trace and bare wires, almost all transmission line elements have some significant capacitance. Discrete elements such as pads and decoupling capacitors have discrete capacitances and continuous elements such as coaxial cable and PCB structures have capacitance that is a function of the length and geometry. The model outputs a scalar value for the line capacitance. There is no specific target value for capacitance, this needs to be evaluated based on the specifications of the DUT. However, declining device output power dictates that lower capacitances are better. Current systems generally have approximately 50 to 100 picoFarads of line capacitance and if the capacitance exceeds the ability of the DUT to drive the line, then an additional buffer must be added to the transmission path to boost the output. However, this is not a feasible solution for high pin count digital devices because space and layout constraints will prevent the use of these devices on every line. While there is no roadmap for DUT output power or

test fixture line capacitance, it is reasonable to expect that these values will scale linearly with device power and supply voltage.

5.4.3 Bandwidth

Bandwidth is a system level attribute and, because of the interactions between the impedances of the different elements, it cannot be tracked on a per element basis. However, if the boundary impedances are assumed to be a nominal standard, such as 50 ohms, then a measure each components bandwidth can be approximated. In addition to the individual component bandwidths, a system measure can be calculated by cascading the individual component scattering matrices through the intermediary use of the ABCD matrix transformation. Bandwidth performance is measured as the –3dB point for insertion loss and the –6dB point for return loss. This represents the cutoff frequency at which half the power of the input signal is transmitted and 25 percent of the input signal is reflected at the element boundaries. The bandwidth cutoff frequency must be several times higher than the fundamental digital frequency to insure fast, accurate edges. The amount of information content decreases linearly with each successively higher harmonic and although there is no universally accepted standard, generally at least the first 5 and no more than the first 11 harmonics are required for signal fidelity. For the purposes of this study, it is assumed that the 7^{th} harmonic determines the highest bandwidth.

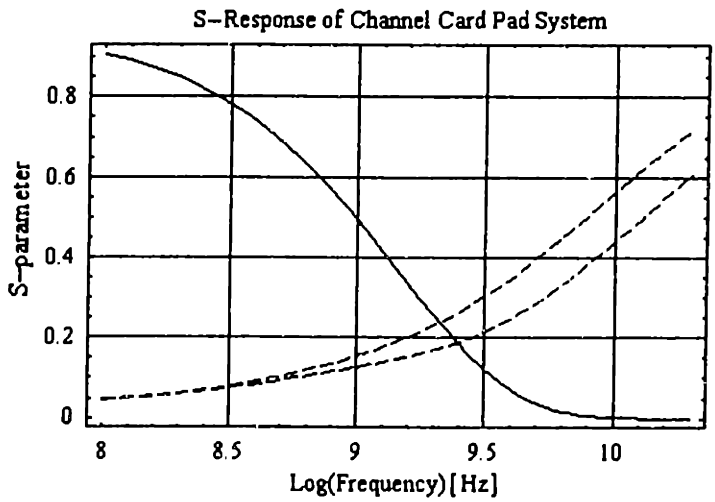


Figure 5.10: Sample Model Output

Figure 5.10 is a sample system S-parameter output plot. The three curves are the insertion loss (S12) and the forward and reverse return losses (S11, S22) The bandwidth is therefore the frequency at which the insertion loss –3dB or the return loss is –6dB. Because the model is completely symbolic, it's value can be solved for in closed form and in this case, the bandwidth is found to be 500MHz. This corresponds to a 70 MHz clock if the 1^{st} through 7^{th} harmonics are considered.

5.5 Results of Baseline Model Simulation

The baseline model is configured to simulate the design of the transmission system of a typical high frequency multimedia digital pin. Several simulations were run using this model to confirm its accuracy and to explore the systems sensitivity to various parameters. The following simulations were run on the baseline model:

- Nominal: all parameters are at their nominal values
- Junction impedance mismatch: all parameters are at nominal values except for the design impedance, which is increased to 60 ohms. This shows simulates the effect of mismatches at component boundaries, which is a phenomenon that the model cannot normally capture.
- Sensitivity analysis: Several of the parameters are varied to see if they have a significant effect on signal integrity and to demonstrate the ability of the model.

The above simulations are not meant to be an exhaustive study of the model, rather, they are meant to demonstrate the way the model can be used as an analysis tool. The results from these simulations are discussed below.

5.5.1 Nominal Scenario

A tremendous amount of insight can be gained just by examining the nominal system model. The most important result from this model is the verification of its results by comparing it to the actual system. For this initial version of the modeling program, the accuracy goal was to be able to simulate a system accurately to within a factor of two. This modest goal will result in a simulation that is accurate enough for basic design analysis, but not for design verification. In addition to accuracy verification, the following results can be derived from the nominal model:

- Insertion loss (S21) as a function of frequency
- Return loss (S11) as a function of frequency
- Capacitance between the DUT and instrument drivers
- Transmission delay between the driver and DUT

5.5.2 Model Verification: System Performance

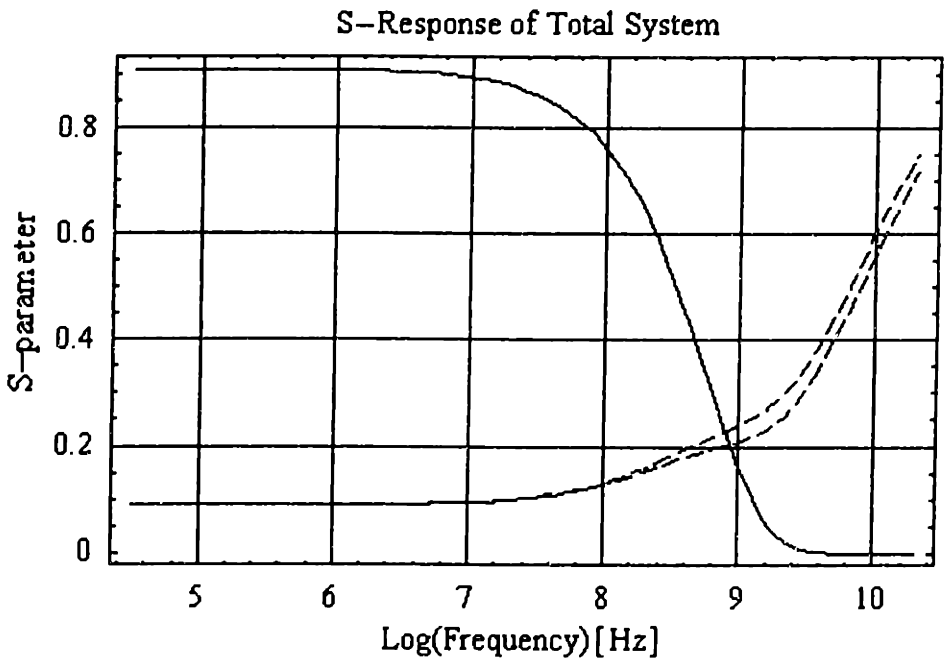


Figure 5.12: Signal Integrity for Nominal System

Figure 5.12 shows the model output for the total system. In this, and all subsequent graphs, insertion loss (S12) is shown as the solid line and return loss (S11) from either port is shown as the two dashed lines. As expected, insertion loss behaves like a low pass filter and

return loss increases with frequency. Because some of the model components are non-conservative, the sum of the insertion and return losses will not equal unity.

The curves in Figure 5.12 indicate that the bandwidth, or the S12 3dB down point, of this system is approximately 350 MHz and the return loss is about 10 % for frequencies below 1 GHz. Line capacitance and propagation delay, tracked separately, are 25 pF and 1.8 nS, which corresponds to approximately 250mm of PCB trace. These values are within a factor of 2 of the specification of the 100 MHz digital channel, on which this model is based, which requires approximately 500 MHz of bandwidth.

5.5.3 Signal Integrity: Insertion and Return Loss

The model can also be used to inspect the performance of the individual components within the system. Figure 5.13-Figure 5.17 show the performance of the transmission components on the channel card and Figure 5.18 shows the total performance of the channel card and a Pareto graph of the relative insertion loss contribution of each component.

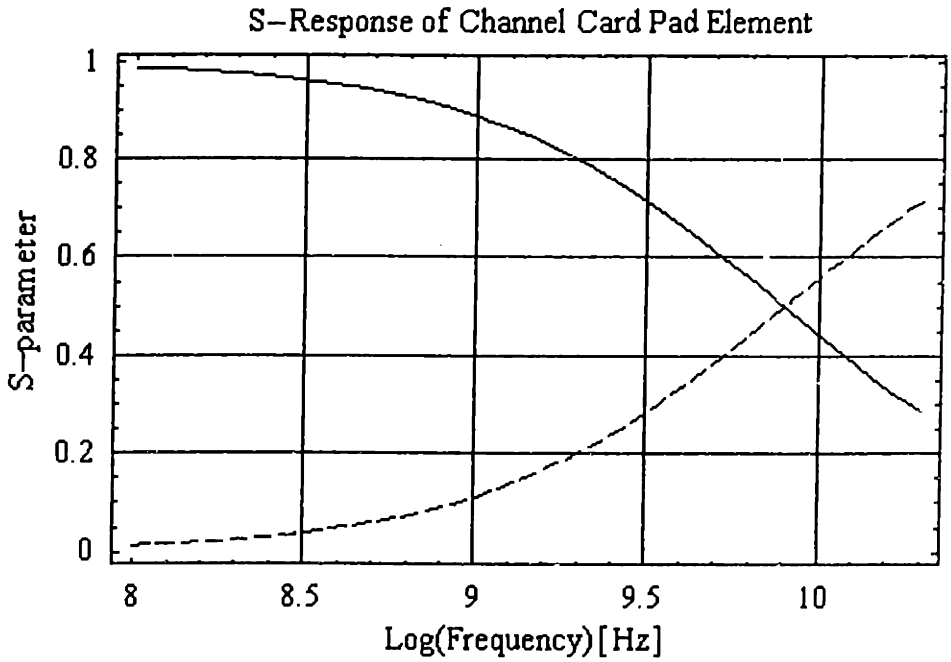


Figure 5.13: Channel Card Solder Pad Performance

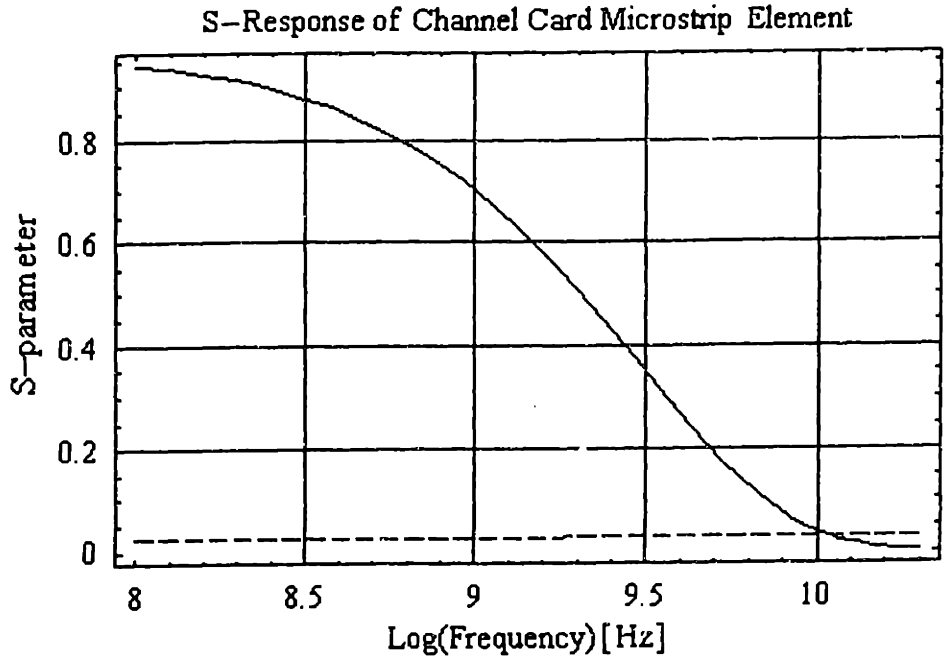


Figure 5.14: Channel Card Microstrip Performance

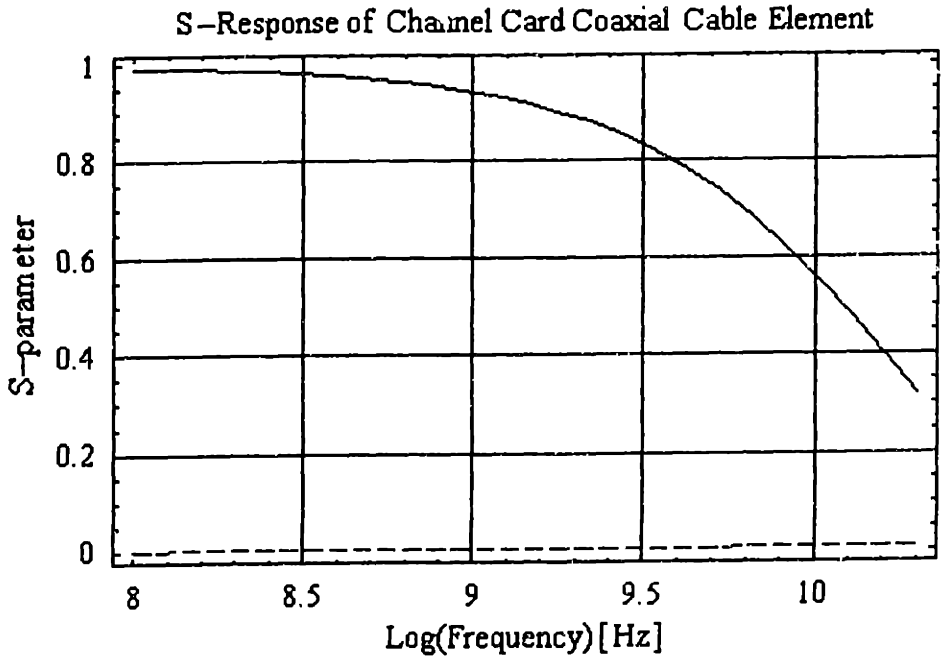


Figure 5.15: Channel Card Coax Cable Performance

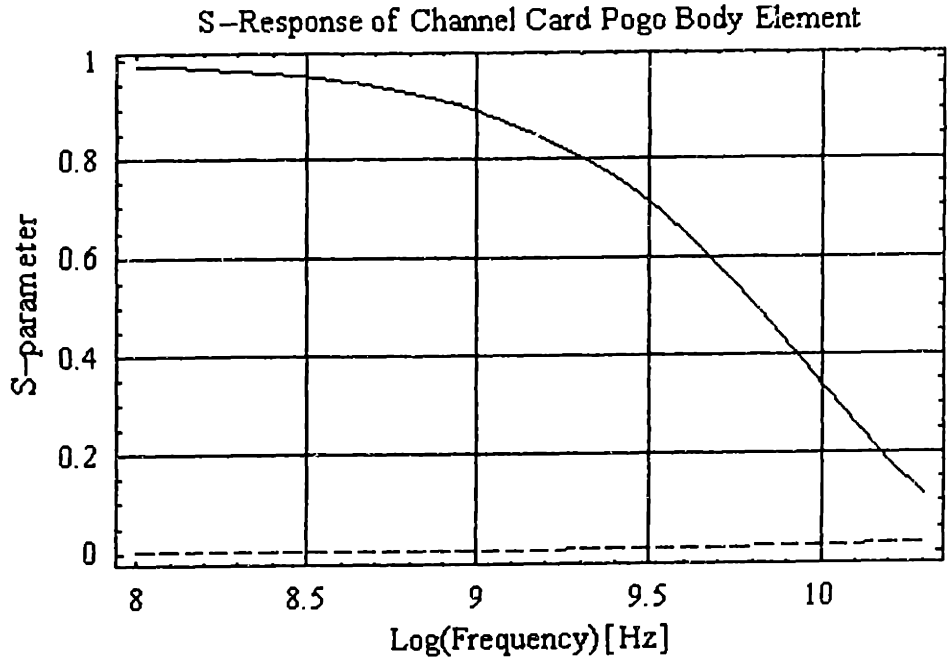


Figure 5.16: Channel Card Pogo Barrel Performance

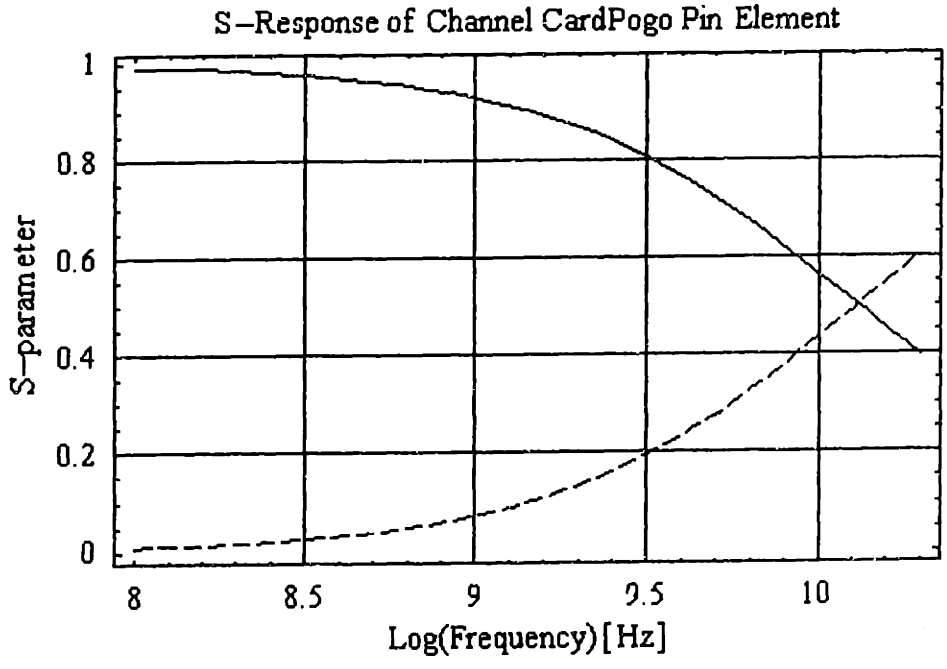


Figure 5.17: Channel Card Pogo Needle Performance

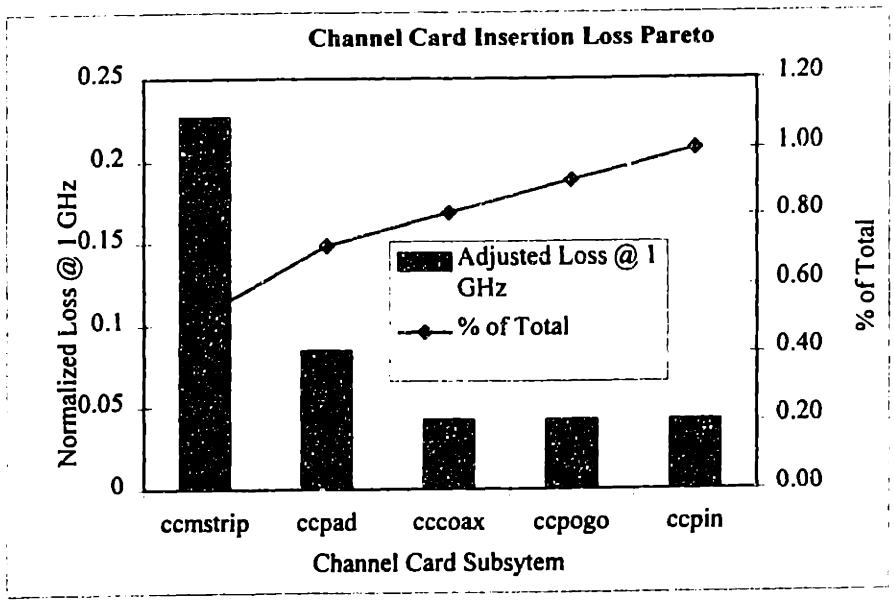


Figure 5.18: Channel Card Subsystem Performance and Pareto

Similar analyses can be performed on the other subsystems of the baseline transmission model. The performance of the channel card, DIB, pogo tower and probe card subsystems are shown in Figures 5.26-5.29 below, followed by a Pareto graph of the total system performance broken down by subsystem in Figure 5.23. The detailed results from the subsystem components can be found in Appendix 9.1.

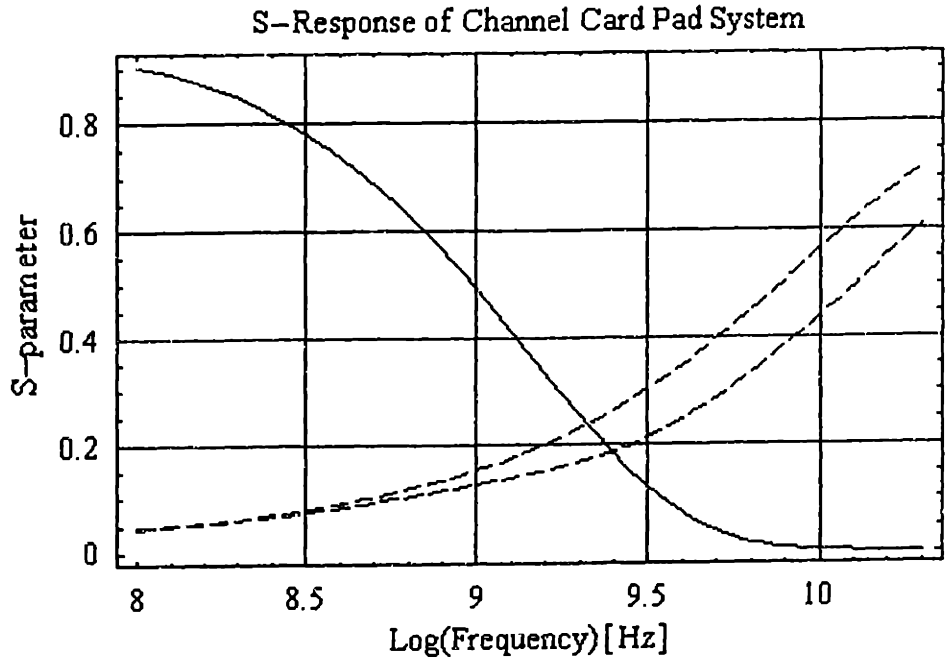


Figure 5.19: Channel Card Subsystem Performance

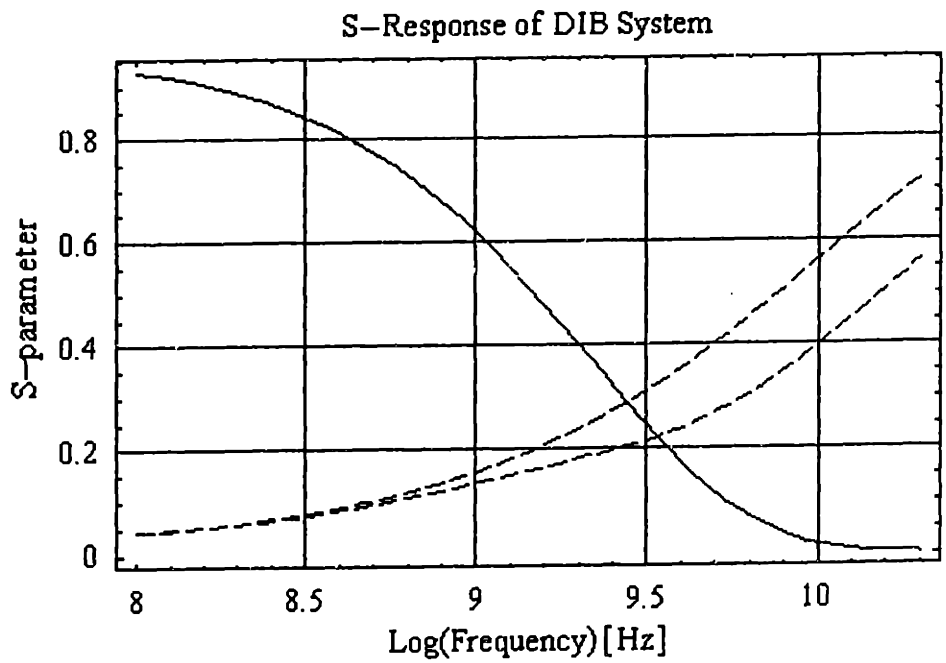


Figure 5.20: DIB Subsystem Performance

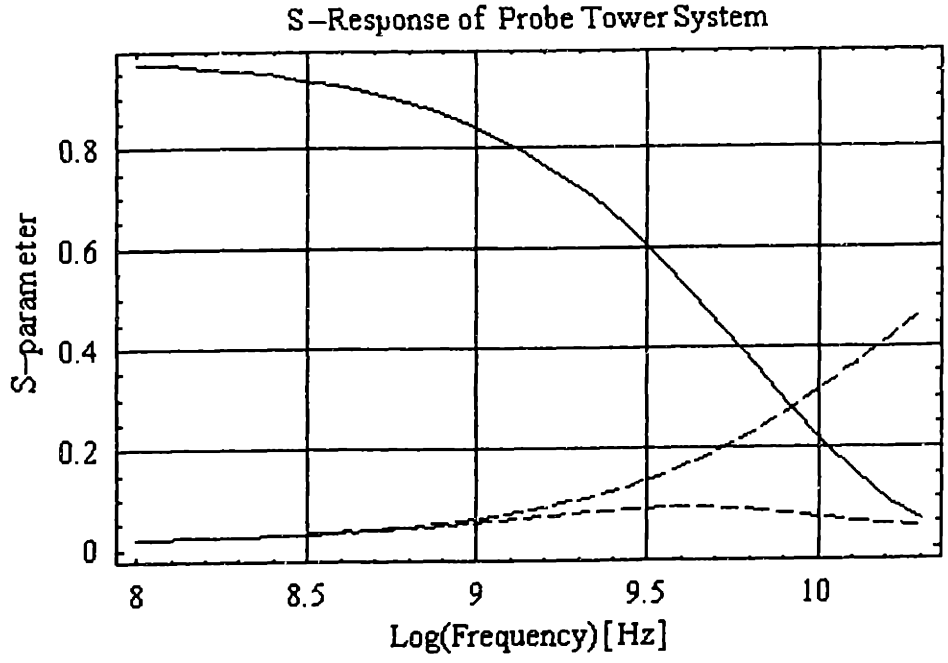


Figure 5.21: Pogo Tower Subsystem Performance

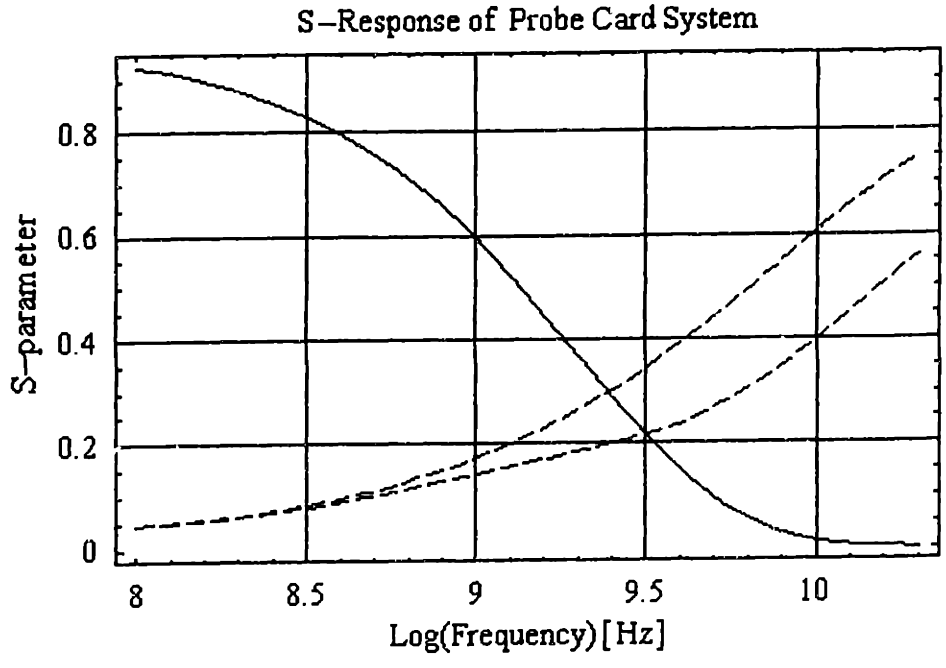


Figure 5.22: Probe Card Subsystem Performance

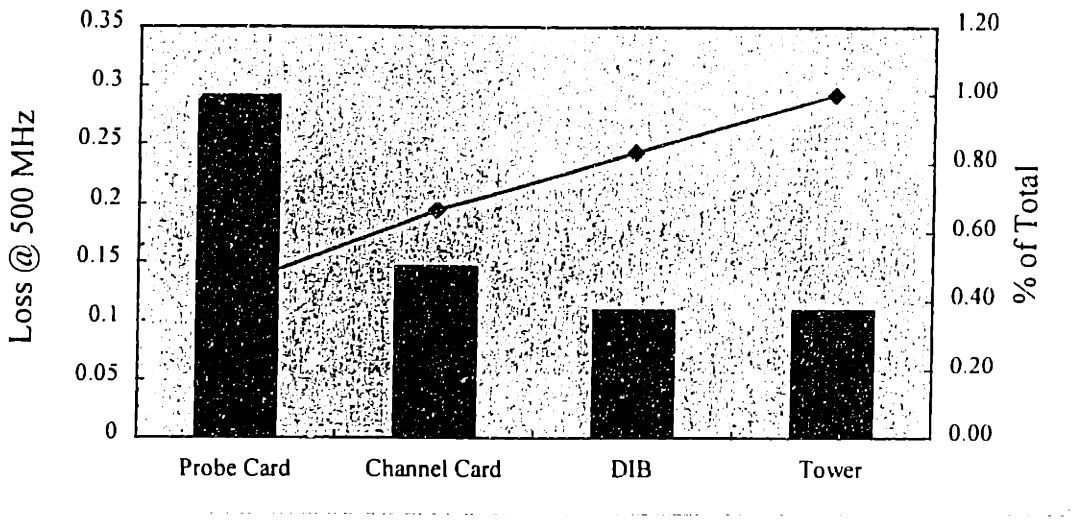


Figure 5.23: System Pareto

The analysis above shows that for the baseline system, the probe card contributes greatest component of bandwidth loss, followed by the channel card, whose losses are dominated by PCB losses. This result agrees with practical experience, and often shortened probe needles combined with bypass capacitors are used to reduce the inductive effects of these elements. The same model can be used to estimate the value of bypass capacitor required by adding the component and explicitly solving for bandwidth = f(bypass capacitance). However, a simple SPICE simulation would likely be easier and more accurate

5.6 Modeling Impedance Mismatches

One significant limitation in the modeling above model is the inability to accurately capture output impedance loading due to a mismatch downstream from a component. For example, the effects of an inductive probe needle do not fully propagate through to all of the element boundaries upstream of the device. One method of approximating the effects of this downstream impedance mismatch is to impose a smaller impedance error at the boundary of each element, causing a small reflection. Because the phase information is not tracked, these reflections cannot be used to reconstruct the wave-form error at the source due to reflections. However, an estimate of the error at the comparator due to reflected energy (S11 for the system) due to this additional impedance mismatch can be approximated using conservation

of energy. This is done by comparing the return loss (S11) of the mismatched system to that of the nominal (50 ohm) model. Both systems will have the identical insertion loss (S12) and the difference in return loss is the additional reflected energy due to the impedance mismatch.

The above baseline model can be modified to perform such an analysis by simply changing the nominal impedance from 50 ohms to 55 ohms, a 10% variation, while leaving the transmission component values and geometry unchanged. This results in a 50 ohm line with a 10 ohm mismatch at each boundary. Figure 5.24 show the output of this analysis, which confirm that the return loss of the system has increased from 0.17 to 0.22 at 300 Mhz. The result of this additional reflected energy is a decrease in system bandwidth from 350 Mhz to 250 Mhz.

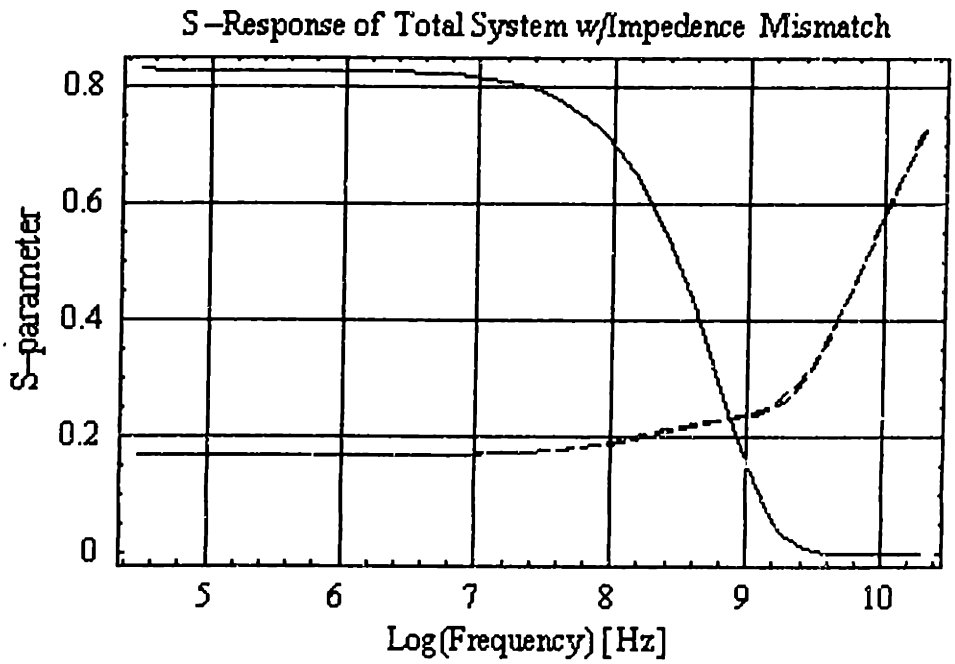


Figure 5.24: Baseline Model with Impedance Mismatch

5.7 Sensitivity Analysis

The simulations presented above show that the model is capable of capturing the fundamental frequency dependent behavior of signal transmission lines. This behavior is

important to the signal transmission designer, unfortunately, the accuracy of the results limits the utility of the modeling tool for this type of analysis. A more appropriate application of these tools is to use them to gain insight into the non-intuitive coupling between mechanical parameters and electrical performance. Sensitivity studies can be performed that predict the effects of design variables that can propagate through an entire system. Because the model is symbolically based, exact relationships between the variables and dependent parameters can be developed.

There are a number of important and poorly understood relationships that can be investigated. Of particular interest are those that are directly related to the mechanical systems roadmaps presented in Chapter 3. This section presents several example sensitivity studies that are related to the predicted mechanical requirements, including a study of the affects of mechanical accuracy, temperature and manufacturing tolerances on signal performance. This analysis brings the device roadmap analysis back from the mechanical domain to that of measuring the DUT.

5.7.1 Sensitivity to Mechanical Motions

There are two basic phenomenon that couple mechanical accuracy with signal delivery performance. The first is the pointing accuracy required to connect the test pin, such as a probe needle, to the signal pin of the DUT. Without sufficient pointing accuracy, testing is impossible and this is the relationship investigated in the accuracy roadmap. The second effect that links mechanical accuracy with transmission performance is the relationship between the geometry and inductance of the interconnects, such as spring pins. At frequencies approaching a gigahertz, even short lengths of unshielded transmission lines can have impedances that are significant enough to cause reflections. If mechanical error motions are large enough, they can cause changes in the exposed lengths of these spring pins and thus cause variations in the impedance. The previous analysis showed that a 10% mismatch in impedance can cause a reduction in bandwith from 350 to 250 MHz. The transmission line models developed above can be used to determine if mechanical error motions can cause impedance changes that can result in significant reflections and thus reduced bandwidth.

Since the effects of a local impedance change on the global system have already been modeled, it is sufficient at this point to model only those components that are directly affected by, or are in close proximity to, the mechanical error motions. For example, the spring-pin to DIB interconnect is an area that commonly experiences mechanical error motion due to docking misalignment. The effects of this misalignment on signal transmission can be approximated by modelling those components that are local to this motion; i.e. the spring pin on the channel card and the pad and trace on the DIB.

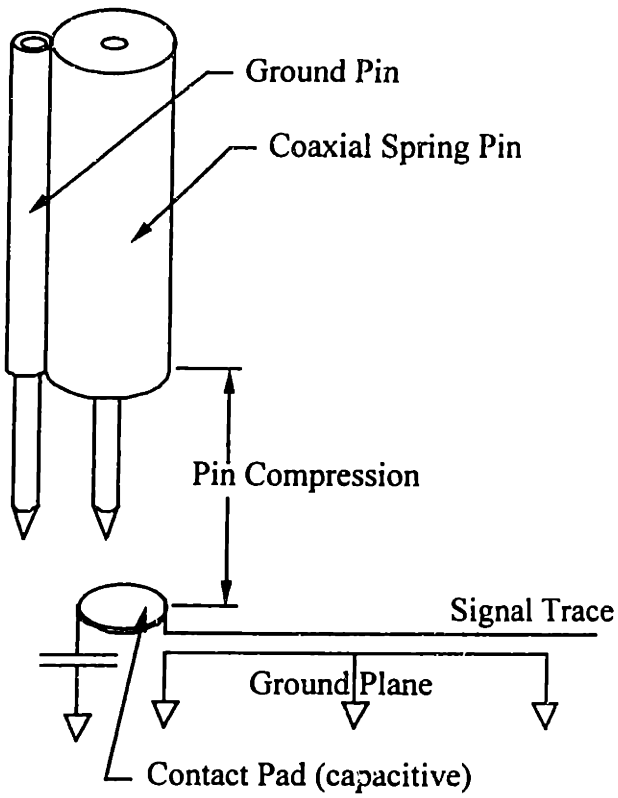


Figure 5.25: Components Affected by Mechanical Accuracy

Figure 5.25 schematically shows the components that are local to the mechanical error motions. Nominal values from the baseline system can be used to assemble a model of these components and the effects of small changes in the length of the exposed spring pin needle can be investigated. Figure 5.26 shows the S-parameter response of the nominal system, with 0.060" of exposed pin, as well as the response of the same system with 0.040" of exposed pin.

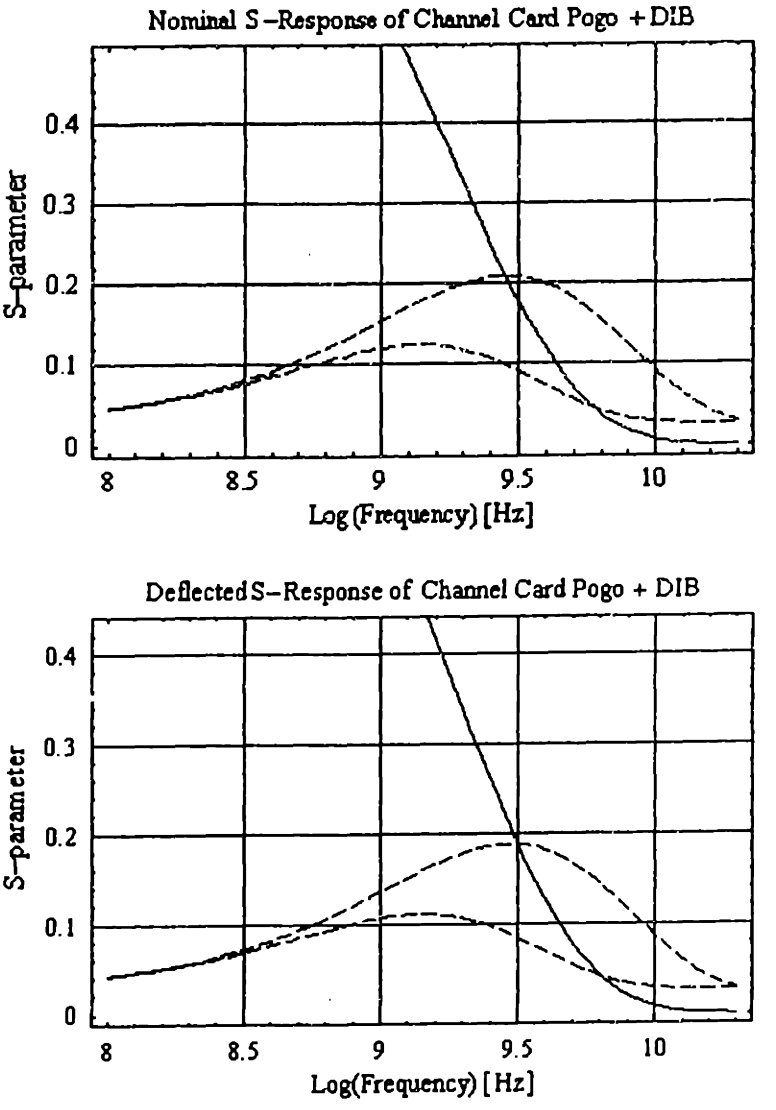


Figure 5.26: Affect of Mechanical Motion on S-Parameters

Although these response curves look similar, there are some slight differences in the bandwidth and return losses of these two simulations. The results of these two models are summarized in Table 5.1 below.

	Nominal System	Displaced System
Insertion Loss (S12) @ 300 MHz	0.82	0.8
Insertion Loss (S12) @1 GHz	0.57	0.55
Return Loss (S11) @ 300 MHz	0.07	0.08
Return Loss (S11) @ 1 GHz	0.13	0.15

Table 5.1: Effect of Mechanical Error Motions

As the previous model predicted, impedance mismatches as small as 10% can propagate through a system and cause significant reductions in overall bandwidth. Unmodeled changes in S-parameters due to manufacturing tolerances and environmental changes are not uncommon and therefore are generally calibrated out in high performance systems. However, errors due to mechanical displacement may occur between calibrations and therefore a 2% shift in S-parameter may have a significant effect on the overall measurement accuracy.

5.7.2 Comparison to Measured Data

In addition to pure simulations, these parameters can be easily measured using a network analyzer and the measured these results compared to those of the model. Figure 5.27 shows a picture of a fixture built to control and measure the compression of a spring pin against a signal pad on a PCB. Custom, controlled impedance PC boards were built and calibrated and the measurements were made using an HP 8753A 300 kHz-3GHz Network Analyzer.

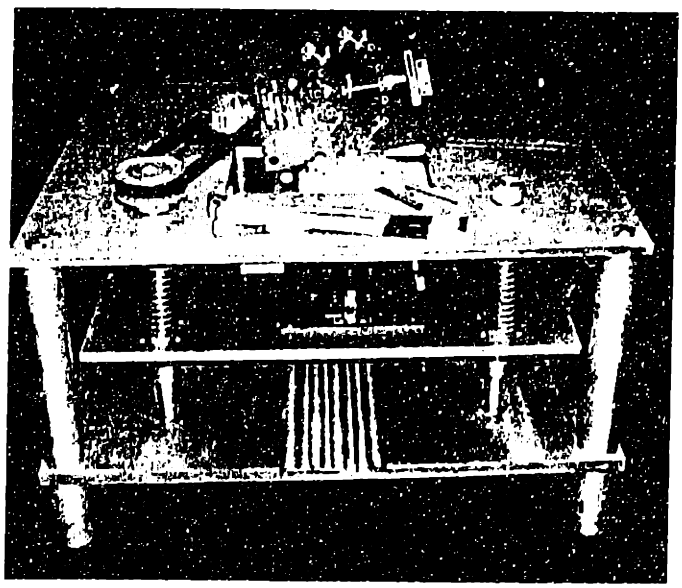


Figure 5.27: Spring Pin Compression Measurement Fixture

There are several ways to compare the data generated from these tests to that of the model.

Figure 5.28 displays a frequency response plot of insertion loss for several different pin compressions. This graph shows that the modeled system captures the overall rolloff trends

of the transmission line, but cannot capture the more local, frequency dependent phenomenon.

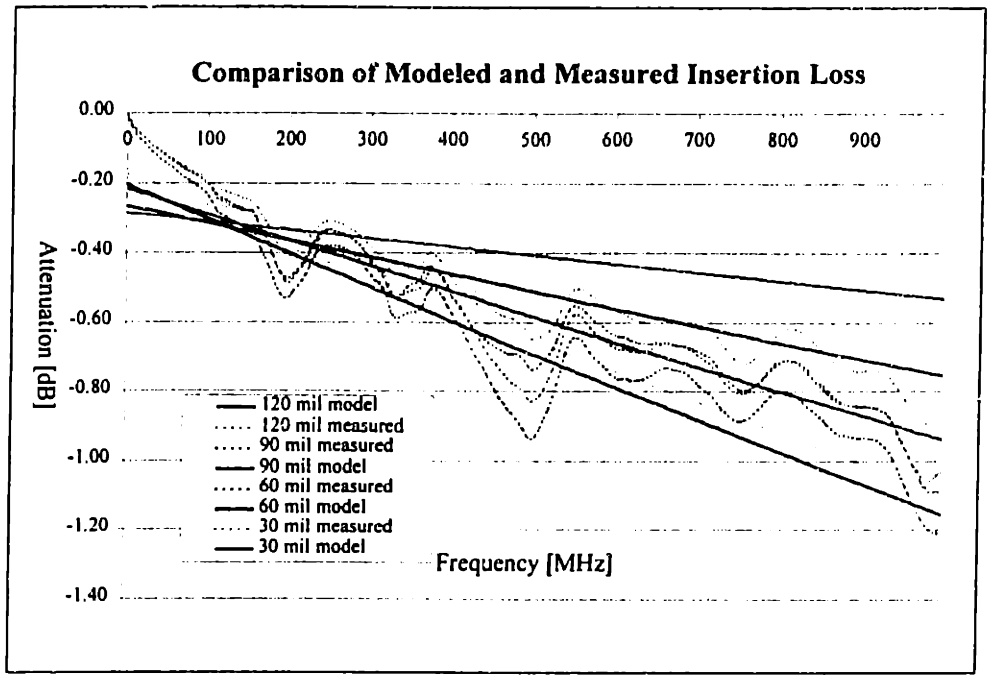


Figure 5.28: Comparison of Modeled and Measured Insertion Loss

A second view of the same data is shown in

Figure 5.29. This displays the insertion loss as a function of pin copression for a number of frequencies. This view of the data more clearly shows the effect of mechanical error motions on signal transmission bandwidth. At low frequencies, these curves are relatively flat and there is little coupling. However, as frequencies increase, the coupling between these two variables increases. This graph also shows that the modeled behaviour roughly agrees with that of the measured data.

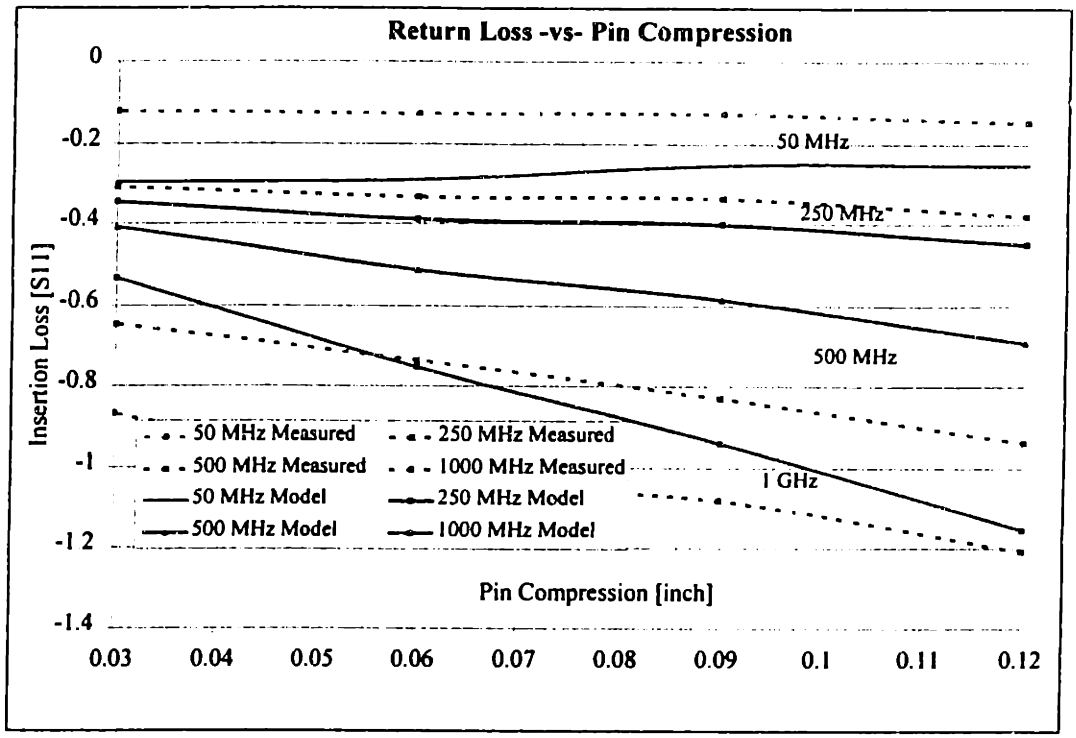


Figure 5.29: Return Loss –vs- Pin Compression

5.7.3 Symbolic Sensitivity Analysis

There is another mode of sensitivity analysis that can be performed using the transmission line models developed above. Because the models are developed using *Mathematica*, which is a symbolic mathematics engine, the relationships between all variables is preserved. Therefore, the relationship between any two variables can be solved for explicitly. For example, using nominal values, the scattering matrix as a function of spring pin length can be solved:

$$\begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} = \begin{bmatrix} \dfrac{.0935(.005+len)(.073+len)}{(.073+len).078+len)} & \dfrac{(.009+.065len)}{(.017+.226len+len^2)} \\ \dfrac{.005+.065len}{(.005+.152len+len^2)} & \dfrac{.842(.005+len)(.073+len)}{(.073+len)(.078+len)} \end{bmatrix}$$

The sensitivity of the scattering parameters can then be found by taking the partial derivative of [S] with respect to the pin length (len):

$$\frac{\partial[S]}{\partial len} = \begin{bmatrix} \dfrac{.068}{(.079+len)^2} & \approx 0 \\ -\dfrac{.065(.074+len)^2}{(.006+.152len+len^2)} & \dfrac{.062}{(.079+len)} \end{bmatrix}$$

The second equation clearly shows that changing pin length primarily affect the reflected elements of the scattering parameters. The coefficients in the above equations are also functions of all of the other gemoetric and physical parameters that define the entire transmission line. Sensitivity studies using these equations can include multiple variables by looking at other partial derivatives and these equations can also be inverted to allow optimization of a particular parameter with respect to any variable.

5.7.4 Other Sensitivity Studies

In addition to investigating the affect of a specific element, such as the compression of a spring pin, the transmission line models can be used to perform more global sensitivity analyses. For example, there are several parameters such as temperature and PCB manufacturing tolerances, whose influence is distributed throughout the transmission system. The dielectric constant of PCB materials such as Kapton can vary up to 10% over the common operating range of test systems. PCB manufacturing processes are able to control feature widths and layer thicknesses to no better than 0.001", but for high density boards with traces less than 0.010" wide, this represents a 10% variation. The easiest method of investigating the sensitivity of transmission bandwidth to these variables is to adjust the `tol` variable in those components whose performance is expected to vary due to thermal or manufacturing variations.

A more exact analyses can be performed by explicitly defining the relationships between temperature and dielectric constant. Also, signal integrity is more sensitive to some PCB geometric variations than others. The width, dielectric thickness, overall length or copper weight can be varied individually for any or all PCB components within the signal path.

5.8 Analysis of Crosstalk and Signal Density

One other significant phenomenon that should be considered is signal distortion due to crosstalk between signal traces. This phenomenon is not explicitly addressed in the model,

but a rough estimate of its effect can be made by examining the coupling between two adjacent, parallel signal lines. This coupling is primarily a problem on high density microstrip printed circuited boards, where the distances between adjacent traces is on the same scale as the distance between an trace and the ground plane. The magnitude of this crosstalk interference is a function of the microstrip geometry, frequency and the length of adjacent parallel line. Commercially available magnetic field solver and PCB simulation can be used to evaluate specific cases, but no general model exists to investigate the relationships between board density and crosstalk or signal integrity in general.

Crosstalk is essentially due to mutual inductance, therefore, its effect can be approximated by treating the coupling between the signal traces as a mutual inductance. Figure 5.30 shows a simple schematic of how two parallel traces can be modeled as a transformer. In most cases, the microstrip geometry is such that the effective turns ratio of an equivalent transformer is 1:1, therefore, the mutual inductance is typically low, and primarily a function of the length (L) which the trace run parallel to each other. A mutual inductance L between these two traces will cause a noise voltage in the second trace proportional to the rate of change of current in the first trace. A schematic of this phenomenon is shown in Figure 5.30 below.

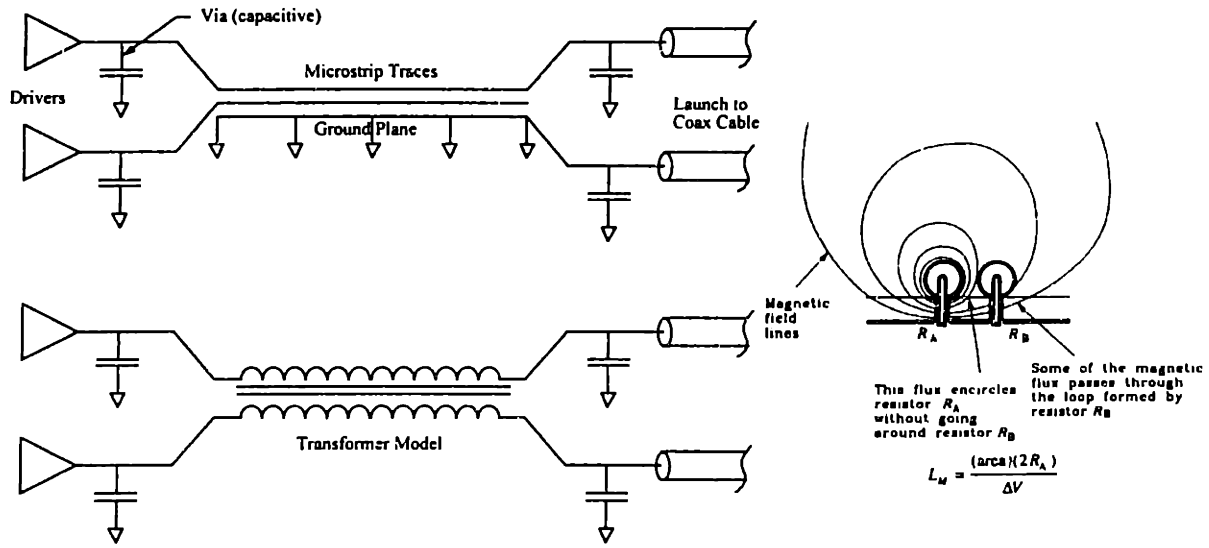


Figure 5.30: Analysis of Crosstalk Between PCB Traces[35]

[35] Johnson

Empirical analysis of crosstalk is also available for the following two cases shown in Figure 5.31.

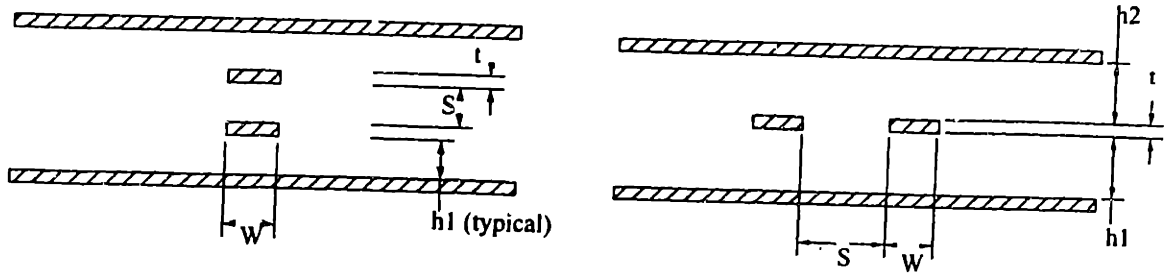


Figure 5.31: Crosstalk Model Geometry[36]

Crosstalk can be calculated as follows:

$$V_0 = V_{in}\left(\frac{C_m}{C_m + 2C_0}\right) \qquad \text{and} \qquad Crosstalk = \frac{V_o}{V_{in}}$$

where:

$$C_m = \varepsilon\left(\frac{W}{S} + \frac{2}{\pi}\ln\left[cIth\left(\frac{\pi s}{4h_1}\right)\right]\right) \qquad C_0 = \varepsilon\left(\frac{W}{h_1} + \frac{B}{\pi}\ln 2\right)$$

While either of the above descriptions can be used to capture the relative effect of crosstalk between adjacent transmission lines, the mutual inductance of a typical microstrip trace is sufficiently small such that the signal transmission model in it's current form is unable to capture the crosstalk phenomenon. The additional refinement required to do this analysis is discussed in the following section on future model development. This additional work is worthwhile, however, because it will capture the relationship between signal degradation due to crosstalk and physical signal density. The output of this analysis will be an upper limit on the signal density of printed circuit transmission lines. This is of significant interest to the ATE systems architect because it may limit the ultimate signal density of PCB based test fixtures, however, it is likely that device application will approach this limit faster than test applications.

[36]Canright

5.9 Future Model Development

The models and simulations presented above represent a first step in the development of a more formal and intuitive understanding of the relationships between the mechanical design and electrical performance of signal transmission systems. While these models address the basic analysis needs in the design of these systems, there is significant additional functionality that should be added to make them more powerful. There are also several other analysis tasks that can be undertaken that would lend significant insight into the design of high performance signal transmission systems. In particular, the following model improvements are suggested:

- Phase information should be retained so that both time and frequency domain analysis can be performed. This will require a slightly different approach to the structure of the model and fewer independent variables to keep the computation times reasonable. This will allow the tracking of reflections at every component boundary and the exact reconstruction of the output waveform.
- The symbolic analysis tools are currently difficult to use. Macros and pre-programmed analysis should be added to unlock the power of this aspect of the model.
- The crosstalk modeling methods presented above should be implemented in code to allow investigation of the relationships between signal density and integrity.
- The constitutive models of the individual components should be refined to include additional second order effects such as crosstalk interference as well as leakage and emissive losses.

5.10 Summary of Chapter 5

This chapter discussed the application of the signal transmission modeling tools developed in Chapter 4. The goal of this modeling effort was to develop an intuitive understanding of signal transmission systems and develop a method of discerning where future ATE problems may lie within the context of the device roadmap presented in Chapters 2 and 3.

The derivation and results of several simulations were presented, beginning with a baseline model that represented a state-of-the-art high-speed digital test fixture transmission path. This baseline model was modified to demonstrate how a local impedance discontinuity can affect the overall bandwidth of a transmission line. Sensitivity studies and supporting empirical data were presented to show how mechanical disturbances couple with transmission performance. The chapter closes with a discussion of additional analyses that should be performed as well as future work that is required to make the model more useful.

6 Case Study of the *Kinematic Docking System*

6.1 Introduction

The first five chapters of this thesis discussed some important trends in the mechanics of ATE and developed an analysis method for analyzing ATE signal delivery systems. These chapters described a growing class of mechanical test problems caused by increases in device pin count, density and speed. One of these problem areas of critical importance is the design of the interfaces and test fixtures that mechanically interconnect the tester with the wafer prober or device handling equipment. The analysis developed in Chapter 4 showed that the mechanical repeatability of this interface has a direct correlation to the electrical performance of the signal delivery system and the analysis in Chapter 5 quantified this relationship. This chapter presents the design and implementation of one solution that addresses the above issues when interfacing test heads with device handling equipment.

The chapter begins with a background discussion of test cell interfaces, including a description of existing designs and their problems. The chapter continues with the development of the functional requirements of a solution, as well as a description of common Precision Machine Design (PMD) tools and techniques such as determinism and the kinematic coupling. The remainder of the chapter describes the development and implementation of the *Kinematic Docking System* (KDS). This includes a description of the basic design concept and details of specific innovations. Overall, this chapter demonstrates how basic PMD concepts were used to address some of the high level trends described in the SIA roadmaps.

6.2 Background

This section defines the function and requirements of mechanical interfaces and narrows the scope of the investigation to that of testhead docking systems. This is followed by an overview of the state-of-the-art in test cell interfaces at the time of the development of the KDS as well as an analysis of their benefits and limitations. The section closes with a discussion of the factors and trends that motivate the redesign current systems.

The analysis and development described below applies to both wafer probing and package device handling test cells. For the sake of simplicity, the term device handler will be used to refer to either of these systems.

6.2.1 Definition of Interface

A test cell mechanical interface is responsible for aligning and holding the testhead and signal interface fixturing with respect to the device handler. These tasks can be partitioned into two major sub-functions: testhead dock and DIB dock. The testhead docking system is primarily concerned with accurately and repeatibly positioning the testhead metrology frame to that of the device handler. DIB docking is concerned with the positioning of the test fixturing to the testhead along with the multitude of electrical interconnections between them. The final link in the metrology loop, alignment between the device handler and the test socket, is completed through a closed loop positioning system in the case of a wafer prober, or is accomplished through open loop, mechanical fixturing in the case of most package handlers.

Interfaces generally perform two separate functions: alignment and compression. The need for alignment is dominated by the signal interconnections between the testhead and DIB. These are typically spring pin to pad connections that require approximately five to ten mil pointing accuracy. The DIB is aligned and fixtured to the device handler and the interface guides the testhead pins to the pads on the DIB. Some higher performance applications require that the DIB be aligned to the head prior to docking and in this case, the interface is responsible for positioning the probe needles or test socket with respect to the device, usually to sub-mil accuracy.

Gross alignment of the testhead to the handling device is also required to properly engage the second interface function; compression. Virtually all interfaces rely on spring pins to complete the signal path between the testhead and DIB. Individually, each spring pin requires only a few ounces of force to compress. However, a complete interface may have several thousand of these pins resulting in hundreds of pounds of pin compression force.

Every interface design must be able to produce both the accuracy and pin compression force described above. In addition to this, the interface must be robust enough to reject any

external disturbances, such as static loads from an unbalanced manipulator or twisted cables, and dynamic loads from vibrating device handling equipment or environmental noise.

6.2.2 Design of Existing Interfaces

Test cell interface design have evolved from simple PCB edge connectors to complex mechanical assemblies with pneumatic or motorized actuation. Almost all of these systems are semi-custom and are redesigned with each new generation of test system and often, new designs are developed for specific applications. Few standards exist and interface components are often modified to meet the specific demands of a particular application. Despite the broad variety of interface designs and components, they can generally be separated into three basic categories based on how they develop the interfacing force; breech lock, pull-down block and vacuum.

Breech lock designs use a ring with cylindrical cams on one half of the interface that engage radially fixtured followers on the mating half of the interface. As the ring is rotated, the cam surfaces force the followers to move axially, compressing the signal pins. These types of interfaces typically encircle the DIB and thus are approximately 12-16" in diameter. An example of this type of interface is shown in Figure 6.1. Breech lock designs are generally simple to implement and offer tremendous mechanical advantage. However, they can be quite expensive due to the size and complexity of the rotating ring and require a large-diameter bearing to support the compression loads.

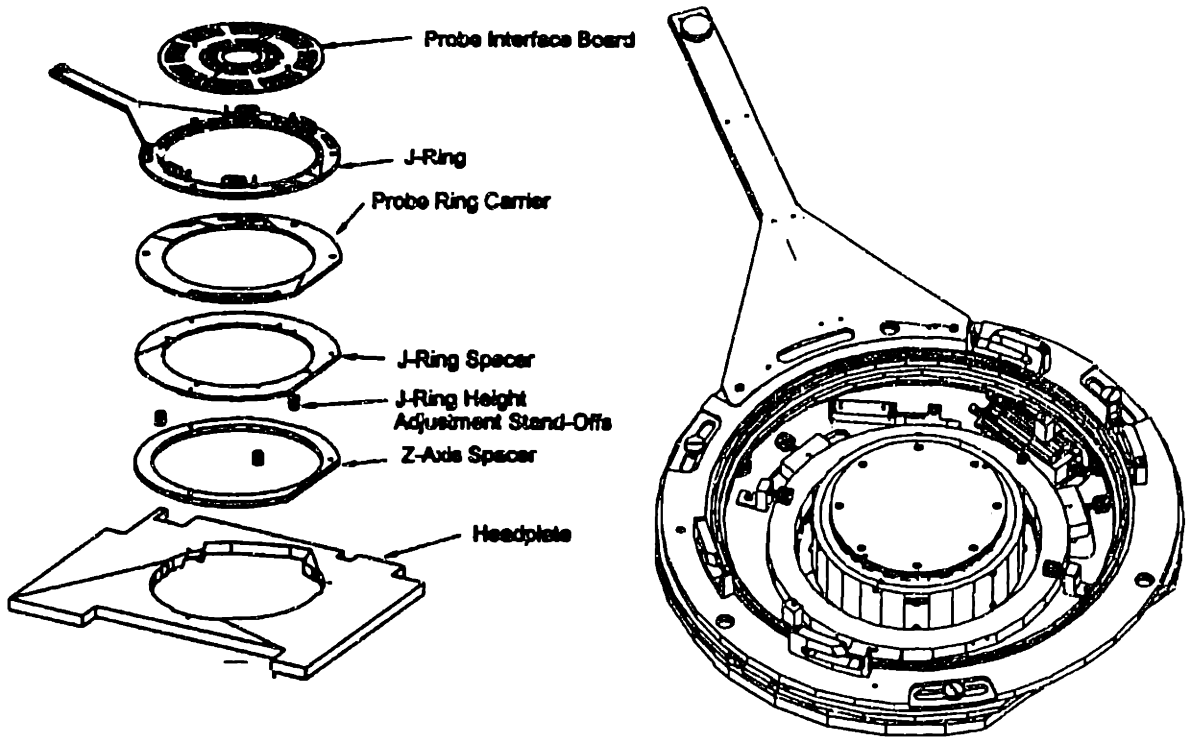


Figure 6.1: Breech-Lock Interface Design

Pull-down block designs use multiple cam/follower assemblies located beyond the diameter of the DIB, between the testhead and the top plate of the device handler. This design eliminates the need for large machined parts and bearings. These assemblies are typically actuated simultaneously through a cable or linkage and pull the two halves of the interface together, compressing the signal pins. An example of this type of interface is shown in Figure 6.2. Because pull-down blocks must be located further from the center of test and do not have a unifying structure, they rely on the stiffness of the testhead and device handler mounting locations. Often, convenient, rigid mounting locations for these blocks do not exist and as a result, performance is compromised.

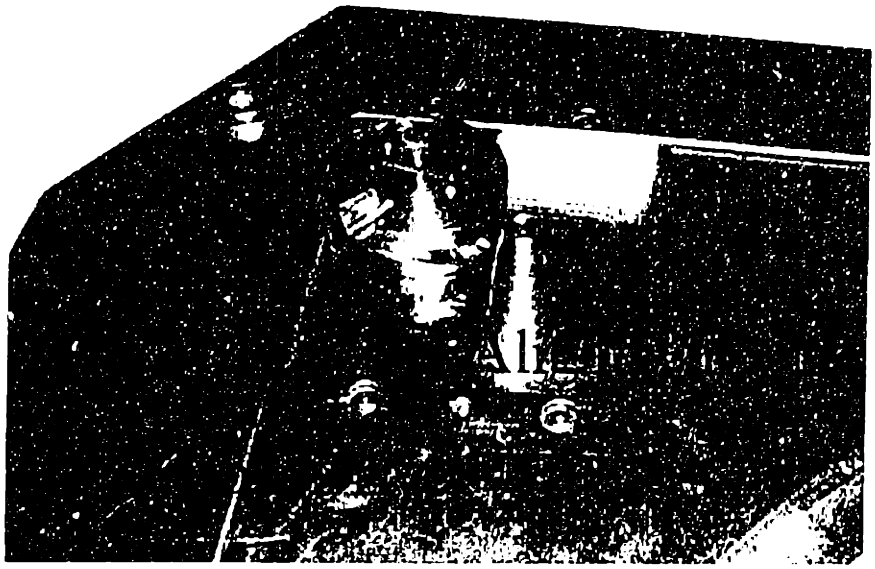


Figure 6.2: Pull-down block interface design

Vacuum is often used in the testhead to DIB interface due to the large DIB area available and in systems with small testheads, it can also be used to hold the testhead to the device handler. However, the sealing, stiffness and alignment required to interconnect a large testhead makes this type of interface impractical for many high pincount applications. Interfaces have been designed that use a vacuum assisted with a mechanical latch, eliminating the need for cammed surfaces and timed actuation. Drawbacks of vacuum systems include problems of seal wear, plumbing and the additional requirement of a vacuum source.

All of the above interface types use similar pin and bushing designs to align the mating components of the interface. In the case of the breech lock, these alignment features are external to the locking mechanism. The vacuum and pull-down block designs often incorporate the pin-bushing alignment features into the design of the block.

6.3 Problems with Existing Interface Designs

Chapter 3 described the evolution of test interfaces and fixtures and how they have evolved from simple sockets, to cables to a small collection of electronics and eventually to testheads. The growth in demand of CMOS devices has increased the need for pin electronics near the DUT and thus has accelerated this trend and as a result, testhead size, weight and volume have increased rapidly. The diversity of test applications has also grown

over the years, requiring test systems to be more configurable and customizible. These needs have been addressed by adding multiple layers of interface between the pin electronics and DUT, thereby increasing the number of interface layers.

These two trends; increasing TH size and increasing interface layers, combined with the ever-increasing need for mechanical accuracy, have combined to create the following problems:

- Large tolerance stack-up
- Multiple structural loops
- Interface instability

Not every interface suffers from all of these symptoms, but all interfaces fail to address at least one of them and compromise on the others. Each of these problems is discussed in detail below.

6.3.1 Interface Tolerance Stack-up

Typical production interfaces can contain as many as 5 PCB interface layers between the pin electronics and the DUT. Each of these PCB layers requires alignment and compression mechanisms, and often, these boards are part of the structural loop and alignment is done using a PCB surface. Figure 6.3 schematically shows the layers of a typical interface.

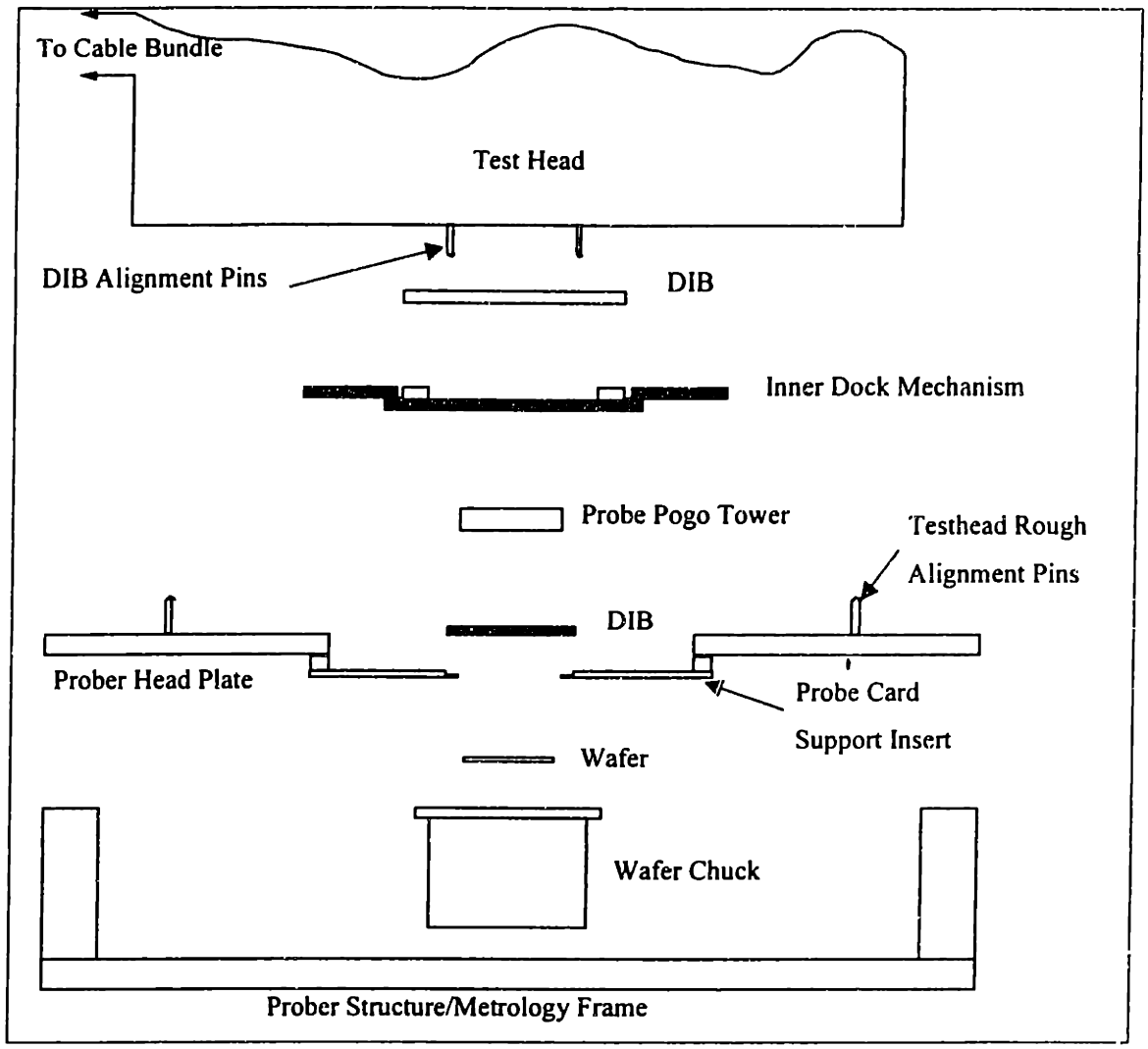


Figure 6.3: Layers of a Typical Interface Structural Loop

The primary problem with this arrangement is that each additional layer of interface adds a layer to the tolerance stack-up and the dimensional accuracy and stability of the PCB material, typically FR4 or G10, is very poor. Board thicknesses can easily vary up to 5 mils across a board and between batches. Locating holes drilled in this material can quickly wear or fail due to the high contact stresses imposed by alignment pins. As a result, achievable tolerances for interfaces that include PCBs in the accuracy loop are generally no better than 5-10 mils/board layer.

In addition to imprecision of manufacture, PCBs also lack the structural stiffness to accurately support the loads imposed by the signal pins. The modulus of G10 is roughly two orders of magnitude less than that of steel and as a result, large errors due to distortion are common. To counter these error motions, DIBs are often mounted on a stiffener or

compressed between two reference surfaces. Both of these solutions require precision mating components and to the depth to the interface, potentially resulting in longer signal path lengths.

6.3.2 Multiple Structural Loops

The problem of poor tolerances and multiple layers of PCBs often makes it impossible to have a single structural loop between the DUT and probe card or socket. For example, the probe card is often rigidly fixtured to the head plate of a prober. This insures that, regardless of the tolerance stack-up in the interface, the probe needles are properly aligned to the DUT. However, the signals must flow from the probe card, through a pogo tower and DIB to the testhead, and therefore additional alignment features are required to insure signal contact. The testhead is also be directly connected to prober structure, resulting in an overconstraint, as shown in Figure 6.4.

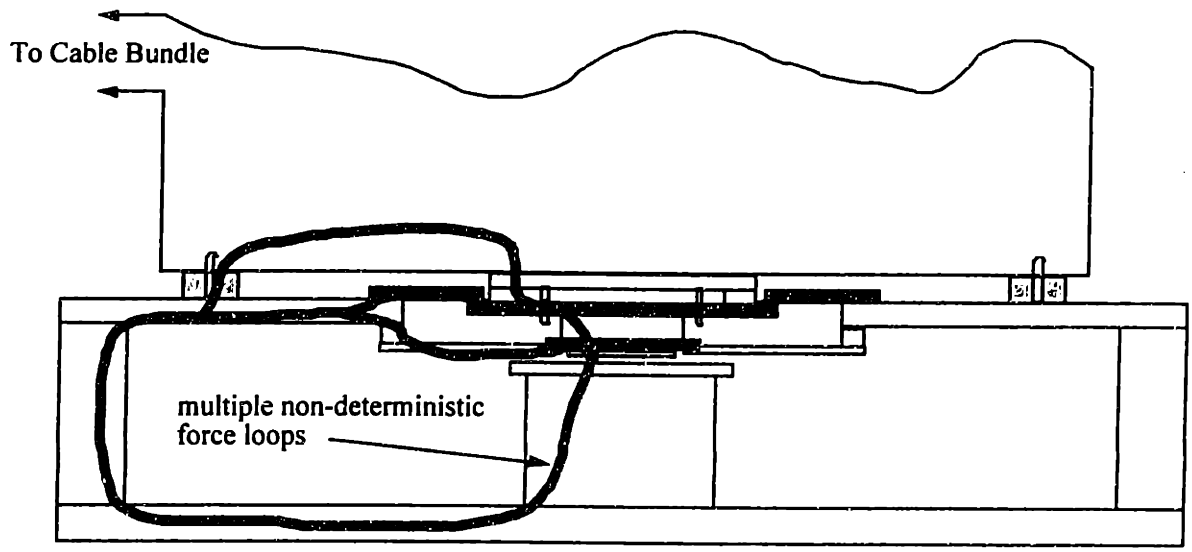


Figure 6.4: Interface Structural Loops

This redundancy causes binding, wear and un-modeled deflections of the interface components, often resulting in premature failure. A larger problem, however, is that the overconstraint reduces predictability of the system. Thus, it is often difficult to make a clear correlation between an interface failure and a remedy. This results in long system downtimes for troubleshooting, experiments and repair.

Another example of interface overconstraint can be found in the design of many pull-down block-type interfaces. Square testheads usually use four of the block assemblies, one

in each corner and often these are designed as an afterthought or retrofit. As a result, the blocks are usually mounted using existing tapped holes and are often mounted to painted or as-cast surfaces and out-of-plane errors as large as 50 mils have been observed in such systems. Because the blocks are mounted to the rigid metrology structure of the testhead and device handler, the overconstraint can only be rectified through large docking forces and inevitable distortion of one or both of these structures. Deflection of the testhead structure may result in the misalignment of the pin electronics boards and bending of the handler structure may result in positioning errors. Both of these problems are extremely difficult to diagnose because they disappear as soon as the system is undocked and the overconstraint is removed.

6.3.3 Interface Instability

One solution to the problem of fixturing overconstraint is the elimination of the alignment between the testhead and handler structures. The probe card is fixtured to the prober head plate, and each successive component is fixtured to the next. In some ways, many existing interfaces used this method; pin and bushing alignment was used between each component with tolerances loosening away from the probe card. Alignment between the testhead and prober structures is loosely constrained in the plane with a pin and bushing and completely unconstrained in the out of plane directions. (Z, roll and pitch) Although this design eliminates the overconstraint described above, it does so at the expense of stability. Again, the inclusion of PCB material in the structural loop results in poor out-of-plane stiffness. But even if this is eliminated, the thin sections of the interface component between the testhead and device handler limit the stability of any interface that does not tie directly to the structure of both of these components. These thin sections are required to maintain short signal paths. Figure 6.5 shows how a centrally fixtured interface can deflect, or 'oil-can' under external moments such as those imposed by the cable bundle.

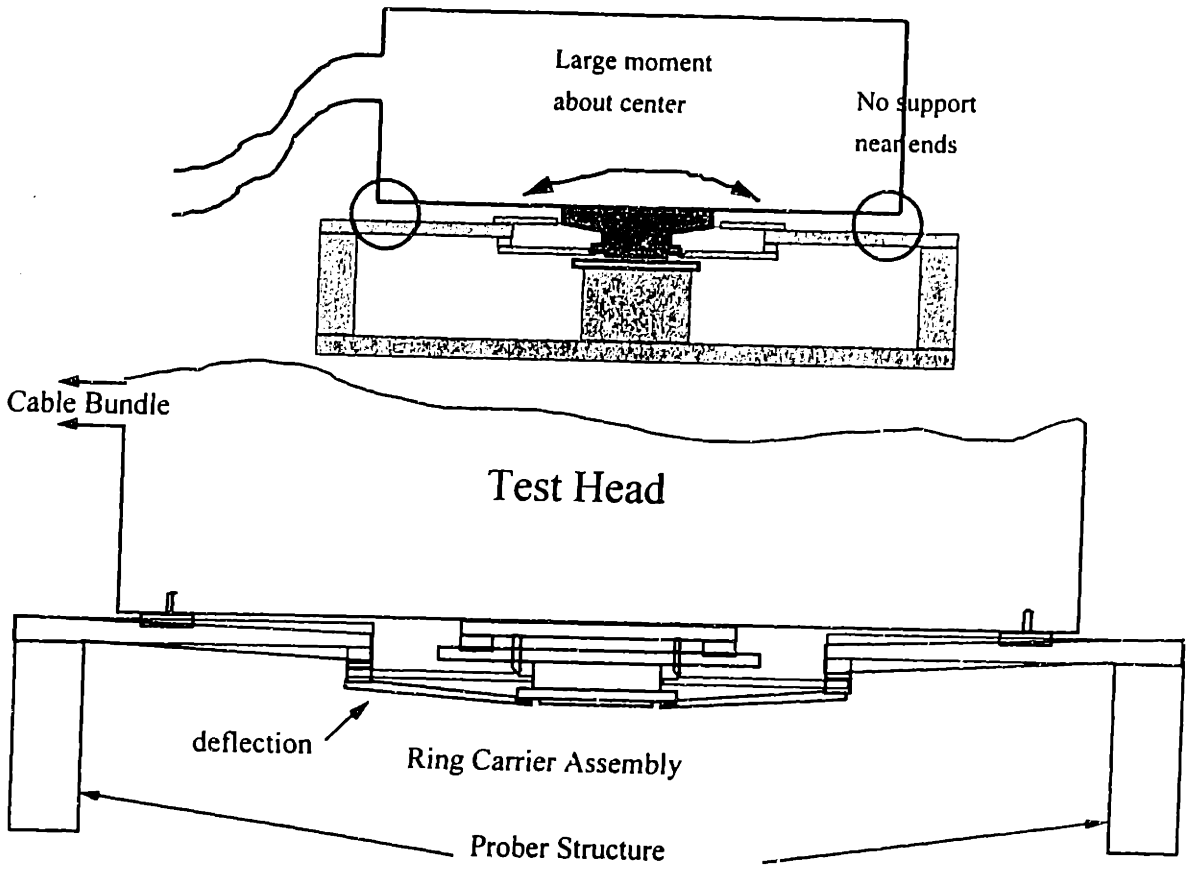


Figure 6.5: Oil-Can Interface Instability

6.4 Motivation for a Solution

If the only result of the above problems were poor design and added cost, it could be argued that efforts to develop a new solution are unjustified. ATE is used in production, and customers want to pay for only the minimum system required to meet their needs and despite their failings, test floor engineers are usually able to get the current crop of interface designs to work.

There are a number of non-technical drivers that help justify the development of a fundamentally new interface system. The most obvious is the accuracy and pin-count needs described by the SIA and ATE road-maps presented in the previous chapters. The existing interfaces described above are barely adequate for the current generation of ATE. This is

witnessed by the increasing number of interface related test problems[37]. These problems will only increase as device pin-counts grow.

A second driver for improved interface development is the reduction of system downtime due to interface problems. The weak link in a test cell is the interface that connects a tester costing in excess of $1M, to a prober that can cost up to $500K. If the interface does not work, both of these machines sit idle. Test cell time can be measured as a function lost production or capitol cost, but either measure results in figures that start at $1000 per hour. Another less obvious result of reduced downtime is the reduction of overcapacity due to the uncertainty. If production interfacing is unreliable, semiconductor manufacturers must build a certain level of overcapacity into their process, resulting in higher capitol cost expenditures.

6.5 Design of the Kinematic Docking System

The redesign of any system presents the opportunity to redefine the performance requirements and broaden the application. However, there are also many requirements that are constrained by the legacy of existing equipment in the field. This section begins by defining the high-level goals and functional requirements of the new interface design. This is followed by an overview of the theory, methods and design concepts used to meet these requirements. The implementation of critical components, control and operation are discussed in the final section.

6.5.1 Functional Requirements

There are three groups of functional requirements that were used to define and constrain the development of the new interface design. The first is that the interface should be a universal solution, compatible with most, if not all existing equipment and not a custom design for a particular application. Second, the interface must meet the anticipated accuracy, repeatability and stiffness performance requirements of all devices that are expected to be tested on the new generation of test systems. The third requirement is that the system present a user interface that is simple, intuitive and does not require a skilled operator. These requirements are described in greater detail below.

[37] Chiu, 1996

6.5.1.1 Universal solution

A universal solution can include many levels of compatibility. For this effort, it was defined as a design that would be compatible with the new generation of testheads, in particular, several that were under development at Teradyne[38]. This required the development and adoption of new standard mounting features that would be common across all of these systems. However, because these new test systems focus on different markets and do not have similar structures, this standard could not constrain their shape or volume. Likewise, the interface solution must be capable of working with the existing and future pool of device handling equipment. This also presented a challenge, because this equipment is designed and manufactured by other independent companies and there are no standard designs. Furthermore, the design should be such that little or no setup or alignment is needed when changing equipment. For example, a common test cell layout may include a single tester with both a wafer prober and package handler. The interface should be designed so that the testhead can be moved between the two pieces of equipment without tearing down or re-fixturing the interface.

6.5.1.2 Performance

There are a number of metrics that describe the performance required of any new interfacing system. In many ways, the test cell is much like a machine tool; if the probe card is thought of as the tool, the DUT is the work-piece and the interface, testhead and prober structures constitute the structural loop. Much like a machine tool, the important parameters are stiffness, accuracy, repeatability and stability. However, because the interface is only a single component in a larger structural loop, it is difficult to assign specific values to these parameters. For this reason, general assumption was made that the interface should contribute no more than ten percent of the budgeted error in the system. Roughly, this translates to an accuracy and repeatability on the order of 25 microns, and a stiffness on the order of 10 N/micron.

[38] Teradyne is a leading supplier of ATE and is located in Boston, MA

6.5.1.3 User interface

The third functional requirement of the new interface is that it have a simple, intuitive operation. Although test floor operators are trained to properly use the equipment, it's operation should not require excessive skill or understanding. This is especially true of equipment that is used in off-shore production facilities. Specifically, the operation should require no tools or adjustments beyond those required for initial setup. Strength, coordination, and height requirements should not exceed those of a 'five percent Asian female' as defined by standard human factors guidelines. Because the equipment will be used in non-english speaking countries, operation should be obvious without written or spoken instruction. Finally, unlike most existing interfaces, the system should provide positive feedback of the status of the docking system. This feedback will reduce the downtime associated with debugging the total test cell.

6.5.2 Design Concepts

A number of core concepts were used in the design of the new *Kinematic Docking System* to address the requirements described above. These include:

- The use of a kinematic coupling to meet the repeatability and stiffness requirements
- Modular design to address the diversity of applications
- An automatic latching and preload mechanism to simplify the operation
- Automated operation to satisfy access, ease of use and operator feedback requirements

These concepts are described in greater detail below.

6.5.2.1 Kinematic Coupling

The KDS interface uses a standard three-ball, three-groove kinematic coupling that is commonly found in precision instruments and optics. Traditionally, kinematic couplings have been used in applications where high accuracy and repeatability are required, but only small loads are encountered. Typical applications include the mounting of optics within an assembly and the positioning of a stylus on coordinate measuring machines. There are also several documented applications of kinematic couplings used successfully in machine tool

fixturing applications[39]. All of these applications allow the location and support of the coupled object without deformations due to overconstraint.

Kinematic couplings are deterministic: in other words they only make contact at a number of points equal to the number of degrees of freedom they are to restrain. The simple and elegant method of using one contact point to constrain one degree of freedom can result in an extremely useful coupling. Two rigid bodies are located relative to each other very accurately by constraining all six degrees of freedom, thus, kinematic couplings have six contact points. Being deterministic, kinematic couplings do not suffer from the problems caused by overconstraint and the accuracy of a coupling depends only on the material properties, geometry and preload force and hence they are extremely repeatable.

In machine tool applications, fixturing has typically been performed using elastic averaging and massive overconstraint. This method is capable of withstanding large cutting forces and offers high stiffness. However, it is generally not as repeatable between setups and the clamping forces and overconstraint may deform the work-piece, resulting in manufacturing errors once the part is removed from the fixture. Properly designed, kinematic couplings can be built with sufficient stiffness to withstand cutting and other external forces.[40] Additionally, it has been demonstrated experimentally that these couplings can repeat on the order of ¼ μm even under the conditions encountered in machine tools.[41]

Application of the kinematic coupling concept to testhead interfaces requires only 3 basic elements: a ball, a groove and a method of preloading the two. The ball and groove elements are relatively simple to implement and performance can be adjusted by modifying ball and groove geometry, coupling diameter and material properties. The application of a preload force to a coupling not only gives it bi-directional stiffness, but to some degree also increases its repeatability performance. How this preload force is applied can also affect the accuracy of the coupling because a large preload force may create deflections in the supporting structure, in this case, the testhead or prober. One method of circumventing this problem is to apply the preload force directly through the center of the ball, as shown in Figure 6.6. This contains the force loop to just the single ball-groove coupling and eliminates the transfer of preload force to the supporting structure.

[39] Slocum, 1992
[40] Slocum, 1988

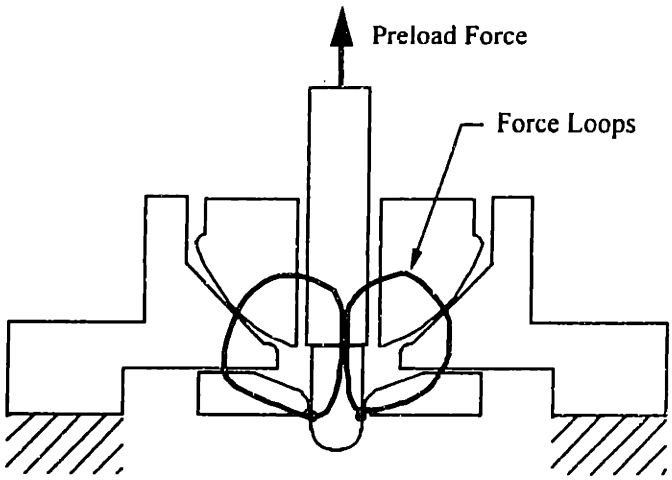


Figure 6.6: Ball-Groove Preload Force Loop

6.5.2.2 Modular Design

The second core design concept used in the *Kinematic Docking System* is that of modularity. The large array of potential testhead and device handler combinations limits the possibility of a single, standardized solution. Instead of a monolithic structure containing the kinematic ball and groove surfaces, two discrete components were designed, a ball module and a groove module. Three ball-modules are mounted to the testhead and three mating groove-modules are mounted to the device handler. To simplify the design, the groove module is completely passive and thus no wiring or communication between the testhead and handler modules is required.

This level of modularity is possible in kinematic fixturing because of the deterministic nature of the coupling; its repeatability is insensitive to small positional errors in the mounting of the modules. If one location is nominally misaligned, the coupling will deterministically find a position that is perfectly constrained. All that is required to properly apply the modules is a relatively stiff structure a reference fixture for initial alignment.

One common drawback of modular designs is the increased cost required to design and manufacture.[42] While this certainly applies to the KDS, it's effect is offset by the relaxed tolerances allowed by the kinematic design and by the economies of scale offered by a more solution that can be applied to multiple test cell designs.

---

[41] Slocum and Donmez, 1988
[42] Cutherell, 1996

6.5.2.3 Ergonomics

The two final design concepts; automatic latch and electronic control, are primarily ease-of-use features and are designed to eliminate the dependence on operator strength and skill. Automatic latch refers to the ability of the ball module to grab and hold, or 'latch', a mating groove-module prior to reaching it's final position. This is a notable feature because two people are usually required to dock a testhead, one to manipulate the testhead and hold it in position while the second operates the interface. The automatic latch feature allows a single operator to maneuver the testhead to the vicinity of it's final docking position, where it will automatically latch. The operator can then release the testhead and operate the interface. Larger test systems may require more operators or a powered manipulator to overcome the significant mass of the testhead and external forces applied by the manipulator and cable bundle.

Each of the three ball-modules contains a gear-motor and an array of sensors. Once these sensors detect that the three ball-modules have latched to their mating grooves, a single button operation instructs the controller to complete the docking process. The controller simultaneously drives the three ball modules into the grooves with a predetermined, repeatable preload force. Once the docking operation is complete, module sensors confirm that the docking operation was successful and the controller indicates this status to the operator. These three automatic control functions: sensing latch, controlled preload, and sensing docking, minimize likelihood of operator errors and therefore decrease the amount of test cell downtime due to interfacing problems.

6.5.3 Design Implementation

This section briefly describes how the above functional requirements and design concepts were implemented. A complete system, shown in Figure 6.7 consists of the following components:

- Ball modules (3)
- Groove modules (3)
- Controller (including cabling and power supply)
- Operator pendant
- Alignment fixture

Figure 6.7: Complete K-dock System

The majority of the mechanical and electrical design is relatively simple, but there are several subsystems and operating features that are of particular interest. These are described in detail below.

6.5.3.1 Ball-Groove Design

The concept of kinematic interfacing is relatively simple and elegant, yet it requires some effort to properly execute. There are a number of aspects of the ball-groove interface design that required close attention to the design details. The first area of interest concerns the method used to latch the ball to the groove. This must be designed in such a way as to not add additional constraints beyond the two ball-groove contact points. This was achieved by decoupling the latching mechanism from the ball and groove. Specifically, the latch receptacle in the groove, a hardened washer, was allowed to float in the plane perpendicular to the latch engagement, as shown in Figure 6.8. This allowed the latching mechanism to engage and preload against the washer without constraining the position of the ball relative to

the groove. The specific design and operation of the latching mechanism is described in Section 6.5.3.2 below.

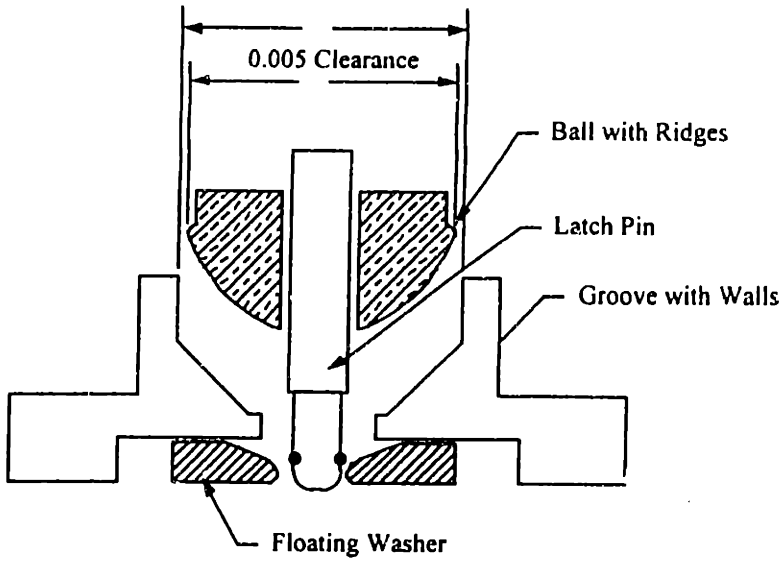


Figure 6.8: Ball and Groove Cross-Section

The selection of the ball and groove geometry and materials is also consequential. A number of different design tools and performance criteria are available to optimize these parameters. Slocum[43] provides a spreadsheet based tool that correlates geometry and material properties to repeatability performance. In this case, the geometry was constrained by system geometry and packaging and a 0.875" SR ball on an approximately 30" diameter bolt circle was chosen. The material was chosen such that the Hertzian contact stress at the ball groove interface did not exceed the maximum compressive yield stress of the material. It was found that 440C stainless steel hardened to 55 RC was sufficient to meet this criteria.

A number of additional features were added to the basic ball and groove geometries to improve the ease of use and performance of the system. Figure 6.8 also shows how vertical walls were added to the grove above the nominal contact point and a rim was added to the outer diameter of the ball. These features limit the lateral motion of the testhead relative to the handling device to +/- 0.010" during the last 0.125" of vertical motion. These additional features prevent the lateral scrubbing and damage of signal pins without overconstraining the coupling when it is fully docked. A molded cover applied over the groove, shown in Figure 6.9, changes the docking target shape from the small hole in the floating washer to a funnel-

[43] Slocum, 1992

shaped orifice of approximately 1.5" diameter. This decreases the amount of skill and precision required by the operator to properly latch the ball module to the groove.

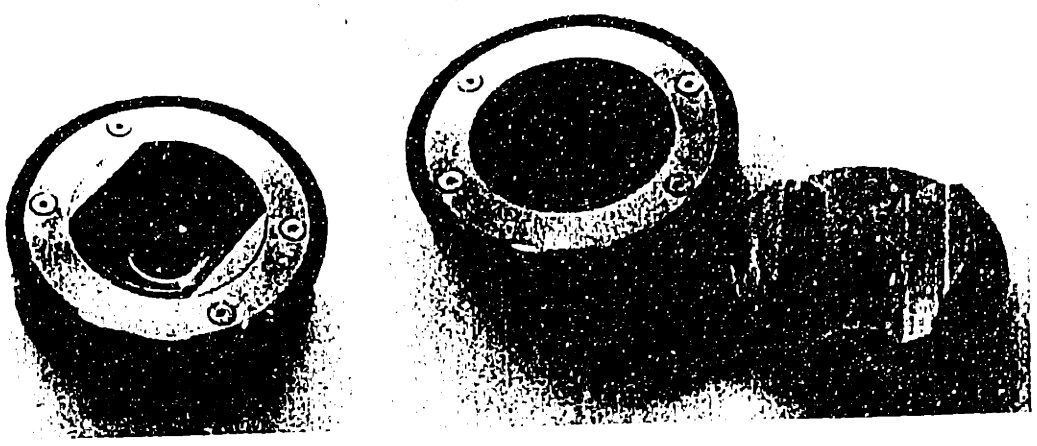


Figure 6.9: Groove with Cover and with Cover Removed

6.5.3.2 Latch-Pin Design

The latch pin subsystem is a simple mechanism that provides the following functions:

- Automatic latch between the ball and groove modules.
- Preloads ball into groove during docking
- Pushes ball out of groove during undocking (pushes testhead away from handler)
- Releases latch between ball and groove when undocking is complete

A number of commercial 'slam-latch' design concepts were investigated to execute this automatic latching feature. All of them require multiple actuators to perform the latch, preload and release operations. Due to size, cost and reliability constraints, multiple actuators are undesirable and instead a novel hybrid ball-lock system was designed that requires a single actuator for all of the above functions. This hybrid ball-lock mechanism is similar to those found in socket wrenches.

There are four main components to this system, the barrel, control pin, balls and receptacle. The position of the control pin is controlled by the preload mechanism, described below, the other three components are passive. A complete cycle of operation is shown in Figure 6.10 and proceeds as follows:

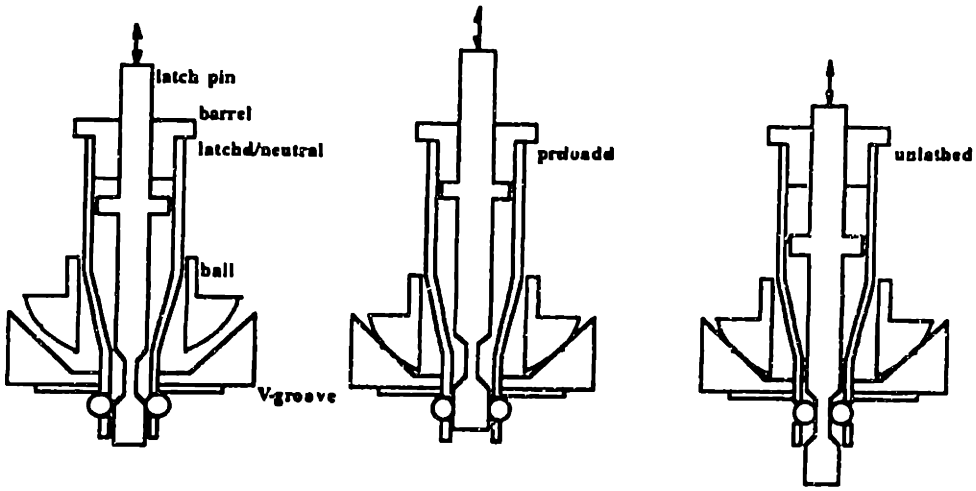


Figure 6.10: Latch and Preload Operation

- Control pin is in the neutral position and system is ready to latch
- Barrel approaches and enters washer. As balls contact washer, barrel moves up relative to control pin.
- As barrel moves up, balls retract into control pin recess, allowing barrel and balls to snap through washer. The system is now latched and cannot be separated. At this point, the barrel geometry prevents the testhead from being driven any closer to the device handler, further preventing scrubbing damage to the signal pins.
- Control pin is retracted, preloading the ball into the groove. Docking is complete.
- Undock button is pressed. Control pin drives to its fully extended position.
- Balls are free to enter recess in control pin and testhead can be withdrawn.
- Ball module is clear of groove, control pin returns to neutral position. (ready to dock)

Components of the latching mechanism were designed to withstand the impact and compressive loads expected during the docking operation. The most severe of these is the brinnelling force of the ball bearing on the circumference of the control pin. Again, the contact stress can be evaluated using the Hertzian contact equations for a sphere on a cylinder. Both of these components were specified as 440C RC 55 to prevent damage.

6.5.3.3 Servo Drive Design

The position of, and forces on the control pin dictate the state of the docking system and the preload force between the ball and groove elements. These positions and forces are controlled by a quasi-open loop servo system that consists of a DC gear-motor and spur gear-train that drives the control pin lead-screw. The motor is the sole actuator in the system, and is driven open-loop between the operating states described in section 6.5.3.2 above.

The servo system, shown in Figure 6.11, consists of the following components:

- Gear-motor (11 W, 66:1 reduction)
- Geartrain (3 spur gears)
- Support bearings (5)
- Clutch release assembly
- Sensors

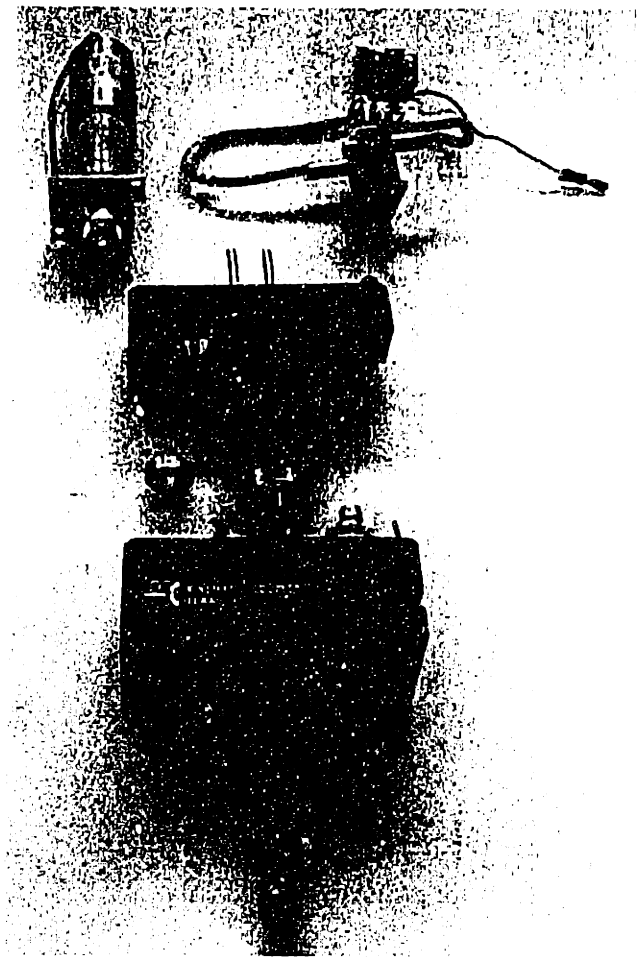


Figure 6.11: Servomechanism Design

The gear-motor was sized so that it met the worst-case operating requirement; driving a testhead away from the docked position. In this case, each docking motor is required to lift approximately one third of the total testhead and cable weight, possibly up to 500 kg total over a maximum operating time of 10 seconds. An 11 watt brushed DC servo motor with 66:1 gear reduction was chosen for this application.

Torque from the gear-motor is transferred to the leadscrew through the small spur gear-train with an additional 2.5:1 reduction. The motor spur drive gear is end-supported with a

radial deep-groove ball bearing to counteract the reaction loads. The main gear-nut is supported by two, fully loaded, angular-contact ball bearings in a face to face configuration. The material for this gear, Nitronic-60, is a high galling resistance stainless steel, and was selected to prevent galling damage at the screw-nut interface. The assembly is lubricated with Krytox, a flourinated lubricant that is well suited for this application due to its extremely low outgassing rate. This allows for use in a cleanroom and also increases the lubricant life, reducing the need for routine maintenance.

An additional spur gear assembly added to the gear-train is used as a manual release system in the event of a malfunction or power failure. This assembly consists of a central shaft supported by deep-groove radial contact bearings. About this shaft rotates a sprague clutch, friction clutch and gear, all arranged serially. During normal operation, the gear spins freely on either the bearings or the sprague clutch. In the event that the system needs to be manually undocked, wrench flats on the shaft can be used to back-drive the gear-train. The sprague clutch allows this to be used only in the release direction and the slip clutch is designed to prevent damage from back-driving into a hard-stop. An additional relay was added to open the motor drive loop when the system was unpowered, to eliminate the back-EMF torque when manually driving the gear train.

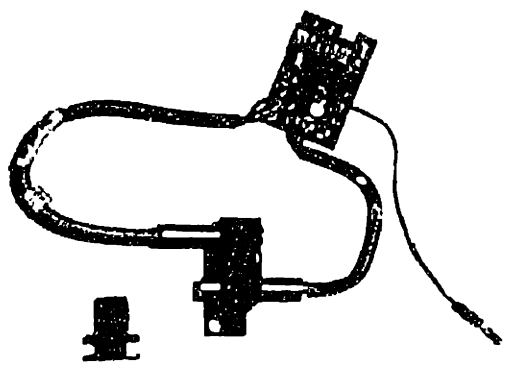


Figure 6.12: Sensor Assembly

The preload servo system does not operate completely open loop. Two sensors are used to determine discrete locations, and a current-limited motor controller is used to set the

maximum output torque, and thus the preload force. The sensor system, shown in Figure 6.12 consists of two inductive proximity sensors and an electrical contact sensor. Non-contact sensors were chosen over mechanical and optical sensors because of their reliability, stability and insensitivity to dirt and dust. The third sensor did not require high switch-point repeatability and thus was designed as a low cost, mechanical contact. The inductive sensors are used to detect the edges of blocks mounted to the control pin and a sensor pin. The control pin sensor is used to determine the if the control pin is above the neutral position (docked) or below (undocked). The sensor pin is used to determine whether the module is latched, unlatched or clear of the groove. By using multiple edges and making binary combinations of these two sensor outputs, more than 4 operating states are possible.

6.5.3.4 Controls

Operation of the KDS is coordinated by an additional control system. There are five primary components of this system; operator pendant, motor controllers, processor, software, and support systems. With the exception of the pendant, all of the above components are housed in an external enclosure with cables connecting the modules and pendant in a hub configuration.

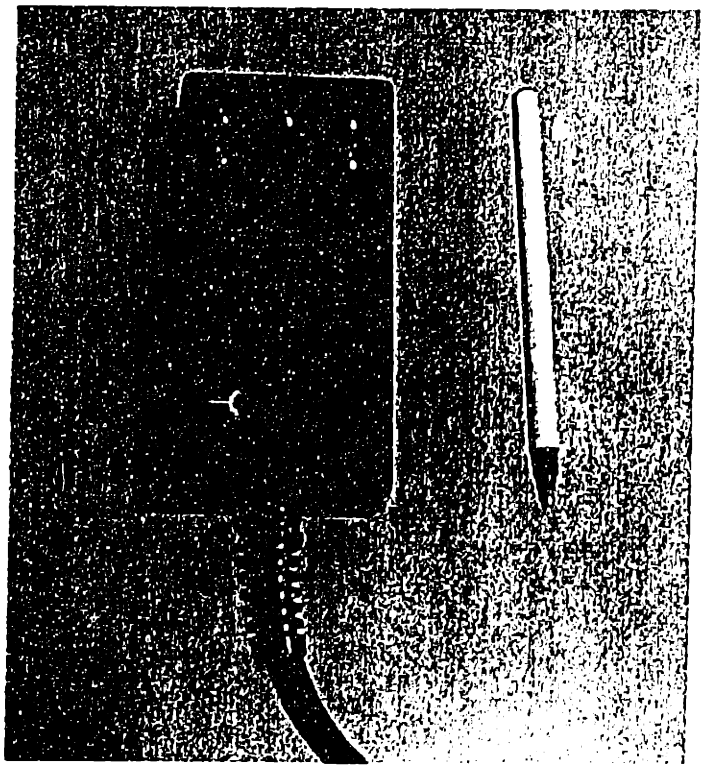


Figure 6.13: Operator Pendant

The complete user interface of the KDS is contained on the pendant shown in Figure 6.13. The two primary requirements of the pendant are to make the operation as simple as possible and to provide the operator with feedback on the status of the system. The two buttons activate the DOCK and UNDOCK operations and as a safety feature, pressing any button during operation temporarily halts all motion. A row of three multicolored LEDs indicate the status of each module, and a fourth LED indicates the status of the system. Green indicates that they system is ready, amber indicates that the system is in motion and red indicates an error. A fifth LED lights green when the docking process has been successfully completed and provides unambiguous feedback to the operator of a good dock.

The basic element used to control the ball module motors is a pulse width modulated (PWM) motor controller (National LMD 18245 DMOS full-bridge motor driver). This controller, shown schematically in Figure 6.14, allows for four-quadrant motor control. The brushed DC motor speed is determined by the 24 VDC input and the motor controllers are used to limit the maximum motor current, thus limiting the motor torque which determines the preload force. The controller is an off the shelf component and also provides thermal and

over-current protection. In addition to 24VDC power and 5 volt Vcc supply inputs, the motor controller has a 6-bit interface to control the motor. Two bits are used to select the quadrant of operation and the remaining four bits are used to select one of 16 discrete current limits. The range of possible current limits is determined through an RC network which is used to determine the pulse width and thus the average current.

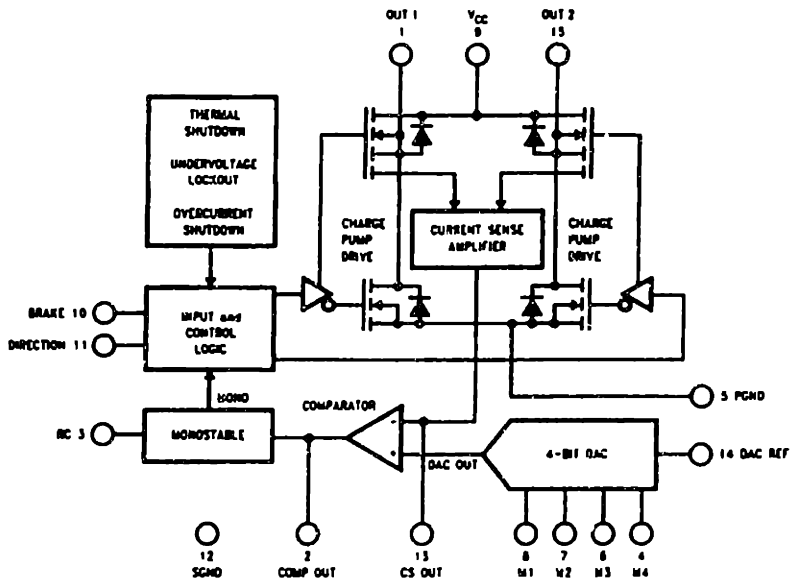


Figure 6.14: Motor Controller Schematic

Coordination of the pendant I/O, motor drivers and sensor status is determined by two finite state machine (FSM) models. Each module is described by an FSM that determines its docked state by monitoring the three module sensors. Each module can be in one of several states, such as DOCKED, NEUTRAL, RELEASED, etc... The complete system is described by a second FSM that determines the module commands by monitoring the state of each module and the output of the pendant. These state machines are described as either truth tables or logic flow paths. Figure 6.15 shows a schematic of the system state model. The entire model, including embedded test and debug routines, consumes approximately 128K of memory.

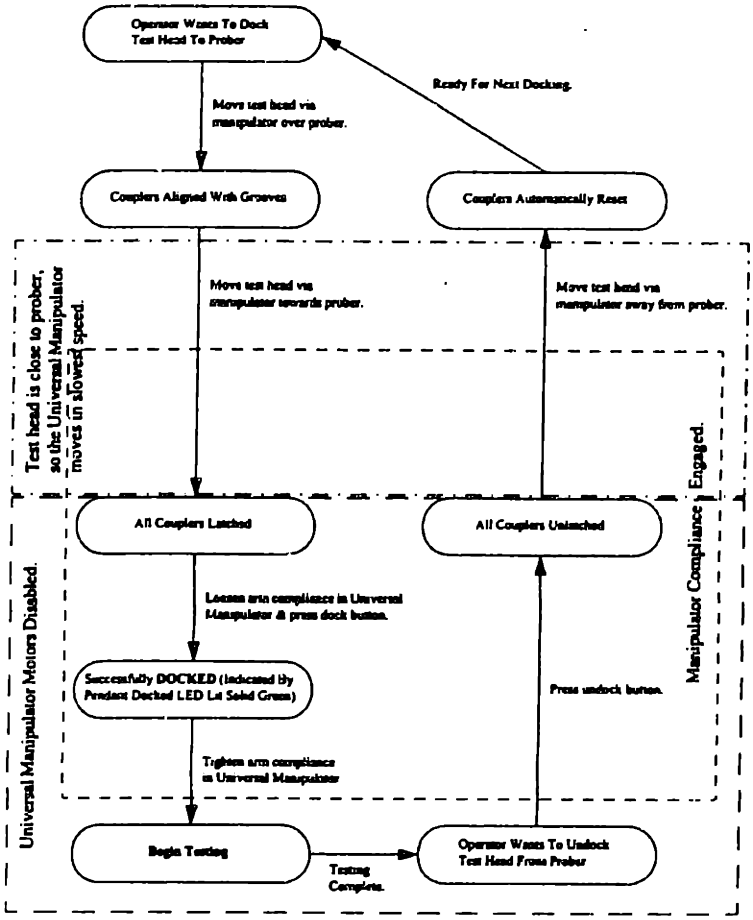


Figure 6.15: Finite State Machine Block Diagram

The FSM described above is implemented using a field programmable gate array (FPGA). Early prototypes of the KDS used discrete logic components (e.g. AND and OR gates) to implement this logic, however, as the features, functionality and number of sensors in each module grew, this approach quickly became large and difficult to implement. The current system has eleven input bits and 33 output bits. The FPGA provides the equivalent of 10K logic gates and can be reprogrammed using an erasable programmable read only memory (EPROM). This design provides the power required to control the large number of I/O's without the complexity of a full microprocessor design.

A number of other components that increase the systems functionality are included in the controller design. A 24V, 3A universal switching power supply is used to provide power to the system and on board regulators convert this to 5V for TTL and 15 V for the LED's. A serial bus is used to communicate between the pendant and FPGA. A second, parallel bus is

also present to allow the controller to be integrated with a testhead manipulator. This provides the ability to have a single, common pendant when used with the manipulator and provides a means to interlock the operation of the two systems and thus prevent accidental damage. A block diagram of the system controller is shown in Figure 6.16.

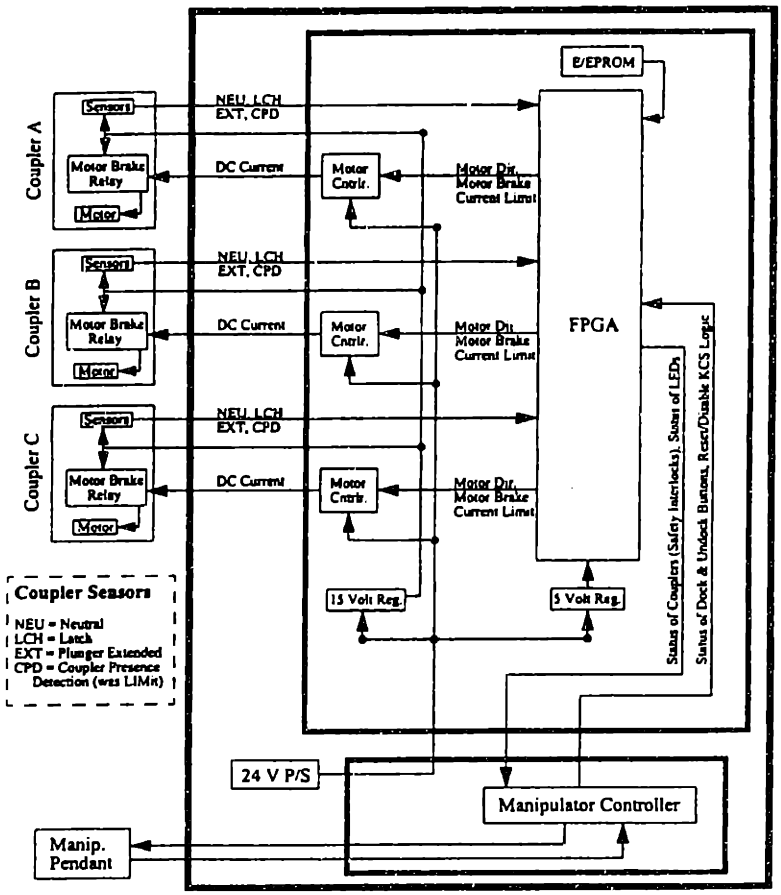


Figure 6.16: Block Diagram of System Controller

6.5.3.5 Adjustment Scheme

Although the kinematic design of the docking system provides high repeatability and stability, this alone is not sufficient to provide a practical docking system. Absolute accuracy of the docking system is also required to permit proper alignment between the testhead and device handler. The KDS is designed to work with any and all types of handling equipment, regardless of their design or manufacturing tolerances. Therefore, a means must be provided to adjust the relative positions of the critical interface components in the field. Ideally, a kinematic fixture requires only six adjustments to properly align the two mating bodies.

Practically, however, additional degrees of freedom are required to accommodate all types of equipment. In particular, typical test floor environments require frequent reconfiguration of test cells. Often a pool of device handlers will be used with a pool of testers and therefore any tester must be able to dock to any handler. For this reason, it is desirable to allow all interfaces to be calibrated to a common standard so that interfaces do not need to be realigned before every docking cycle. This fixture is used to align the X, Y and Z positions of each ball to a datum on the head and the rotational and yaw position of the grooves to a common datum on the plane of the handling device. The balls are aligned using a 3 ball contact or inverted tetrahedron and the grooves are aligned with a cylinder. Figure 6.17 shows the design of a single fixture that contains all of these features and can be used to support an entire test floor. The fixture shown below has four such features, to accommodate multiple ball/groove locations on different testheads.

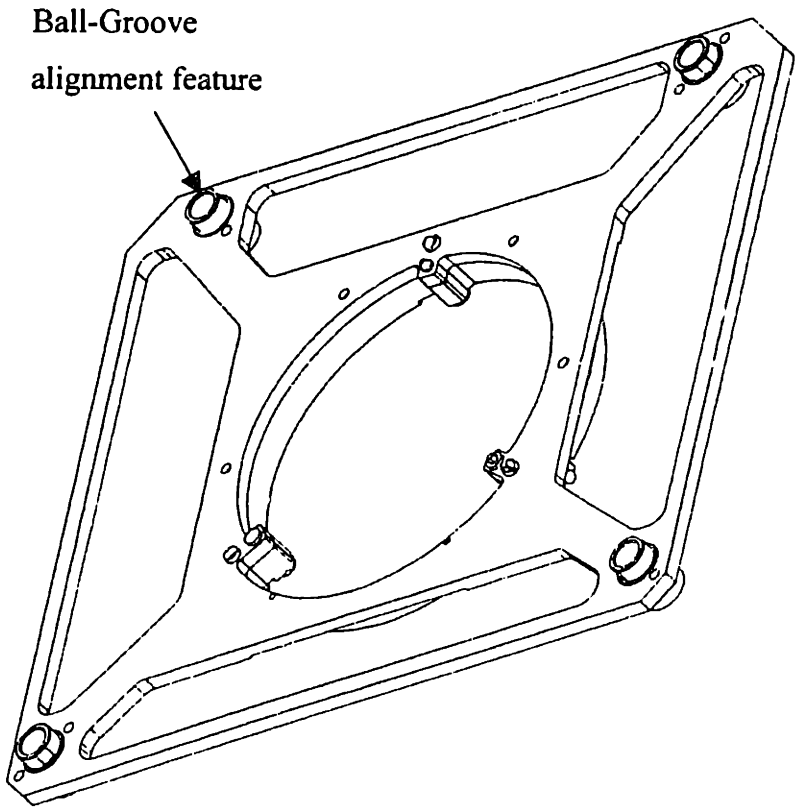


Figure 6.17: Setup and Alignment Fixture

The adjustments needed to calibrate a docking system require motion of both the ball and the groove. A groove adjustment range of +/- 0.075" is sufficient to accommodate

positional deviations from nominal due to handler setup and manufacturing errors and is implemented using oversized clearance holes for the mounting screws. X-Y adjustment of the ball requires the same +/- 0.075" range and is implemented with a modified bolt through the center of the ball with a large clearance hole for the latch pin. Ball Z-adjustment range must be approximately ½" to accommodate varying thicknesses of interface components. A preloaded lead screw is used to control the adjustment and crossed roller bearings are used to support the motion. Additional lock-down screws are provided to eliminate load transmission though the leadscrew. Flexures are cut around the screw mounting locations to eliminate overcompression of the roller bearings.

6.5.3.6 Design for Manufacture, Assembly and Use

Considerable design effort was given to the manufacture, assembly and use of the KDS components. Common design for manufacturability (DFM) guidelines were used to reduce the manufacturing cost and complexity of the individual components and to ease assembly efforts. For example, the majority of components can be machined in a single setup, features are added to prevent improper part orientation and fasteners are limited to 3 common sizes.

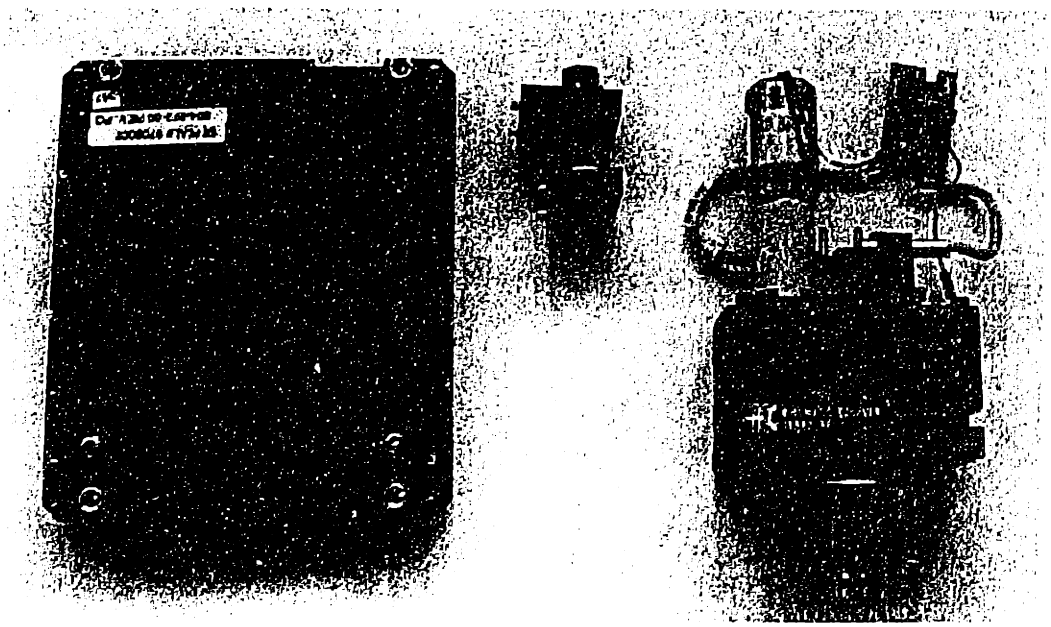


Figure 6.18: Ball Module Sub-Assemblies

In addition to component level DFM considerations described above, the assembly of the KDS ball module is further simplified through the use of component subassemblies. The five

subassemblies; the bearing stage, gearbox, z-adjust, wiring harness and covers, are shown in Figure 6.18. This decomposition allows each subassembly to be stocked, kitted, assembled, calibrated and tested separately.

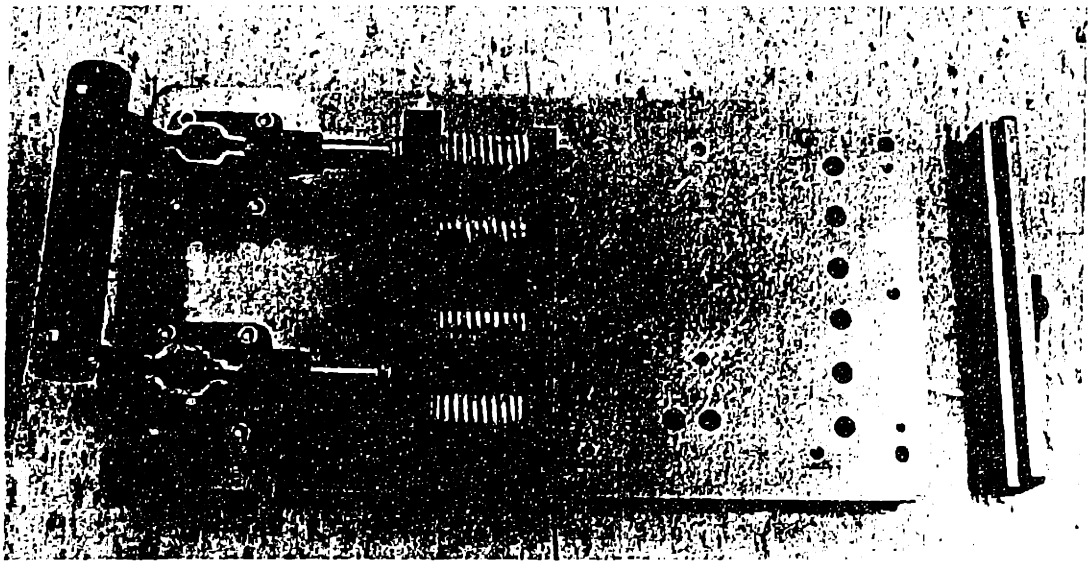


Figure 6.19: Bearing Preload Assembly Fixture

A number of fixtures, such as the bearing preload fixture shown in Figure 6.19 were designed to facilitate the assembly of the ball module and to reduce the dependence on operator strength, training and skill. The first fixture is used to simultaneously preload the crossed roller bearing stage and orient it relative to the datum defined by the mounting plate. A second fixture is used to mount and calibrate the two inductive proximity sensors, thus eliminating any adjustment procedures. The third fixture, used to assemble the slip-clutch mechanism, prevents assembling the components in the incorrect order and eases the insertion of a delicate snap-ring. A final fixture, shown in Figure 6.20, is used to verify the proper assembly and performance by cycling a module and measuring the preload force and cycle time. A majority of faulty components or improper assembly will be detected in this final Q/A fixture.

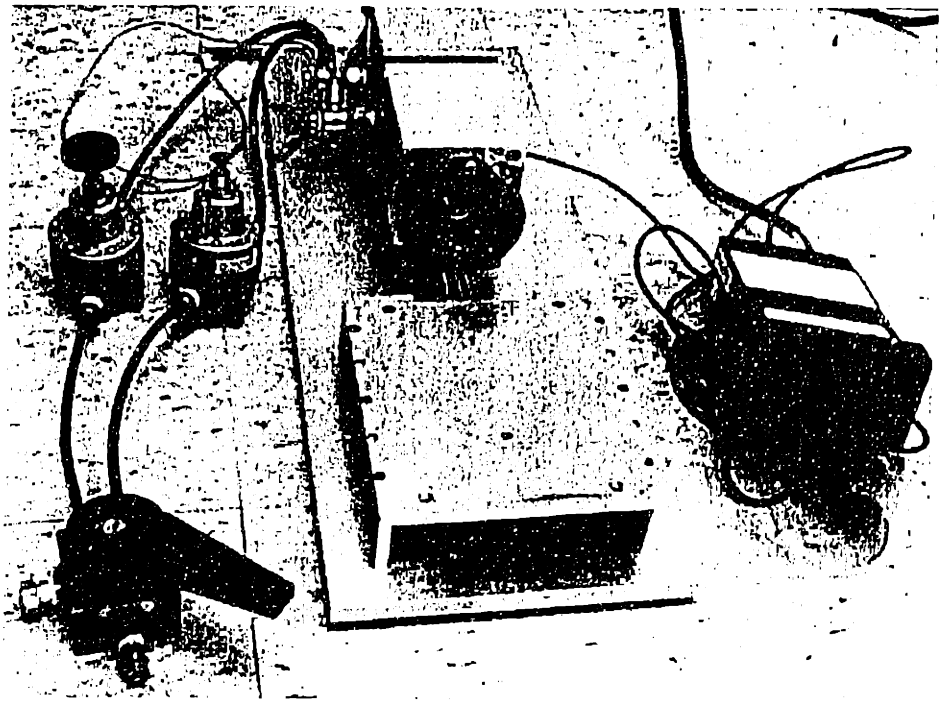


Figure 6.20: Final Test Fixture

6.6 Summary of Chapter 6

This chapter describes the need for test cell interfaces and presents an overview of the state-of-the-art in production interfaces along with some of the primary design trade offs and flaws. This is followed by a definition of the functional requirements and design concepts used in the development of a new interfacing system. In particular, the concepts of kinematic coupling and deterministic constraint were used extensively throughout the design. The bulk of the chapter is devoted to describing the execution of the *Kinematic Docking System*, including the relevant subsystem designs, material selections, operating principles and manufacturing processes.

7 Results of *Kinematic Docking System*

7.1 Introduction

The *Kinematic Docking System* concepts described in Chapter 6 were developed into a finished product as part of this thesis effort. At the time of this writing, approximately 100 systems have been shipped and are in use in semiconductor test floors throughout the world. At several points within the development of this project, the design was tested to verify correlation between the predicted and actual performance. The data and observations collected from these tests were used to support and improve the development of the design concept into a working, shippable product.

This chapter discusses the theoretical and measured performance of the *Kinematic Docking System*. It begins with the development of a simple error budget analysis of the test cell structural loop to help develop a base-line understanding of what level of accuracy is possible for an interfacing systems. The bulk of the chapter contains data collected at three distinct stages of the development. Tests were performed at the proof of concept stage to verify the conceptual design. Further tests were done on the first working prototypes at a customer site to confirm that the basic implementation was valid and to uncover any application specific problems that were overlooked. Finally, a set of tests were performed on the released version to measure the actual performance of the product in a controlled environment, independent of external influences. The chapter closes with a discussion of future work and some derivative designs.

7.2 Theoretical performance

The mechanical accuracy performance requirements of the KDS discussed in Chapter 6 prescribe a repeatability of 25 microns and stiffness of 10 N/μm. These parameters are considered to be typical worst-case requirements for a wafer probe test cell. Prior to the development and testing of a prototype system, some investigation was undertaken to determine if these levels of mechanical performance were achievable.

The most rigorous method of analyzing the accuracy and stiffness of a mechanical assembly is through the use of homogeneous transformation matrices[44] (HTM) to develop a complete error budget of a machine. This method has previously been used to predict the performance of semiconductor processing equipment[45] and uses a 4x4 matrix to describe the geometry and error motions of each component within the structural loop. Thermal, Abbe and deflection errors can all be described as combinations of linear and angular elements of these matrices. In general, HTM models only apply to deterministic systems, overconstrained structures require FEA or iterative methods to solve. In addition to this, HTM methods most easily apply to rectilinear chains of axis, such as those found in a vertical machining center or a prismatic robot arm.

Since the KDS consists of only one primary interface element, the kinematic coupling, a more relaxed procedure can be used to evaluate its accuracy performance. Because the tooling is located at the center of the coupling, the system is insensitive to Abbe errors and a simple error budget consisting of the linear sums of expected error motions is adequate for this analysis. The following factors were considered in the KDS error budget:

- Repeatability and stiffness of kinematic coupling.
- Accuracy and repeatability of fixturing for assembly and setup.
- Stiffness of interface components in structural loop.
- Thermal environment.
- Uncontrolled components within the structural loop include the prober structure and motion stages, the prober head plate.

Because the KDS is aligned to the wafer prober using a fixture and because the prober has closed-loop positional control, the manufactured accuracy of the static mechanical components within the structural loop can be ignored. Only the accuracy of the fixturing and the stiffness and repeatability of the kinematic interface components affect the overall accuracy of the interface. The contribution of these components to the overall error budget are discussed below.

[44] Slocum, 1992
[45] Van Doren, 1995

7.2.1 Kinematic Coupling Analysis

There are several methods of evaluating the repeatability and stiffness performance of a mechanical assembly. Most commercial FEA packages have some features for tolerance studies and bolted joint analysis, however, these are generally difficult to apply to 3-dimensional structures with non-prismatic features such as a 3-ball, 3-groove kinematic coupling. However, because the kinematic coupling is completely deterministic, it is possible to directly calculate it's stiffness and error motions in closed form. Stiffness of the coupling is determined by the material properties, preload force and geometry of the kinematic contacts. The error motions can then be found by applying an external disturbance force to the coupling. A spreadsheet based tool developed by Slocum[46], shown in part in Figure 7.1, is able to predict both the stiffness and error motions of any generically defined coupling. This spreadsheet was used in the concept development stage to predict the Kinematic Dock contribution to the overall interface error budget.

XY plane is assumed to contain the ball centers						
For standard coupling designs, contact forces are inclined at 45 to the XY plane						
Standard 120 degree equal size groove coupling?	TRUE					
For non standard designs, enter geometry after results section						
Dbeq=	0.0381	Equivelent diameter ball that would contact the groove at the same points				
Rbminor =	0.01905	Minor radius	0.009525			
Rbmajor =	0.01905	Major radius	#NUM!		Material properties	
Rgroove =	1.00E+06	Groove radius (negative for a trough)			*Hertz stress*	
Costheta =	FALSE	Is ball major radius along groove axis?			SiN	6.75E+09
Dcoupling=	1.000	Coupling diameter			Aluminum	5.00E+08
Fpreload =	1000	Preload force over each ball (N)			RC 62 Steel	1.81E+09
Xerr =	0.000	X location of error reporting			*Elastic modulus*	
Yerr =	0.000	Y location of error reporting			SiN	3.11E+11
Zerr =	0.000	Z location of error reporting			Aluminum	7.50E+10
Auto select material values assume that metric units are used (mks)					RC 62 Steel	2.04E+11
Matlab =	4	Enter 1 for aluminum ball, aluminum groove			*Poisson ratio*	
		Enter 2 for Steel ball, aluminum groove			SiN	0.27
		Enter 3 for SiN ball, RC 62 Fe groove			Aluminum	0.27
		Enter 4 for RC 62 Fe ball, RC 62 Fe groove			RC 62 Steel	0.29
		Enter 5 for other values and enter them for each ball and groove				
Min. yield strength (Pa, psi)	1.81E+09	262500				
Applied forces' Z,Y,Z values and coordinates		**Coupling centroid**				
FLx =	150.00	XL =	0.000	xc	0.000	
FLy =	150.00	YL =	0.000	yc	0.000	
FLz =	150.00	ZL =	0.000	zc	0.000	
Results: Hertz stresses and deformations						
Error displacements at the point of interest (micron)						
DeltaX	3.11E-07	DeltaY	3.11E-07	DeltaZ	3.04E-07	3.11E-07

Figure 7.1: Kinematic Coupling Analysis Spreadsheet

[46] Slocum, 1992

Preliminary analysis was performed using the following nominal parameters:

- Coupling Diameter: 1 m
- Ball Diameter: 38 mm (1.5")
- Material 440C stainless, RC 58
- Preload Force 1000 N (200 lbs)
- Disturbance Force 250 N (50 lbs in XYZ direction)

The above analysis predicts that the coupling should be contribute +/- 5 microns of repeatability error to the overall system when subjected to a 50 Newton reversing load and have an overall minimum stiffness of approximately 50 N/μm.

7.2.2 Fixture repeatability

The KDS setup fixture is intended to be used as a standard to calibrate all interfaces on a test floor. Therefore, it does not require tight manufacturing tolerances, it need only position the interface components to within a nominal tolerance. However, the fixturing process does need to be highly repeatable so that the mechanical calibration of the interfaces is consistent across an entire test floor. Again, the best method for achieving the highest level of repeatability performance is to use a kinematic fixturing design. The fixture, described in detail in Chapter 6, must align both the balls and the grooves to a common set of datums at the center of test. Each ball must be aligned in three degrees of freedom (x, y, z) and each groove must be aligned in two (2 planar angles). The ball is constrained in the fixture by contacting three smaller balls that located about the OD of a bore, resulting in 3 contact points constraining the 3 degrees of adjustment. Each groove is quasi-kinematically located by using a cylinder that constrains both the planar angles of orientation. Because the fixturing of both the balls and grooves are done kinematically, this operation should have a repeatability similar to that of the actual coupling analyzed above. The actual fixturing repeatability should be somewhat better because of the lack of external disturbance forces during the alignment process.

7.2.3 Stiffness of interface components

A deterministically designed structure implies that there is a single force path supporting each degree of freedom. Therefore, each component in the structural loop of the interface must be sufficiently stiff to meet the design goal. Close attention must be paid to the stiffness the structural ball and groove module components because the interface is only as

stiff as it's most compliant link. In all but the simplest geometries, designing for stiffness is requires either FEA analysis, highly conservative deflection models or a traditional build and test approach. Most mechanical interface components are generally designed such that they will not yield under design loads, which usually provides sufficient stiffness. However, some components require closer attention, in particular those that permit motion or adjustment such as bearings, leadscrews and lock-down bolts. There are several such components in the ball module portion of the kinematic interface, including the crossed-roller bearings, Z-adjust leadscrew and ball lock-down bolt. All of these components must be chosen, mounted and applied in such a way as to produce sufficient stiffness and eliminate any free-play or lost motion.

The standard method for eliminating free play between moving components is to apply a preload force to the bearing surfaces. As long as the preload force is greater than any disturbance force, there will be no off-axis relative motion between the two components. Both the crossed-roller bearings and the Z-adjust lead screw are preloaded to eliminate lateral motion and screw backlash during the adjustment operation. In addition to this, lock-down screws are used to fix the Z position after adjustment to reduce the load supported by the leadscrew. All of these features effectively reduce the module stiffness analysis to that shown in Figure 7.2, which is primarily dictated by the bolted joints. Preliminary estimates of 250 N/μm of the module assembly stiffness, measured between the ball and mounting plate, were later confirmed by experimental stiffness measurements. This measurement was performed by mounting a module to a plate, applying a 500 N reversing lateral load to the ball and measuring it's deflection relative to the plate. This measurement captures the stiffness of all the components in the structural loop of the module.

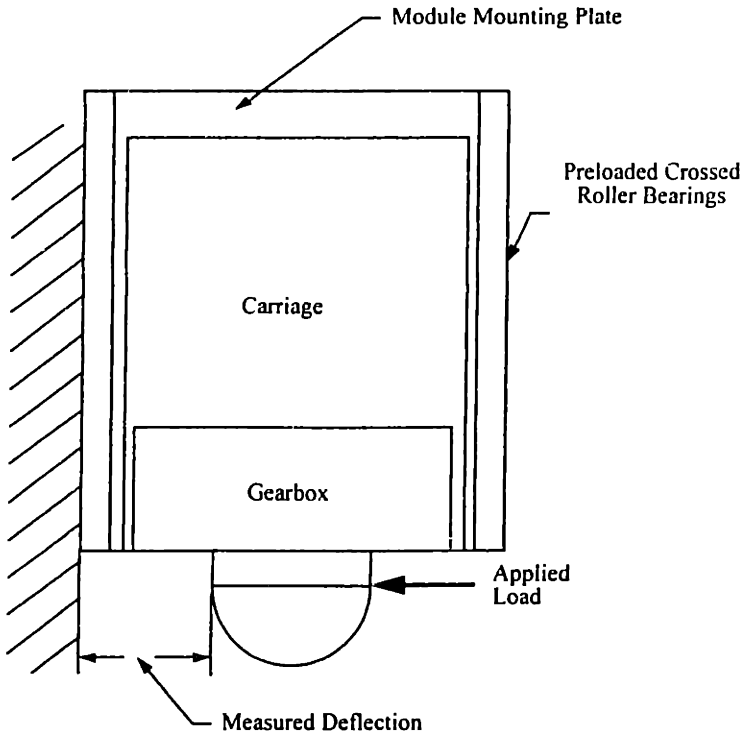


Figure 7.2: Ball-Module Stiffness Measurement

There are other, uncontrolled components with the structural loop include the prober and testhead structures, as well DIB and DIB stiffener and inner docking mechanisms. Manufacturing inaccuracies in these components are eliminated with the use of the setup fixturing. These components are essentially static during the test operation and therefore do not contribute any tolerance stack-up to the error budget. They do contribute errors due to deflections under load, however, the primary source of error motion under load, 'oil canning', is eliminated through the 3-point attachment of the testhead to the prober structure via the KDS. Beyond this, there is little that an interface can do to improve the overall stiffness of the structural loop. Therefore, structural error motions are limited to linear deflections of the testhead and prober structures, both of which are typically robust enough to support a 250 N load with less than 2.5 microns of deflection.

7.2.4 Thermal performance

Test floor environments generally need to be controlled to better than +/- 2° C to maintain instrument calibration. However, local gradients within and around the interface may exist due to DUT thermal conditioning equipment in the device handling equipment or

from heat generated within the testhead. Both the environmental temperature changes and local gradients may cause differential thermal expansion structural loop and lead to error motion.

Most errors due to thermal expansion can be eliminated by matching the coefficient of thermal expansion (CTE) of all materials within the structural loop and by using thermocentric design. Both testhead and prober metrology frames are typically constructed of aluminum as well as most interface components, with the exception of the hardened bearing surfaces. Nominal layout of the of the modules about a testhead is also in a symmetrical, 120 degree pattern, and thus is thermocentric. Thermal errors are therefore limited to those caused by local thermal gradients, which are both unlikely to occur and whose effects difficult to predict. However, because there is only a small area of contact, and thus a narrow heat path, between mating halves of a kinematic coupling, local gradients may take a long time to dissipate. It is estimated that 2° C gradient across a 1m coupling diameter will cause less than 5 microns of lateral and Z error motion at the center of test.

7.2.5 Overall Predicted Error Budget

If the above error sources are summed, the result is the worst-case, expected repeatability of a the KDS. The estimates described above suggest that this system should be capable of a repeatability of approximately 15 microns and a stiffness of approximately 40 kN/mm. These estimates seem reasonable for a light structural assembly on the order of several meters cubed. This level of predicted precision is sufficient to meet current and near future test cell interfacing requirements. The remainder of this chapter describes a series of tests that were performed at various stages of the KDS development to assess the actual accuracy, repeatability and stiffness of the KDS system.

7.3 Kinematic Docking System Tests and Results

Tests were run on prototype docking systems at three distinct stages of its development. At the proof-of-concept stage, a testbed consisting of prototype modules, testhead and prober structures, was used to verify that the kinematic theory was applied correctly and to experiment with alternative design concepts and options. The second phase of tests were done on a customer test floor using pre-production modules to verify the design in a production environment, measure the effect of the design on test cell throughput and signal

integrity and to gather customer feedback. Finally, a third set of tests were performed on a production system in a controlled environment to verify that the final design was performing to specification. The test fixtures and procedures used in these tests and their results are discussed in the sections below.

7.3.1 Proof of Concept Tests

Early within the development process, a set of tests were performed on preliminary proof-of-concept prototypes to verify that the kinematic concept and proposed design implementation were sufficient to produce the required repeatability performance. The prototypes were also used to investigate the feasibility of alternative design concepts and functional enhancements. Data from these tests confirmed that the modular, three ball, three groove kinematic coupling concept is capable of producing mil repeatability in testhead interfacing applications. In addition to this, an enhanced version of the module that allowed closed-loop, servo-controlled planarization of the testhead tooling with respect to the DUT was tested. This additional feature eliminates the need for any manual adjustments and fixturing when docking to wafer probe systems. The results of these tests are discussed below.

7.3.1.1 Test Fixture and Procedure

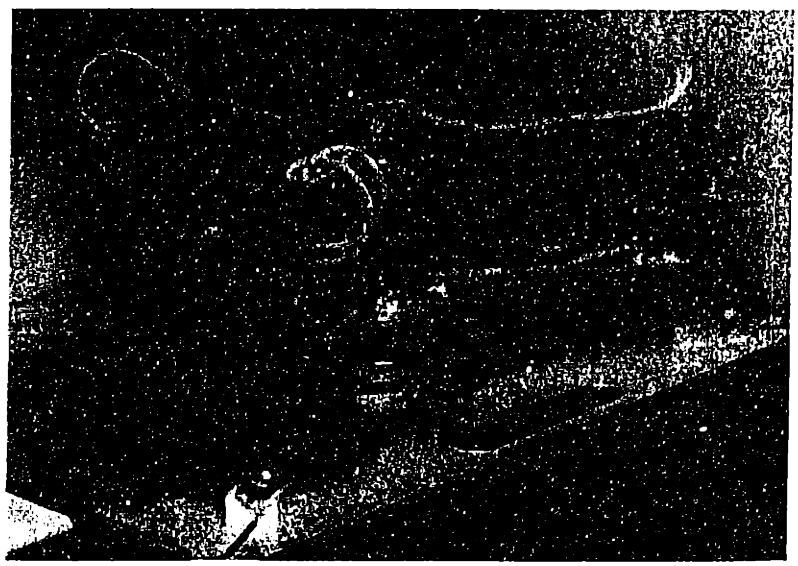


Figure 7.3: Proof of Concept Prototype Test Fixture

Two testbeds were developed for the proof of concept prototype evaluation. A fixture was built for preliminary tests to simulate the testhead and prober metrology frames. Prototypes and instrumentation were later transferred to an actual testhead casting once it was available. The prototype modules and fixtures used in the preliminary tests are shown in Figure 7.3. This testbed consisted of the following components:

- Prototype modules with pneumatic preload and manual latching mechanisms.
- Prototype aluminum grooves.
- ½" thick aluminum jig plate (simulates prober structure).
- 30" diameter by 10" tall aluminum weldment (simulates testhead structure).
- 3 digital linear encoders to measure the z, roll and pitch positional repeatability.

Both of these frames were designed with a minimum bending stiffnesses of 10 kN/mm in the critical directions. The maximum design loads of 100 N would yield at most 10 microns of error motion and the linear encoders used had a repeatability of 2.5 microns. Therefore, the total expected accuracy of the preliminary test results were approximately ±12 microns, or 0.5 mils. Initial test results yielded a repeatability of better than 125 microns. Most of this error can be attributed to deformation in the aluminum grooves and to motion of the non-preloaded linear bearings used in the prototype modules. These issues were addressed in future designs by using heat treated 440C stainless steel for the ball and groove interfaces and by changing to preloaded crossed roller bearings.

A second set of concept evaluation tests were performed using an actual testhead structure and updated prototype modules. These tests improved on the previous repeatability tests as well as confirmed the viability of the autoplanarization concept. Figure 7.4 shows the components used in this set of tests.

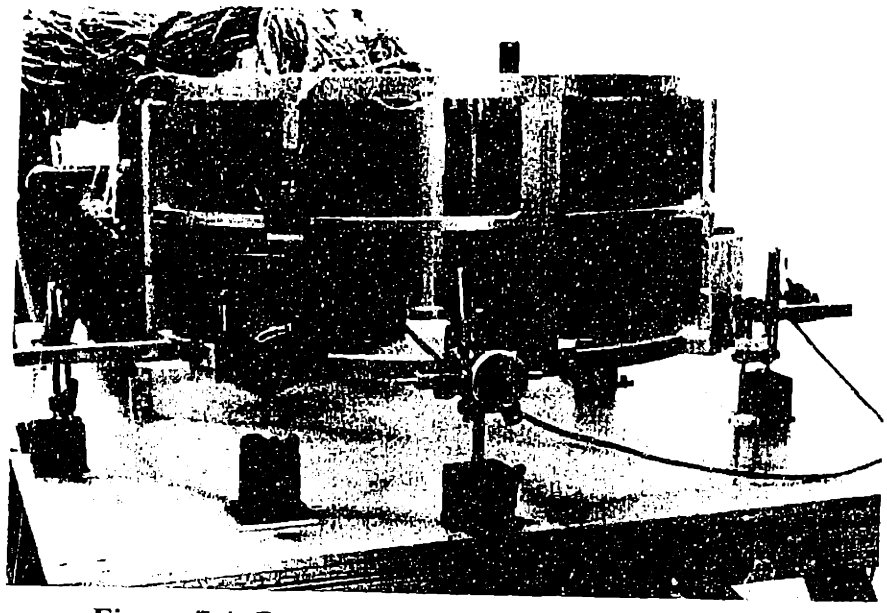


Figure 7.4: Repeatability Using Testhead Casting

This test setup yielded considerably more accurate results and higher module repeatability performance. In addition to the substantially stiffer testhead casting, a 6" deep torsion box was welded to the bottom of the groove mounting plate to increase bending stiffness. Also, the modules were upgraded by adding hardened steel ball and groove components and preloaded linear bearings. Three more linear encoders were also added to capture x, y and yaw repeatability data. An additional servo system was also added to the module z-adjust mechanism to automatically adjust the testhead planarity after the docking was complete. The upgraded modules are shown in Figure 7.5.

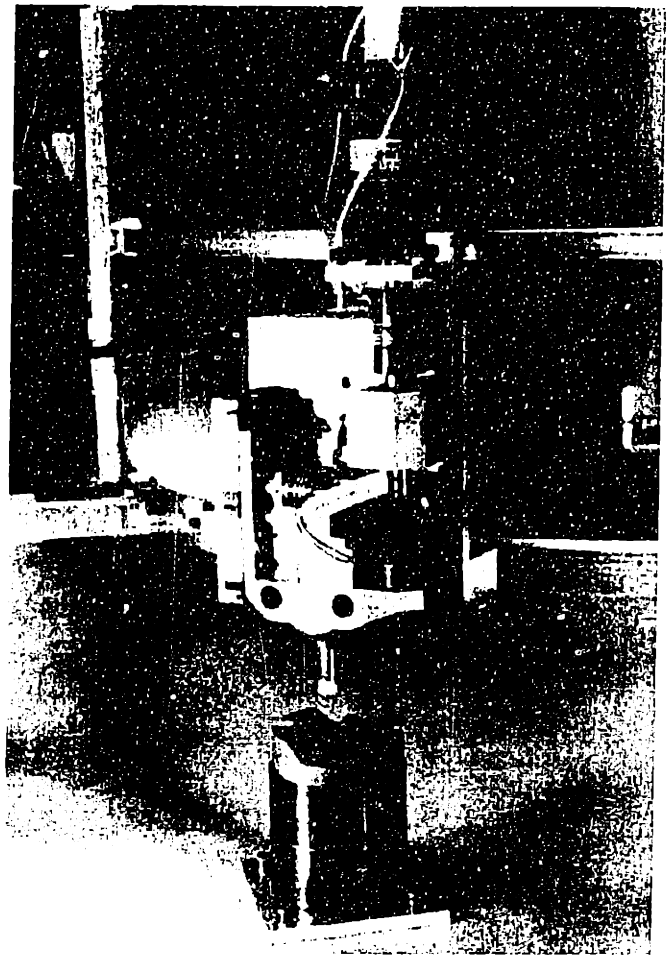


Figure 7.5 Autoplanarization Prototype Module

Repeatability measured using this test setup improved; linear errors were less than 2 mils in the x, y and z directions and angular errors were less than 20 arc seconds. The autoplanarization system used feedback from the linear encoders to get positional information on the height, roll and pitch error motion of the testhead. An algorithm in the motor controllers then calculated the required corrective motion of each module and drove the servos to the proper position. This code was implemented on an external PC using a DSP board and the LabView programming environment to execute the calculations and code. Using this system, repeatability increased to that of the resolution of the servo system, which was approximately 2.5 microns. Test on this system confirmed that a closed-loop, autoplanarized kinematic testhead coupling system is capable of better than 10 microns and 5 arc-second positional repeatability. Additional data on this system can be found in Appendix 9.2.

7.3.2 Pre-Production Field Tests

A third series of tests were performed on a pre-production system to verify that the design implementation would function in a production environment. Field tests were performed on the *Kinematic Docking System* as well as two existing systems, the Teradyne J-ring and a third party interface based on a pin-bushing pull-down block (PBD) design. The following data was collected:

- Test head docking repeatability.
- Probe card planarity.
- Interface static stiffness.
- Interface dynamic stiffness.

Each of the sections below briefly describe the test, its relevance and present some representative data. In addition to the above mechanical tests, yield data and operator feedback was collect and is briefly discussed below.

7.3.2.1 Test Environment

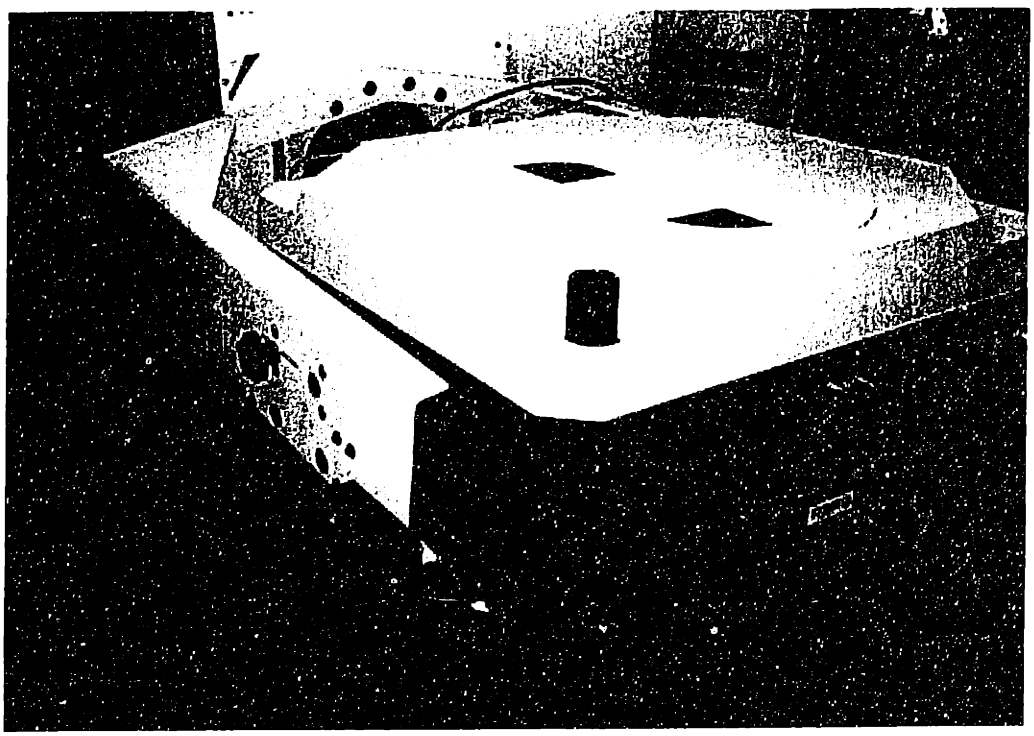


Figure 7.6: AMS testhead with KDS Ball-Modules

The tests were performed by retrofitting two existing test cells to accept the KDS. The first application used a Teradyne A580 Advanced Mixed-Signal (AMS) test head docked to a

KLA 1201 8-inch wafer prober. Additional structures were added to the test head and prober to accommodate the mounting of the modules and grooves. The KDS replaced an existing J-ring interface. Mechanical measurements and yield data were collected on both interfaces.

The second application also used a Teradyne AMS test head but it was instead docked to a DeltaFlex 1010 pick-and-place handler. Support structures similar to those above were also added to the test head and handler to support the grooves and modules. In this case, the KDS replaced an existing PDB interface. Mechanical data was collected for both of these interfaces, but no yield data was available. A photo of the modified test head is shown in Figure 7.6and schematics of the prober and handler applications are shown in Figure 7.7.

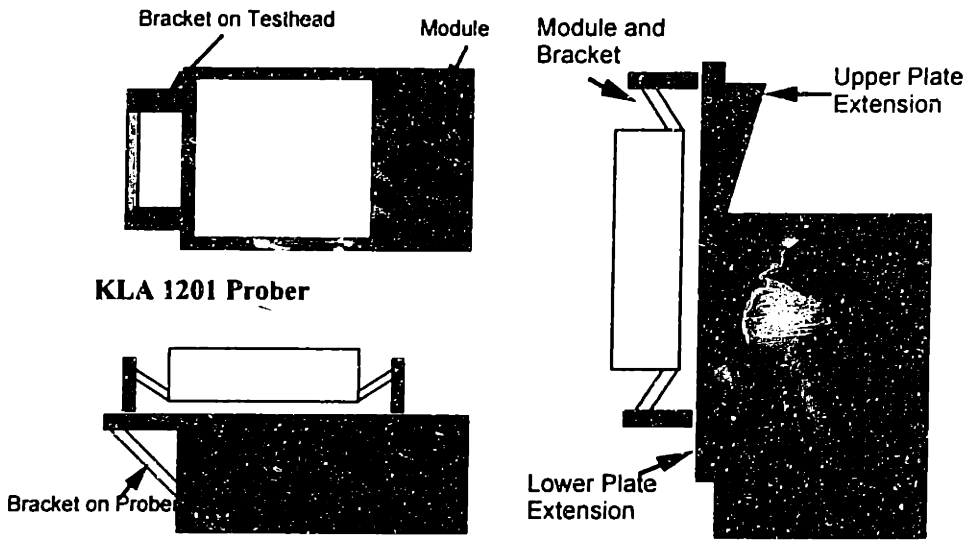


Figure 7.7: KDS Prober and Handler Test Applications

7.3.2.2 Positional Repeatability Tests

The positional docking repeatability of the test head relative to the prober and handler was measured using digital linear encoders mounted at the position shown in Figure 7.8.

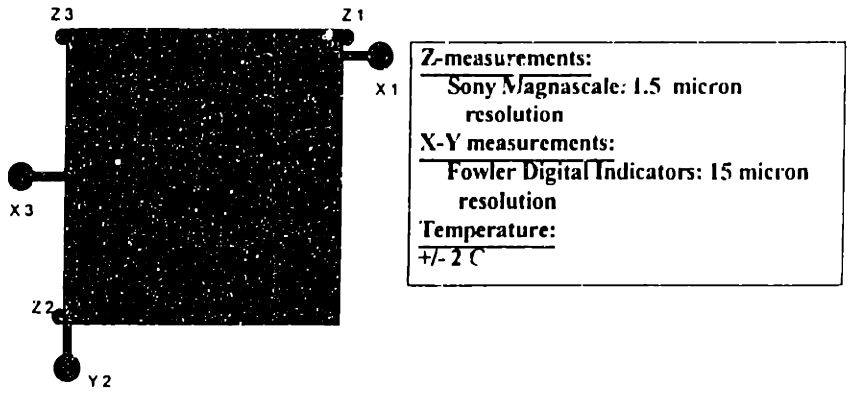


Figure 7.8: Measurement Locations and Instrumentation

Probes were zeroed on setup and measurements were made between each of 30 docking cycles. These tests were repeated for the J-ring and PDB interfaces. Figure 7.9 shows the results of these tests. The *Kinematic Docking System* had a bilateral repeatability of approximately 60 microns. Much of this error was due to movement in the testhead and prober metrology frames due to cable and manipulator loads. This was 5 times better than the PDB interface and nearly 2 orders of magnitude better than the J-ring.

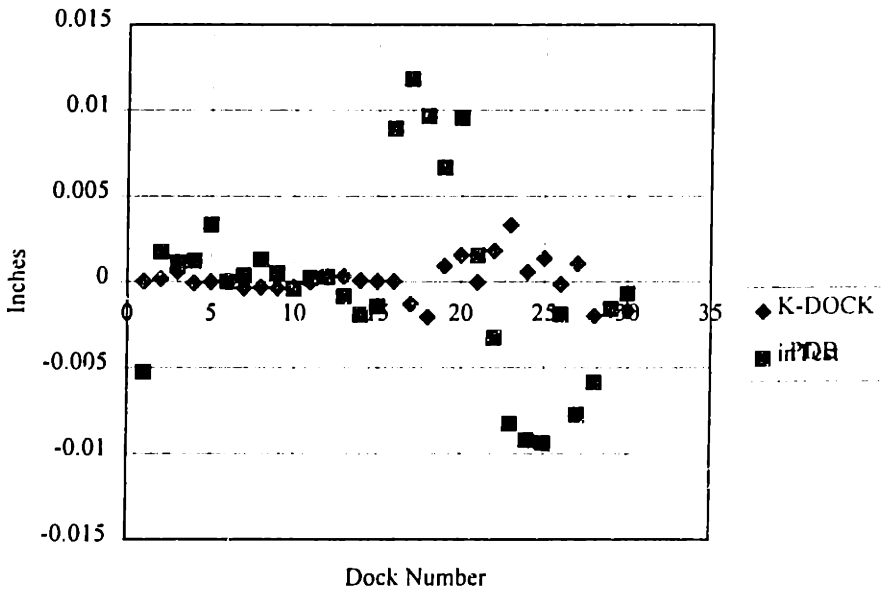


Figure 7.9: Results of Docking Repeatability Measurements

7.3.2.3 Probe Card Planarity

An additional set of repeatability tests was performed on the probe application. This test measures the effect of docking errors on probe card deformation. Even if proper signal continuity is achieved at all other layers of the interface, small deformation of the probe card, on the order of 20-40 microns, can adversely affect the probe to pad pointing accuracy. Measurements were made using a capacitance probe attached to the chuck and stepped across the probe card surface. The setup is shown in Figure 7.10.

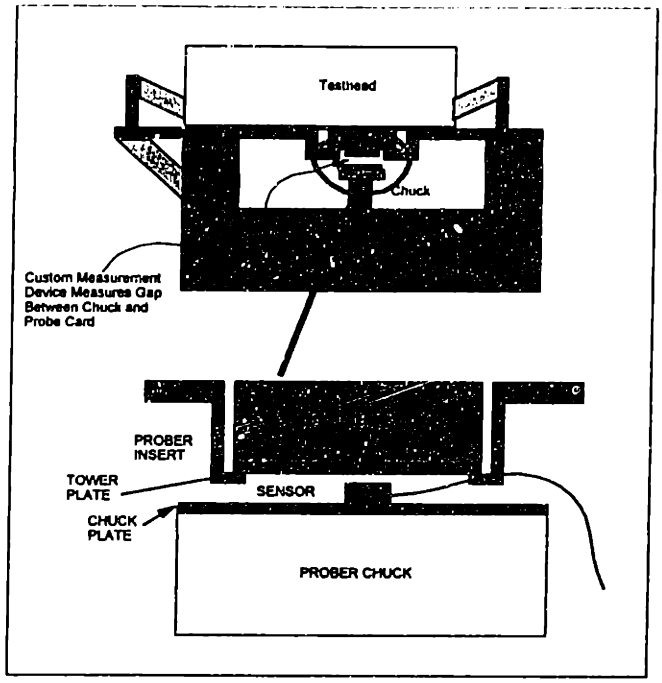


Figure 7.10: Probe Card Planarity Test Setup

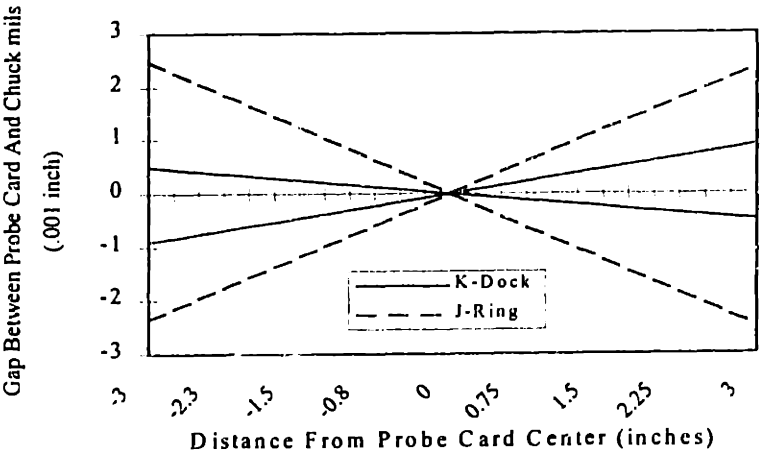


Figure 7.11: Planarity Test Results

The results of the planarity test, shown in Figure 7.11, demonstrate that the *Kinematic Docking System* provided a 3x improvement in probe card planarity over the existing J-ring design. No planarity data was taken on the handler interface because there is not an equivalent measurement.

7.3.2.4 Static Stiffness Tests

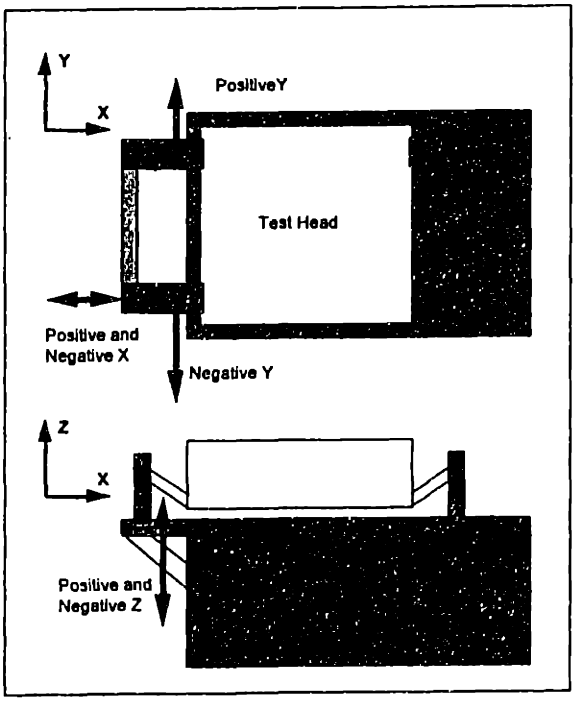


Figure 7.12: Static Stiffness Test Setup

The static interface stiffness was measured for the probe application by applying to the test head while it was docked. The error motions and deflections of the test system were measured using the same metrology as the repeatability and probe card planarity measurements described above. A 200 N load was applied using a load cell and weights in orthogonal direction shown in Figure 7.12. This test is an indication of how well an interface can tolerate external static loads, such as cable hysteresis, improper manipulator setup or an operator leaning on a test head.

The repeatability under load tests demonstrate that the kinematic interface maintained the two orders of magnitude improvement over the J-ring interface. Figure 7.13 above shows the probe card planarity under load data. The kinematic interface maintained a 10 microns/mm (1.0 mils/inch) slope window, whereas the j-ring maintained a 30 microns/mm (2.8 mils/inch) slope window. The kinematic interface had nearly three times less deflection of the probe card due to external forces than the j-ring.

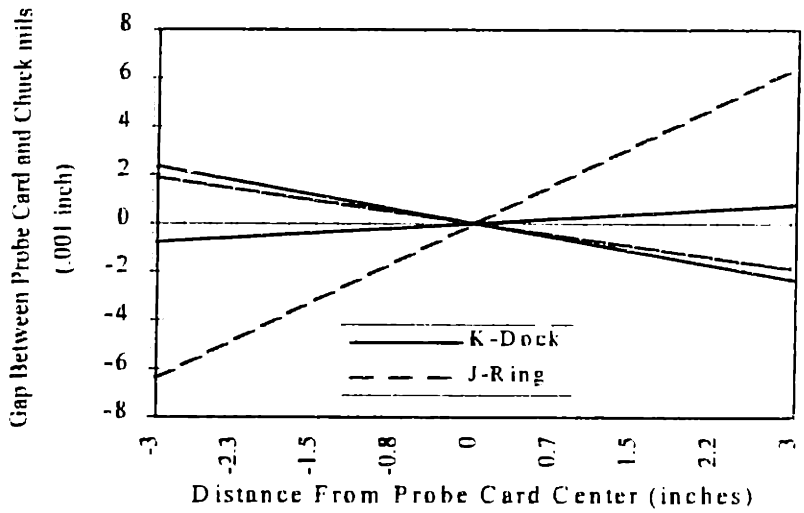


Figure 7.13: Probe Card Planarity Under Load

7.3.2.5 Dynamic Stiffness Tests

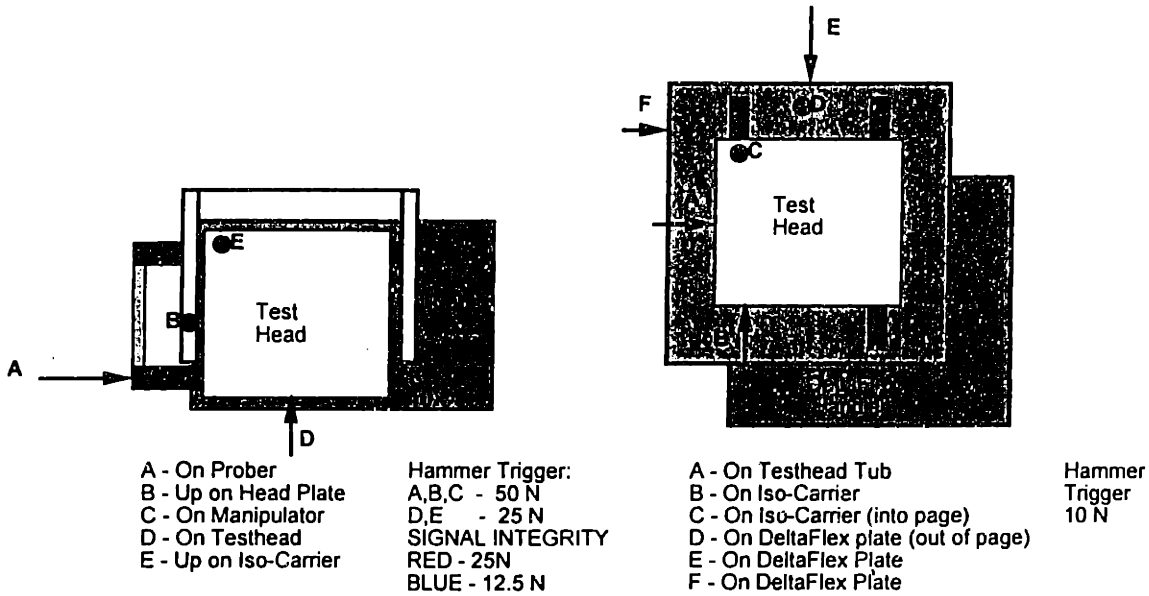


Figure 7.14: Dynamic Stiffness Measurement Setup

Dynamic stiffness of the interface was measured using standard impulse and swept sine techniques. Energy was applied to the testhead with a shaker and calibrated impulse hammer at the locations shown in Figure 7.14 and the output was measured with an accelerometer at a location near the probe card. Stepping motion of the prober chuck, insertion forces from the handler and other random vibrations transmitted through the floor and manipulator can all cause motions at the probe card level. This series of tests is useful in measuring the relative impact of those vibrations on the probe card. Figure 7.15 is a plot showing the transfer function of the probe card displacement to the impulse input for the Kinematic and J-ring interfaces. Also on the same graph is the coherence data, which is a measure of the data integrity at the frequency of interest. For vibrations of frequency greater than 100 Hz, the kinematic interface and j-ring exhibited the same magnitude of vibration at the probe card. At low frequency (below 50 Hz) the kinematic interface has about an order of magnitude less deflection at the probe card than the j-ring. The higher motion of the J-ring interface is due to structural motions of the testhead cantilevered off of the manipulator arm. The large apparent deflections at frequencies below 10 Hz are artifacts of the integration of acceleration data. The relative performance between the two interfaces at these frequencies is still valid, but the absolute values of displacement are not.

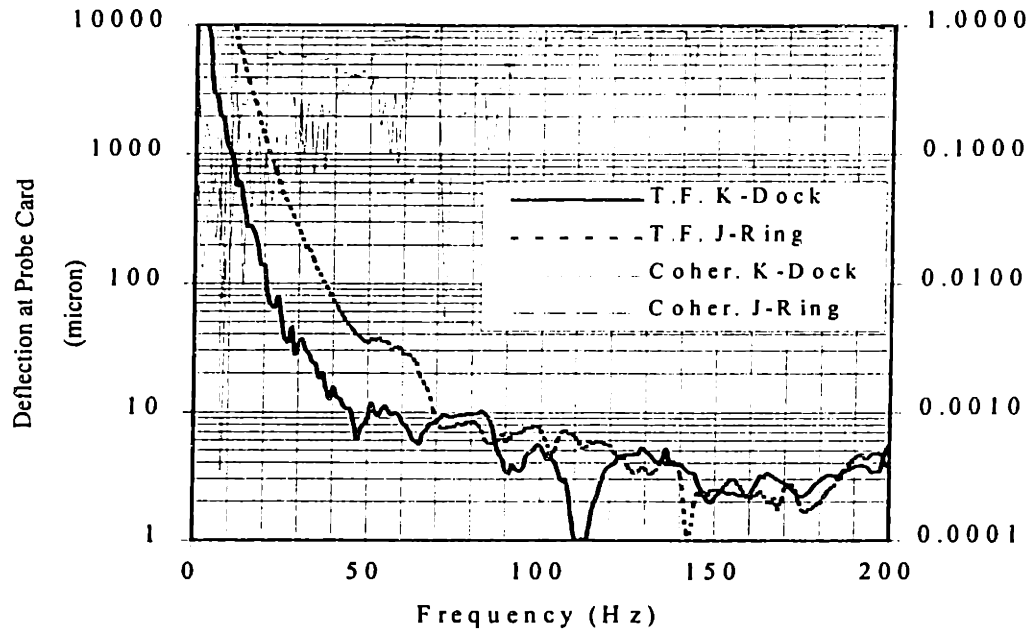


Figure 7.15: Vibration Transmission and Coherence for KDS and J-ring

Results for the vibration measurements of the KDS to the PDB interfaces are similar to those shown above and are shown in Figure 7.16. The KDS demonstrated an order of magnitude increase in dynamic stiffness over inTest up to nearly 200 Hz. The lower coherence of the PDB data is likely due to lost motion at the pin-bushing interface.

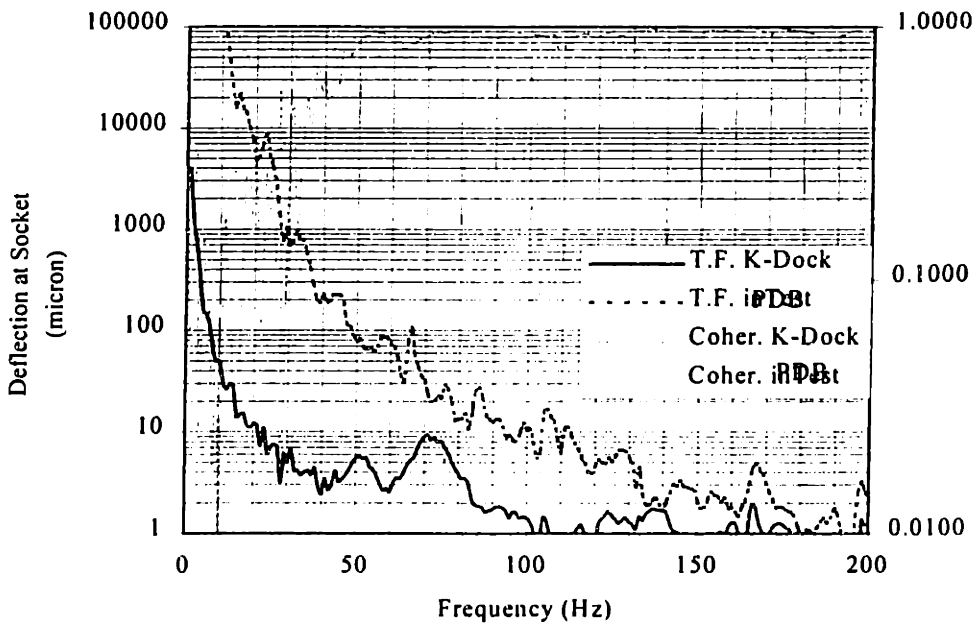


Figure 7.16: Vibration Transmission and Coherence for KDS and PDB Interfaces

7.3.2.6 Yield Data and Operator Feedback

Data on test cell throughput and was collected during the field trial by tracking the test times and yield data on two lots of wafers that passed through the test cell. Preliminary results show, for the probe application, the k-dock improved yield 1.1 % and tightened the yield distribution by 0.8 % over the j-ring. Approximately 20K die over 80 wafers were tested on each interface. The sample size of 80 wafers is not large enough to eliminate other variations such as front-end processes and operator influence. However, if further supporting data can be collected, the 1 % increase in yield is of tremendous significance. For a device with 90 percent yield, a 1 % increase represents a 10% false negative rate due to poor interface design. While high performance devices are often re-tested to eliminate these false negatives, this improvement would still yield a .01% yield increase. For a $130B industry, this improvement represents a potential $13M increase in yield.

User feedback was also collected via surveys and interviews. The experience and analytical skill level of the operators made it difficult to collect objective, quantitative data. Most of the feedback focused on the operator interface and the design of the interface support structures. This data was used to further improve the design before it was released to production.

7.3.3 Verification testing

A final set of tests were run after the final design was released to production to verify that the product met the initial performance specifications. Two separate sets of tests were performed on the KDS, the first repeated some of the field tests described above, but in a more controlled environment to eliminate the constraints imposed by testing on a production floor. The second set of tests measured the life-cycle performance of the docking system to verify that the module meets the design MTBF. The fixtures, procedures and results of these tests are described below.

7.3.3.1 Experimental Setup

The repeatability of any type of docking system will depend on the repeatability of the forces acting on it during and after docking. To insure that the forces in the test setup were known, a counterweight system was used to suspend the test head above the simulated

prober/handler. A schematic is shown in Figure 7.17. The weight of the test head was matched by the counterweight, so the only vertical forces on the test head were due to friction in the counterweight system. The pulley for the counterweight system was aligned such that it was directly above the CG of the test head when docked to eliminate lateral forces on the test head. Controlled external loads were applied to the test head with additional weights and pulleys.

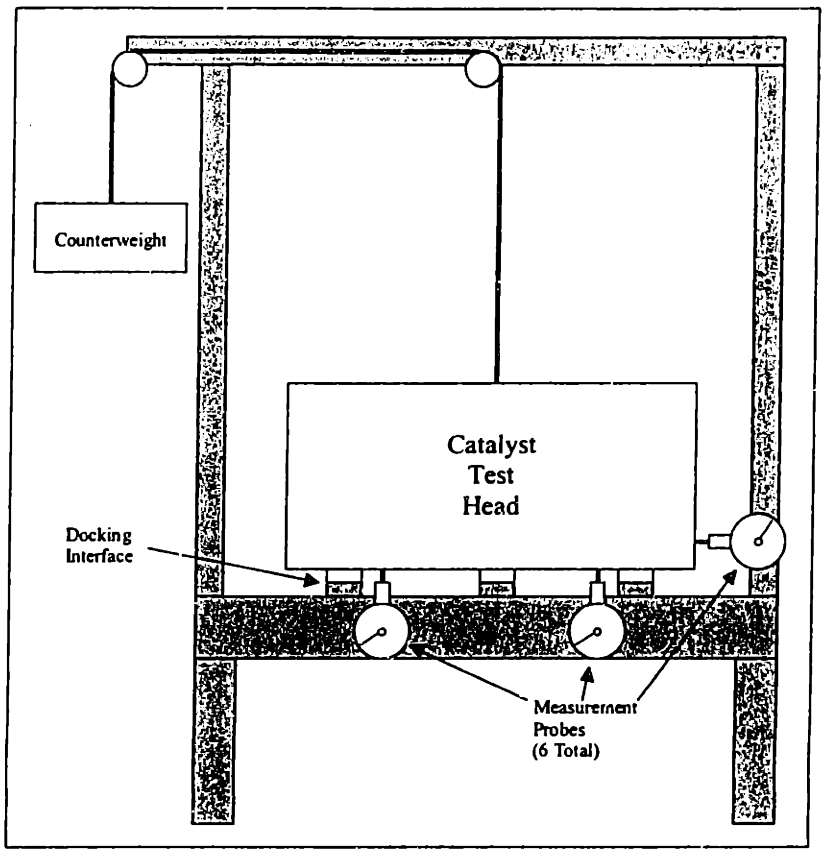


Figure 7.17: Measurement Test Setup

The position of the test head was measured after each dock using six sensors, one for each degree of freedom. Three probes were situated on the test head, extending downward. These measured the vertical position of the test head as well as its pitch and roll rotations. Three more probes with retractable probes were secured to the test frame to measure the position in both lateral directions as well as yaw rotation. All of the probes contacted hardened steel targets on the corresponding surface to insure repeatability of the surface. A picture of the measurement setup is shown in Figure 7.18.

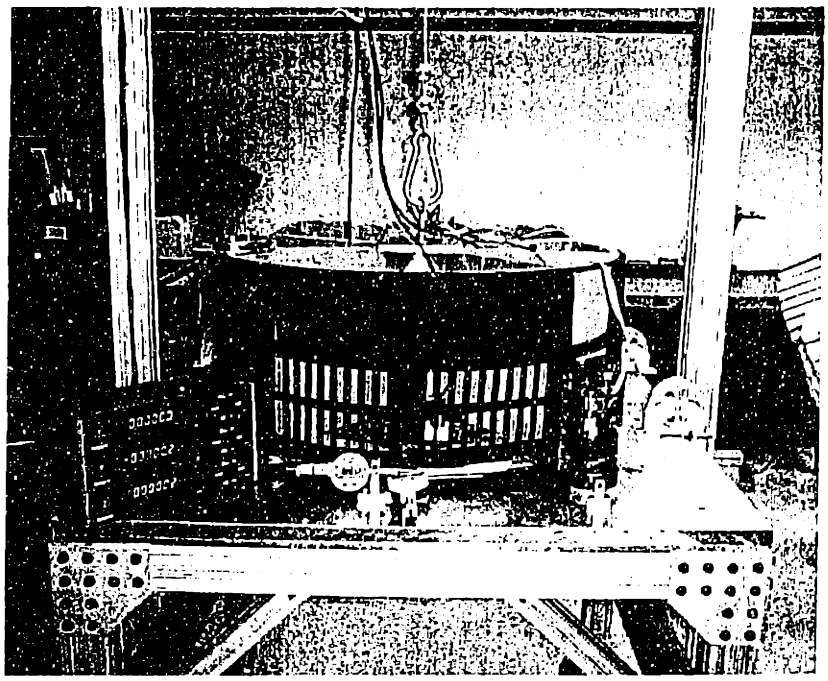


Figure 7.18: Measurement Equipment

The total non-repeatability of the test setup may be conservatively rated at ±2.5 microns (0.0001"). This allows for meaningful results in the range of the intended docking system non-repeatability of ± 25 microns (0.001"). Measurement errors are introduced by accuracy of the measurement probes, deflections of the metrology frame and thermal expansion errors.

Two different probes were used in this setup. Those that read the position of the test head in the vertical direction have a resolution of 0.5 micron (0.000020"). The probes used for horizontal measurements have a resolution of 1.28 micron (0.000050"). Errors due to loading manifest themselves as a change in sensor reading when there is no relative motion of the test head to the base. This may be due to flexing of the test base when the weight of the test head is supported, or when external loads are applied. The stiffness of the base structure strongly influences the level of loading errors. The base was designed with ribs welded to the underside and a stack of a ½" and a 1" plate for the top. The worst-case stiffness of this structure is estimated to be better than 85 kN/mm (500 lbs/mil), and thus a nominal 40 N (10 lb) load would result in a 0.5 micron (0.00002") deflection error. Aluminum was used for the construction material to match the thermal expansion rate of the test head. The metrology is therefore immune to systemic temperature changes, but may be affected by large gradients. Window shades and doors were closed to reduce solar gain and

convective airflow and the room temperature was characterized prior to, and monitored during testing. It is estimated that structural and thermal effects contributed no more than 1μm to the error budget.

Several sets of tests were performed on the KDS to verify that it performed to specification, including repeatability with and without external loads and static stiffness. Similar tests were performed on the PDB interface design.

The most basic test for non-repeatability of the docking system was made with no external loads applied to the test head. This was done to measure the inherent non-repeatability of the interface. A second, more realistic repeatability test is that which is performed with lateral loads on the testhead during docking. These loads may be due to the cable bundle, the manipulator, or another outside influence. Two general conditions of external loading were simulated in this investigation.

- To test forces of relatively large magnitude, a load of 350 N (75 lb.) was applied in three orthogonal directions, aligned with the natural axes of the test head. A load of this nature would most likely be present within a particular test head and prober/handler setup. In this situation, the test head is moved only slightly between undock/dock cycles, so it is likely that the force vector does not change significantly in direction. The loads may be of comparable magnitude due to the effects of manipulator and cable stiffness.
- The second loading condition examined the effects of a small magnitude, reversing load. The load was applied as a side to side force at the cable bundle exit. A load such as this may be encountered when changing the test head from one setup to another. Although the loads may be of equal magnitude to those in the previous case, for the purposes of this test the smaller magnitudes used were adequate. This test should measure the slop present in underconstrained, pin-bushing alignment mechanism.

To test interface static stiffness, an incremental load up to 350 N (75 lb.) was applied in a specific direction, and then reduced in the same incremental steps. The position of the test head was measured at each step in the force application and reduction. This allowed a measure of any hysteresis in the interface, or motion that is not undone when the load is removed. The result for static stiffness was obtained by finding the slope of a best fit linear approximation to the data.

7.3.3.2 Repeatability Test Results

For all the tests, a set of 30 dock/undock cycles were performed and a 3-sigma (3σ) criteria was used to analyze the data collected from the docking experiments, which means that 99.73% of all docks will fall in the range of $\pm 3\sigma$.

The results of the KDS repeatability are shown in Table 7.1 and results from the PDB system are shown in Table 7.2.

	X (in.)	Y (in.)	Z (in.)	θx (arc-sec)	θy (arc-sec)	θz (arc-sec)
No external loads	.0004	.0002	.0002	2	2	2
Side Cable Load	.0006	.0014	.0006	13	13	27
Longitudinal Cable Load	.0032	.0019	.0019	25	40	14
Vertical Cable Load	.0027	.0006	.0009	16	11	16
Reversing Side Cable Load	.0116	.0037	.0003	24	26	74

Table 7.1: Repeatability of *Kinematic Docking System*

	X (in.)	Y (in.)	Z (in.)	θx (arc-sec)	θy (arc-sec)	θz (arc-sec)
No external loads	.0004	.0006	.0000	2	1	4
Longitudinal Cable Load	.0004	.0002	.0001	5	4	4
Reversing Side Cable Load	.0029	.0033	.0001	3	3	49

Table 7.2: Repeatability of Pin-Bushing System

As may be seen, the performance of the KDS is significantly degraded by the presence of constant external loads whereas that of the pin-bushing system is not. This is due to the nature of the two interfaces. The PDB system actually performed better with a constant cable force, since it tends to seat the pins to one side of the bushings. While performance for a pin and bushing arrangement is improved when a preload force is consistently applied, the nature of the interface requires that there be some clearance, or slop. When a reversing load is encountered, it will tend to show the extent of the clearance in the interface. In addition, the performance of the pin-bushing interface is highly sensitive to its manufacture and installation, and thus test on additional systems could yield significantly different results.

With the reversing side cable load, the performance of the KDS was significantly reduced, particularly in the lateral (X-axis) direction. The performance of the pin-bushing

system was also reduced, although it did not show the same degree of change as did the KDS.

Several mechanisms were investigated to explain the reduction in KDS repeatability performance under applied loads, including insufficient preload and friction at the ball-groove interface. For a kinematic interface to work properly, all six ball-groove contact points be allowed to find their natural position without influence from the others. It is likely that the reversing lateral loads during the docking process, combined with the friction and large preload force, causes stiction at the ball groove interface, thus reducing the repeatability. Two additional sets of experiments were performed to further this hypothesis. First, the coupling preload of each ball-groove interface was increased approximately 80% through a simple software change. The increase in preload had a dramatic effect on the non-repeatability of the PD system. With the reversing side to side cable load, the lateral non-repeatability was reduced 90%. The y-axis cable load allowed for 30% reduction in lateral and 63% reduction in vertical non-repeatability.

The second set of experiments attempted to reduce the friction between the ball and groove materials. As an initial trial, the ball-groove interface was lubricated with oil to reduce the coefficient of friction. Due to promising results from this test, other methods of lowering the friction were investigated. Titanium nitride (TiN) coated tooling inserts were glued to the grooves at the ball contact points to serve as hard surfaces. The coefficient of friction between steel and TiN is approximately ½ that of steel on steel. In addition, kinematic coupling models suggest that higher groove hardness (compared to ball hardness) should improve repeatability. Lubrication of the ball-groove interface resulted in a reduction in lateral non-repeatability of 85% and vertical non-repeatability of 57% under a y-axis cable load. The TiN coated inserts resulted in a 78% reduction in lateral non-repeatability during the reversing side to side cable load.

A graph summarizing the performance of the standard and modified *Kinematic Docking System* as well as the results from the pin-bushing system, is shown in Figure 7.19.

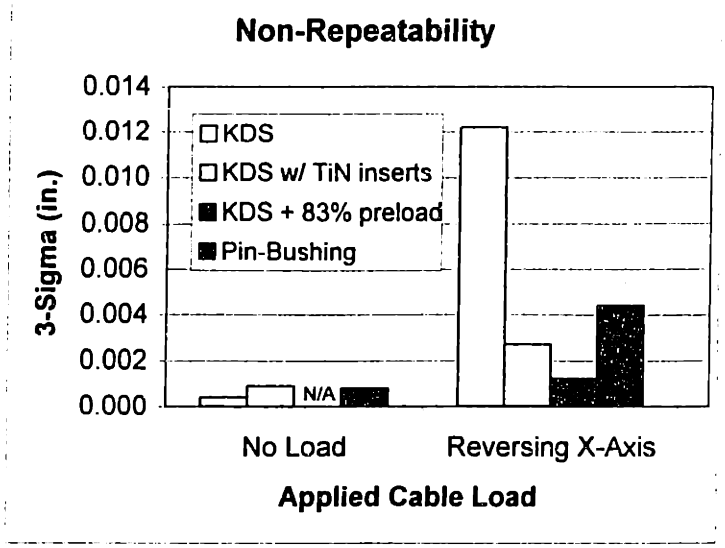


Figure 7.19: Results of Preload and Friction Experiments

7.3.3.3 Stiffness Test Results

The results of the KDS stiffness tests are shown in

Table 7.3, and those of the PDB system are in

Table 7.4. The tables provide not only the direct stiffness of an axis, but the cross coupling between axes as well. As may be seen, all stiffness results meet the requirements discussed in Chapter 6. It is interesting to note that the levels of hysteresis seen in the two systems were noticeably different. The KDS interface displayed essentially no hysteresis while the pin-bushing interface showed approximately 0.002". This indicates the large amount of slop contained within the inTEST interface.

Force applied	Axis of Stiffness					
	X (lb./in.)	Y (lb./in.)	Z (lb./in.)	θx (lb./arc-sec)	θy (lb./arc-sec)	θz (lb./arc-sec)
X	1.00×10^5	4.87×10^5	4.25×10^5	42.6	24.8	43.9
Y	1.00×10^5	2.98×10^5	1.60×10^5	7.5	14.6	8.2
Z	∞	6.69×10^5	2.61×10^5	15.3	26.1	69.9
θx	7.79×10^5	∞	1.02×10^6	24.3	123.5	204.1
θy	3.11×10^5	8.98×10^5	5.68×10^5	89.3	20	15.9
θz	1.35×10^5	3.36×10^5	2.49×10^5	20	208.3	17.7

Table 7.3: Stiffness Results for *Kinematic Docking System*

Force applied	Axis of Stiffness					
	X (lb./in.)	Y (lb./in.)	Z (lb./in.)	θx (lb./arc-sec)	θy (lb./arc-sec)	θz (lb./arc-sec)
θz	8.90×10^{4}	7.18×10^{4}	9.73×10^{3}	27.5	21.9	1.3

Table 7.4: Stiffness Results for Existing Docking System

If the modifications to address the poor repeatability under load performance are adopted, the performance of the KDS meets the specifications prescribed in Chapter 6. In addition to this, these test show that the KDS does produce superior stiffness and repeatability to that of a standard pin and bushing interface.

7.3.4 Life Cycle Tests

Semiconductor test equipment typically has a useful life of 5-10 years before it is obsolete and therefore, an understanding of the KDS maintenance requirements and failure modes is important if is to perform reliably. The MTBF goal for the *Kinematic Docking System* is 10 years, which represents approximately 20K cycles if cycled 5 times per day. This is a very conservative specification for two reasons, first, MTBF for ATE is historically in the 2 to 6 month range and second, it is highly unlikely that a system will be cycled regularly 5 times a day. In addition to this, the fixture built to perform the tests created a 'worst case' load scenario during each docking cycle.

A test system, shown in Figure 7.20, was built to automatically cycle a sample module, apply the design loads and record the performance data. This system contained the following subsystems:

- Pneumatic air cylinder, valves and regulators: These are used to simulate the docking motion of the head and is used to generate a position-independent load on the module during docking and undocking.
- Load cell and meter: These are used to measure the docking forces generated by the module, as well as the loads applied by the air cylinder.
- Module controller: This is a simple PCB that contains the same basic electronics used in the final KDS system controller. It has been designed to run a single module and accept inputs from an external controller.
- System controller: A PC running *LabView* was used to coordinate the motions of the subsystems and collect and record the data.

Figure 7.20: Module Life-Cycle Test Fixture

Figure 7.21 is a graph of the life-cycle preload performance, which shows that the module maintained a preload force of approximately 300 lbs. +/- 25 lbs. through the first 10K cycles. The line at approximately 2700 cycles represents a sensor error and the dips at 5K and 8.5K are indications of lubrication failure.

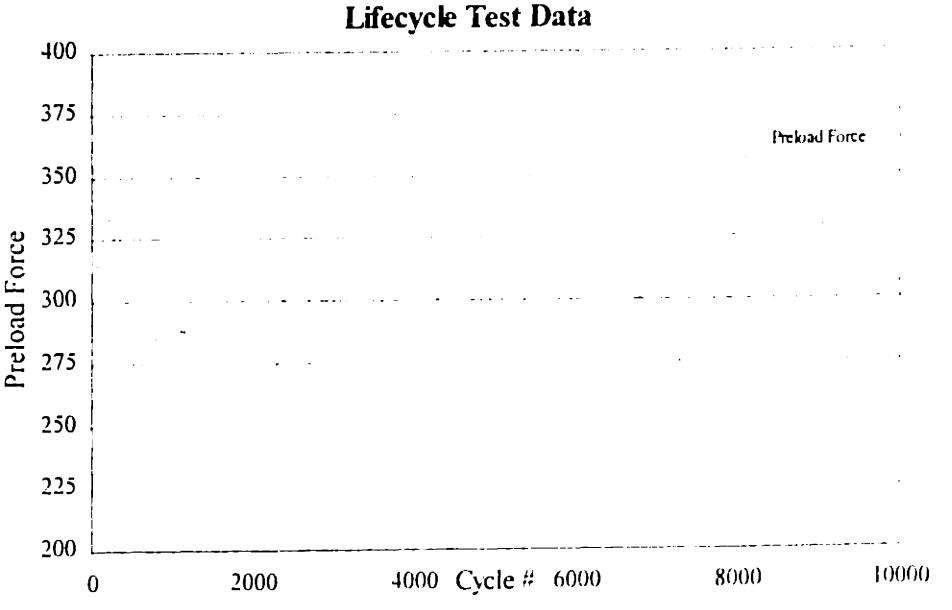


Figure 7.21: Life Cycle Preload Results

The sensor failures were due to oxidation and dust buildup on a beryllium-copper electrical contact. This was addressed by gold plating the contact surface and by changing the contact geometry to add a slight wiping action during contact. Lubrication failure occurred at the lead-screw-nut interface. This is likely due to the high sliding static loads applied during docking. The friction causing the lubrication problem is the same that prevents the screw from back-driving once it is docked, so there is little that can be done to eliminate the friction issue. Grease fittings were added and a maintenance schedule recommended to address the lubrication problems. After 20K cycles, the module was disassembled and the drive train components were inspected for wear. None of these components was outside of nominal specifications.

7.3.5 Future work

A number of derivative design projects that enhance the applicability and performance of the KDS have been pursued in parallel with the work described above. Two of these, which are in the proof-of-concept phase, are described below.

7.3.5.1 Auto-Planarization

The error budget discussed in Chapter 6 shows that the test head docking stability is the primary source of interface errors. The *Kinematic Docking System* addresses this primary error source, but does not solve all interfacing issues. Error motions and misalignment in the other components of the structural loop can still cause interfacing problems. This is more important in probing, than handling applications. Handlers usually use additional alignment features local to the device socket to insure proper placement. Probers do have optical systems that can accurately measure misalignment, but the motion of the probe stage is limited to four axes, X, Y, Z and yaw. Therefore, despite the fact that a prober can measure pitch and roll error of a probe card, it is not capable of compensating for these errors.

Currently, these errors are eliminated using height adjustment screws in the probe interface and a dial indicator to check the runout between it and the chuck. This yields sufficient accuracy as long as the interface is adjusted properly periodically and if all probe cards are built with similar planarity. However, these requirements are often difficult to meet in a manufacturing environment and it is difficult to get repeatable results from probe cards that are essentially fiberglass. These problems increase with increasing die size and with the

use of multi-site probe cards. The result is an increasing need to planarize interfaces and probe cards to insure that proper probe-to-pad contact is achieved.

One method of addressing this concern is to add a controlled axis of motion to the existing docking module. If each module is capable of adjusting the height of the height of the ball relative to the head, then the pitch and roll of the probe card can be controlled without touching the interface. The current modules already have z-stage bearings and a lead screw for adjustment, adding a drive train, feedback and a controller is all that is required to achieve automatic test head planarization. Prior to the mechanical design of this motion system, an error budget was developed to quantify the benefits and determine if the supporting structural errors exceed the resolution of such a system.

The error budget was developed by modeling the relationships between the structural errors and probe card planarity. From this model, the required z-positional resolution can be calculated and feasibility can be assessed. This study was made to determine out how much deflection there would be at the mounting points of the kinematic module grooves due to forces related to docking the head. These deflection were then used to determine the planarity of the probe card with respect to prober chuck. If the slope of the probe card with respect to the chuck is beyond a specified range then the testhead would need to be planarized. The spreadsheet above was developed for this study.

The model used in to determine the amount of deflection at the groove mounting locations uses simple beam bending theory. Forces applied to the groove, which is mounted to the prober head plate, bend the plate about it's support point on the prober. For simplicity, the prober structure is assumed to be infinitely stiff, which is reasonable when compared to the head plate. Given a force, the amount of deflection at the grooves is calculated and once the deflections at each groove are known a transformation matrix is used to calculate the probe card planarity. Several sample cases were run with this analysis tool using different structural scenarios. These tests showed that, for a properly designed interface, there is sufficient structural stiffness to support autoplanarization with sufficient resolution to affect probe card planarity.

To simplify development, manufacturing and field upgrade, the autoplanarization system needed to be designed in such a way that it used as many of the existing parts and as few modifications to the existing design as possible. This was achieved by replacing the manual

Z-adjust components with a new servo-driven version. All other components remained unchanged.

This servo drive mechanism, shown in Figure 7.22, contains the following components:

- DC gear motor similar to preload motor
- Sprockets and chain drive
- Leadscrew and bronze nut
- Linear encoder assembly

Belt tension is adjusted via a setscrew during assembly and the leadscrew is preloaded using a set of bellville washers. The entire assembly is bolt compatible with the existing manual Z-adjust assembly and upgrades can be done in the field. The assembled system is shown in Figure 7.23, alongside a standard module.

The more difficult subsystem of this design is the control system. This controller needs to be able to control six separate motors, three for preload and 3 for autoplanarization, as well as communicate with the prober for probe card planarity data. This system has not yet been designed, but will not require a 6-axis coordinated controller, because it is possible to interleave the motions of each individual motor. Therefore, a simple, low cost PLC-based controller is probably sufficient. This has been accomplished already in the initial prototype development. Additional information on the autoplanarization system can be found in Appendix 9.2.

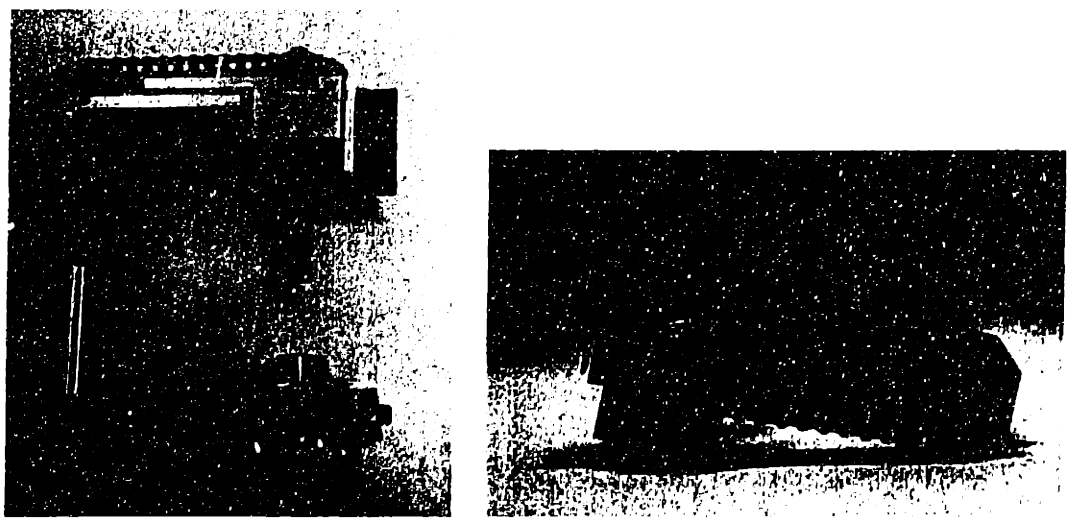


Figure 7.22: Autoplanarization Upgrade Assembly

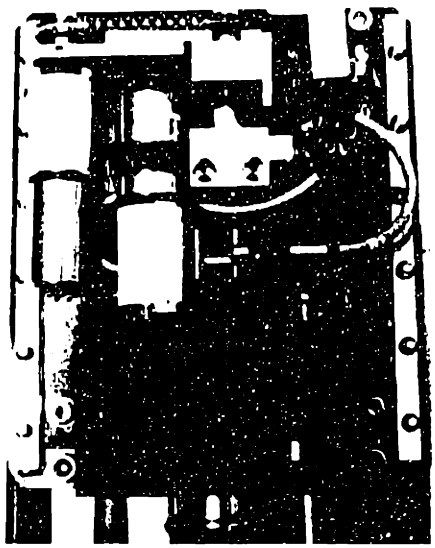


Figure 7.23: Docking Module with Autoplanarization Option

7.3.5.2 Mini-Dock

The design specification for the module described above was developed to meet the needs of several new testhead interface designs, which left the module packaging relatively unconstrained. This allowed the incorporation of the many options that make the module flexible and easy to use, as well as the ability to plan for upgraded functionality, such as autoplanarization described above. One drawback of this approach is that the final design was difficult to apply to existing product, as can be seen by the additional structure required to dock the AMS head in Figure 7.6 and Figure 7.7. While the KDS was not designed to work with older systems, there are many reasons why this would be advantageous. The obvious benefit of improved docking and throughput are generally not sufficient to convince production floors to retrofit working equipment. However, incompatibility between new and old interfaces can cause enough logistical problems to justify retrofit expense. The introduction and acceptance of a new test head design, such as the ones using the kinematic modules, does not happen immediately and during this period of transition, the use of older generation testheads will continue alongside the new. These two disparate test systems will

need to share common device handling equipment and it is desirable to do so with as little hardware setup as possible.

Full compatibility would require that both the old and new test heads use the same X-Y-Z groove mounting location on the test head and handling device. Unfortunately, this was not possible because the existing square test head did not provide similar mounting locations as those of newer, round test heads and the thickness of existing interfaces was too small to accept the tall module package. Rather than modify either of the test heads or the docking system, a new 'mini-module' was designed that is shorter and can be applied in tighter spaces. Compatibility with the interface grooves can be maintained by borrowing a few key concepts from the kinematic interface. The mini-module is based upon the latch pin and ball-groove mechanism of the *Kinematic Docking System.* Instead of utilizing a DC motor, sensors and support electronics, this design uses a single cable drive with to provide the preload and release forces to all three modules simultaneously. Input to the cable drive can be operator activated with a lever, or a single DC motor can be used to drive all three. Because the modules are all driven simultaneously, all of the control electronics can be eliminated and therefore, the mini-module interface is considerably less expensive than the full module. This module design is shown in Figure 7.24 below.

Figure 7.24: Mini-Module Design

7.4 Summary of Chapter 7

This chapter discussed the several stages of testing performed on the *Kinematic Docking System*. These tests were used to verify the design concept and implementation, as well as to quantify the performance and expected life of the system. The results of the pre-production tests confirmed that the design yields a substantial improvement in stiffness and repeatability over existing interfaces. Tests on the final system show that production interfaces based on the KDS meet the stiffness and repeatability performance goals set for the design. However, repeatability under load performance could be improved by implementing one of several small design modifications. In addition to the test procedures and results, the initial designs of two derivative products were discussed that can be used to expand the performance and applicability of this interface product.

8 Summary and Conclusions

8.1 Introduction

This chapter briefly outlines the core concepts of ATE mechanical architecture already presented in this thesis and reviews the fundamental contributions of this work. Additionally, the chapter draws some general conclusions on the current state and future needs of ATE mechanics. Finally, the thesis concludes with a discussion of future areas of focus that may have significant value to the development of ATE and the electronics and semiconductor industries as a whole.

8.2 Summary of Mechanical Systems Design for the Semiconductor ATE Industry

A number of trends in the packaging and performance of semiconductor devices will impact the requirements and design of future automatic test equipment. The exponential growth in complexity suggested by Moore's Law is realized through increasing the number of device pins, clock speeds and power consumption. Testing of these devices in the future will require tighter integration in the design of automatic test equipment. In particular, the area of test cell mechanics, including fixturing, positioning, thermal and signal transmission systems will be critical to the success of future designs. Therefore, it is important to understand how device attributes and trends drive the mechanical requirements of test cells. If these relationships are understood, insight into the challenges and roadblocks of future test cells can be developed.

This thesis developed three core concepts:

First, a general roadmap of the semiconductor industry requirements as defined by the Semiconductor Industry Association was presented. This roadmap predicts the future trends, goals and challenges from an electrical performance perspective. These device trends impose requirements on the design of test equipment and therefore a similar roadmap outlining the mechanical requirements of ATE can be developed. Critical parameters such as thermal performance, mechanical accuracy and mechanical systems cost must be tracked to insure that the goals of the SIA device roadmap can be met. These mechanical roadmaps can be developed by understanding and modeling the relationships between the device parameters and mechanical requirements.

The second concept presented was the critical issue of signal transmission fixturing as well as the physical implementation of these systems in test equipment. The challenges of high frequency signal transmission grow with increasing clock frequency and device pincount, and this issue has the potential to limit the overall performance of ATE. A large body of theoretical and empirical work exists that describes the analysis of these systems in electrical terms and this work remains primarily in the electrical domain. However, the physical implementation of these signal transmission systems, and much of their performance, is determined by their mechanical implementation and few tools exist that can be used by the mechanical systems designer to understand and optimize the design of these components. A framework for analyzing this coupling between mechanical design and electrical performance is presented as well as models of a typical signal transmission system.

The third concept discussed is the design of test fixturing and interfacing systems used in automatic test equipment. These interfaces are a critical component the signal transmission systems discussed above and their design can have a significant impact of the overall performance and throughput of the test equipment. The design requirements and shortcomings of existing systems are presented. This is followed by the description of a new interfacing system based on the kinematic coupling, which offers greater stability and repeatability. Test data from prototypes, design verification efforts and field tests on this new system is also presented.

These three concepts form a coherent body of work that is relevant to the design of ATE mechanics for current and future devices. This work is by no means a completely comprehensive analysis of ATE mechanics, or a completed roadmap of the future of test cells. Rather it is intended to be a starting point in the development of a new class of test equipment that better addresses the growing mechanical test requirements.

8.3 Summary of Contributions

All of the work presented in this thesis is related to the three concepts presented above. The fundamental effort described in this thesis relate to the design, analysis and road-mapping of the mechanics required to test current and future semiconductor devices. Specifically, these contributions are as follows:

1. The translation of the SIA data from a device-based to a mechanics-based roadmap offers a needed perspective on the growing requirements of ATE mechanical systems. This

analysis includes a prediction of the thermal, mechanical accuracy and mechanical systems cost requirements of future systems. These mechanical roadmaps are valuable in that they give a new perspective on the future challenges facing the ATE systems architect. These models can also be used as tools to optimize the design of current tester architectures by providing insight into the mechanical systems principles that affect tester performance.

2. The development of detailed signal transmission models that relate the electrical performance to the mechanical implementation is a step forward in the effort to design higher performance test systems. These models lends insight into a highly coupled problem that historically has not been modeled in detail or accounted for by both the electrical and mechanical design communities. The model can be used to:

- Gain intuitive insight into the relationships that govern transmission lines
- Optimize the design of existing or proposed signal transmission systems
- Investigate the sensitivity of mechanical tolerances such as geometry, accuracy and temperature, on the electrical performance of the transmission line
- Predict the limitations of existing technology

All of these tasks are critical to the development of test equipment that can meet the needs described by the test cell mechanics and SIA roadmaps.

3. Finally, to help address the mechanical aspects of these problems, a novel interface product, the *Kinematic Docking System* (KDS), was developed from concept to product introduction. Although this interface was developed using the deterministic principles and tools that are common to precision machine tools, it represents a significant step forward in the design and performance of test cell interfaces. This design was taken from concept to product development to release and shipment as part of this thesis effort. Field tests confirm that this system has superior repeatability and stability over existing state of the art interfaces. This system is currently being used on hundreds of test systems throughout the world.

All of these contributions are significant for several reasons. First, the semiconductor industry plays a significant role in the economy and in our everyday lives and the performance of automatic test equipment plays an integral role in the success of this industry. The models and roadmaps described in this thesis will help semiconductor ATE manufacturers meet the growing demands of the industry. Second, the proliferation of high-speed digital devices and microwave communication products is pushing the limits of

existing signal transmission line structures. The analyses developed in this thesis represent a first step in understanding the limitations of the current technology and provide a structure in which new technologies can be evaluated. Finally, the *Kinematic Docking System* solves a relevant and important problem that faces the ATE industry by using sound engineering principles and practices from the precision machine design industry.

8.4 Future Areas of Research in the Design of Test Cell Mechanics

This work is only an initial step in the effort to advance the development of semiconductor test mechanics. In addition to this, much of the effort described in this thesis apply to areas beyond ATE and beyond the semiconductor industry. Much more work needs to be done in each of the three areas discussed in this thesis.

For example, there are several changes and enhancements that can be effected in the design of the *Kinematic Docking System* that will improve its performance, applicability and ease of use. In addition to this, much of the background analysis performed in the development of this system can be applied to the design of other test cell components.

The signal transmission models presented are of a basic nature and should be developed in greater detail. Many changes can be made to improve the accuracy, such as the inclusion of higher-order phenomenon and phase information. However, the real utility in these tools is in their ability to give the designer insight into the coupling between seemingly unrelated mechanical and electrical parameters. Improvements should be made in the ease of use of this tool as well as it's ability to symbolically evaluate and optimize the performance of signal transmission systems. Ultimately, tools that integrate electrical performance and mechanical implementation need to be included in the mechanical and electrical CAD layout tools used by designers and engineers. One can envision a *Pro/Engineer* module that integrates netlist data from schematic capture tools with solid models of interface mechanics and can use this data to predict and optimize transmission line bandwidth performance.

Perhaps the greatest effort should be focused on maintaining and advancing the mechanical roadmap analyses initiated in this work. Hundreds of engineers, designers, technologists and visionaries spend countless hours on the biannual SIA roadmap of device requirements. This effort yields a return of continuous, stable growth of semiconductor technology. Similar efforts, on somewhat less of a grand scale, are required to insure that

ATE mechanical systems designs keep pace with these devices and that test cell architecture progresses in a continuous, stable manner.

ATE is only one segment of the larger semiconductor industry, and although semiconductor sales have surpassed $200 billion, the industry is still relatively small and immature relative to more traditional heavy industries. Although Gordon Moore's observation of exponential growth is predicted to continue well in to the next century, the semiconductor revolution, like all of it's predecessors, must eventually slow down and mature. The greatest challenge and real work for the future is to help guide semiconductor equipment design through the transformation from instrumentation to that of industrial process equipment.

9 Appendices

9.1 Signal Transmission Model

This appendix contains a copy of the baseline model, equations and results, including all of the *Mathematica* code used. This is followed by the equations for calculating the nominal impedance of a coplanar waveguide as described in Section 4.3.2.

Digital Transmission System Model

This model simulates the transmission path of a typical Digital probe application

Setup Environment

Packages

```
<< EE`
<< Graphics`
```

Formatting

```
$TextStyle = {FontFamily -> "Times", FontSize -> 10};
```

```
red = {RGBColor[1, 0, 0]};
green = {RGBColor[0, 1, 0]};
blue = {RGBColor[0, 0, 1]};
```

Load Model Parameters

Constants

μo, ηo, ϵo
ϵteflon, ϵexpteflon, ϵair, ϵFR4
σcopper, ρteflon, ρFR4

- Speed of light [meter/sec]

```
light = 3 10^8;
```

- Permeability, $\mu=\mu_0$ [Henry/meter]

```
μo = 1.257 10^-6;
```

- **Permittivity (dielectric constant), ϵ [Farad/meter]**

```
εo = 8.854 10^-12;
εteflon = 2.1;
εexpteflon = 1.5;
εair = 1;
εFR4 = 4.5;
```

- **Characteristic Impedence (of free space)**

Wadell, pg 16

```
ηo = Sqrt[μo / εo];
```

- **Conductivity, σ [Siemens/meter] -or- [1/ohm-meter]**

Source: Seeger

```
σcopper = 5.183 10^7;
σsilver = 6.173 10^7;
σgold = 4.098 10^7;
```

- **Resistance, ρ [ohm-meter]**

Source: Seeger

```
ρteflon = 10^15;
```

This is an estimate from the resistance of bakelite (Seeger)

```
ρFR4 = 10^9;
```

- **Component Parameter Estimates**

```
padcap = .4 10^-12;      (* 0.4 pF *)
```

This can range from 0.1pF for a well designed (relieved ground) pad to 2pF for a moderately designed pad to 4 pF for an unrelieved, poorly designed pad+via. (Manchester, Corwith)

```
viacap = .2 10^-12;     (* 0.2 pF *)
```

This can range from 0.05 pF to 1 pF (Manchester, Corwith)

```
stripthick = 0.0000438;    (* 0.0017 " *)
```

This is the thickness of a common, high performance microstrip (surface). This is approximately 1/2 oz. copper with 1 mil. of plate. (Manchester)

```
pogocr = .05;
```

This is an estimate of the contact resistance of a pogo pin hitting a pad. (Corwith)

- **Other useful stuff**

This is the ratio of outer to inner diameter for ideal power transfer in a coax line (Wadell)

```
DiaRatio = 1.65;
```

Channel Card parameters

Channel Card Pads (ccpad)

```
ccpadcap = padcap;
ccpadnum = 2;
ccpadtol = 1;
```

Channel Card PCB Traces (ccmstrip)

```
ccmstripdout = .00018;          (* .072 " *)
ccmstripdin = .000254;          (* .010 " *)
ccmstripdin2 = stripthick;
ccmstriplen = .025;             (* 1.00 " *)
ccmstripdiel = εFR4;            (* 4.5 *)
ccmstripcond = σcopper;
ccmstriptol = 1;
```

Channel Card Coax Cable RG-178/U (cccoax)

```
cccoaxdout = .00137;            (* .054 " *)
cccoaxdin = .00051;             (* .020 " *)
cccoaxlen = .1;                 (* 4.00 " *)
cccoaxdiel = 1.406;
cccoaxcond = σsilver;
cccoaxtol = 1;
```

■ Channel Card Pogo Pin, Coax Barrel (ccpogo)

```
ccpogodout = .00216;          (* .085 " *)
ccpogodin = .00061;           (* .024 " *)
ccpogolen = .035;             (* 1.40 " *)
ccpogodiel = 2.3;
ccpogocond = σgold;
```

■ Channel Card Pogo Pin, Exposed Needle (ccpin)

```
ccpincr = .1;
ccpindout = .00254;           (* 0.10 " *)
ccpindin = .00046;            (* .018 " *)
ccpinlen = .00178;            (* .070 " *)
ccpindiel = 1;
ccpincond = σgold;
```

DIB parameters

■ DIB Pogo Pad (dibpad)

```
dibpadcap = padcap;
dibpadnum = 2;
dibpadtol = 1;
```

■ DIB PCB Traces (dibmstrip)

```
dibmstripdout = .00018;       (* .072: *)
dibmstripdin = .000254;       (* .010 " *)
dibmstripdin2 = stripthick;
dibmstriplen = .075;          (* 3.00 " *)
dibmstripdiel = eFR4;         (* 4.5 *)
dibmstripcond = σcopper;
dibmstriptol = 1;
```

■ DIB Vias (dibvia)

```
dibviacap = viacap;
dibvianum = 2;
```

Probe Pogo Tower parameters

- Pogo Tower Pogo Pin, Exposed Needle (ptpin)

```
ptpincr = .1;
ptpindout = .00254;          (* 0.10 " *)
ptpindin = .0013;            (* .051 " *)
ptpinlen = .0025;            (* .010 " *)
ptpindiel = 1;
ptpintol = 1;
```

- Pogo Tower Pogo Pin, Coax Barrel (ptpogo)

```
ptpogodout = .0030;          (* .118 " *)
ptpogodin = .0012;           (* .047 " *)
ptpogolen = .04;             (* 1.6 " *)
ptpogodiel = 1.14;
ptpogocond = σgold;
ptpogotol = 1;
```

Probe Card parameters

- Probe Card Pogo Pads (pcpad)

```
pcpadcap = padcap;
pcpadcr = .01;
pcpadnum = 1;
pcpadtol = 1;
```

- Probe Card PCB Traces (pcmstrip)

```
pcmstripdout = .00018;       (* .072 " *)
pcmstripdin = .000254;       (* .010 " *)
pcmstripdin2 = stripthick;
pcmstriplen = .050;          (* 2.0 " *)
pcmstripdiel = εFR4;         (* 4.5 *)
pcmstripcond = σcopper;
pcmstriptol = 1;
```

■ **Probe Card Probe Needle (pcneedle)**

```
pcneedledin = .000125;          (* .005 " *)
pcneedlelen = .025;             (* 1.00 " *)
pcneedlecr = .01
pcneedlecond = σcopper;
pcneedletol = 1;

0.01
```

Terminating Impedence

```
zterm = 50;
```

Load Generic Models

Transformations

■ **Transforms**

TransSABCD[s_,zo_]
TransABCDS[a_,zo_]

S matrix to ABCD matrix (Medley's transform is incorrect, modified to agree with Pozar)

```
TransSABCD[s_, zo_] :=
 {{{((1 + s[[1, 1, 1]]) (1 - s[[1, 2, 2]]) +
       s[[1, 1, 2]] s[[1, 2, 1]]) / (2 s[[1, 2, 1]]),
    zo ((1 + s[[1, 1, 1]]) (1 + s[[1, 2, 2]]) -
        s[[1, 1, 2]] s[[1, 2, 1]]) /
      (2 s[[1, 1, 2]])},
   {((1 - s[[1, 1, 1]]) (1 - s[[1, 2, 2]]) -
       s[[1, 1, 2]] s[[1, 2, 1]]) /
     (2 s[[1, 2, 1]] zo),
    ((1 - s[[1, 1, 1]]) (1 + s[[1, 2, 2]]) +
       s[[1, 1, 2]] s[[1, 2, 1]]) /
     (2 s[[1, 1, 2]])}},
  s[[2]], s[[3]]}
```

ABCD matrix to S matrix (Medley's transform is incorrect, modified to agree with Pozar)

```
TransABCDS[a_, zo_] := {{{(a[[1, 1, 1]] +
        a[[1, 1, 2]] / zo - a[[1, 2, 1]] zo - a[[1, 2, 2]]) /
      (a[[1, 1, 1]] + a[[1, 1, 2]] / zo +
        a[[1, 2, 1]] zo + a[[1, 2, 2]]),
    2 (a[[1, 1, 1]] a[[1, 2, 2]] -
        a[[1, 2, 1]] a[[1, 1, 2]]) /
      (a[[1, 1, 1]] + a[[1, 1, 2]] / zo +
        a[[1, 2, 1]] zo + a[[1, 2, 2]])},
   {2 / (a[[1, 1, 1]] +
        a[[1, 1, 2]] / zo + a[[1, 2, 1]] zo + a[[1, 2, 2]]),
    (-a[[1, 1, 1]] + a[[1, 1, 2]] / zo -
        a[[1, 2, 1]] zo + a[[1, 2, 2]]) / (a[[1, 1, 1]] +
        a[[1, 1, 2]] / zo + a[[1, 2, 1]] zo + a[[1, 2, 2]])}},
  a[[2]], a[[3]]}
```

Cascades two data vectors (Matrix multiplies ABCD portion and adds scalars)

```
Cascade[a1_, a2_] :=
 {a1[[1]] . a2[[1]], a1[[2]] + a2[[2]], a1[[3]] + a2[[3]]}
```

■ **Some useful functions**

Hertz[wr]

Radss[wh]

Convert radians/sec to Hertz

```
Hertz[wr_] := wr / (2 Pi)
```

Convert Hertz to radians/sec

```
Radss[wh_] := wh 2 Pi
```

Reflection coefficient, Rho

```
Rho[zline_, zterm_] := (zterm - zline) / (zterm + zline)
```

```
freq = Radss[10 ^ fmag];
```

S-model Generators

■ **Series Impedence**

SeriesABCD[zs]

SeriesS[zs, zo]

These are the models derived from Max Medley 's book

Inputs:

zs = impedence to be modelled [ohms]
zt = terminating impedence [ohms]

ABCD matrix

```
SeriesABCD[zs_] := {{1, zs}, {0, 1}};
```

S matrix (Medley's matrix is incorrect, modified to agree with Besser and Pozar)

```
SeriesS[zs_, zo_] := {{zs / (zs + 2 zo), 2 zo / (zs + 2 zo)},
  {2 zo / (zs + 2 zo), zs / (zs + 2 zo)}}
```

- **Shunt Impedence**
  ShuntABCD[zs]
  ShuntS[zs, zo]
  ShuntSreal[zs, zo]

These are the models derived from Max Medley 's book

Inputs:

zs = impedence to be modelled [ohms]
zo = terminating impedence [ohms]

ABCD matrix

```
ShuntABCD[zs_] := {{1, 0}, {1 / zs, 1}}
```

S matrix

```
ShuntS[zs_, zo_] :=
 {{-zo (1 / zs) / (2 + zo / zs), 2 / (2 + zo / zs)},
  {2 / (2 + zo / zs), -zo (1 / zs) / (2 + zo / zs)}}
```

This eliminates the (-) sign on S11/S22 so that the answer is non-negative

```
ShuntSreal[zs_, zo_] :=
 {{zo (1 / zs) / (2 + zo / zs), 2 / (2 + zo / zs)},
  {2 / (2 + zo / zs), zo (1 / zs) / (2 + zo / zs)}}
```

- **Transmission Line (Simple)**
  SimpleLineABCD[zc]
  SimpleLineS[zc,zo]

- **Transmission Line (General)**
  GenLineABCD[zc, gl]
  GenLineS[zc, gl, zo]

These are the models derived from Max Medley 's book

Inputs:
gl (gamma*length, γl) = Electrical length: Propagation constant (complex) * length
zo = terminating impedence [ohms]
rho = reflection coefficient [unitless]

ABCD matrix

```
GenLineABCD[zc_, gl_] :=
 {{Cosh[gl], zc Sinh[gl]}, {(1 / zc) Sinh[gl], Cosh[gl]}}
```

S matrix

```
GenLineS[zc_, gl_, zo_] :=
 {{((zc - zo) / (zc + zo)) (1 - Exp[-2 gl]) /
     (1 - ((zc - zo) / (zc + zo)) ^ 2 Exp[-2 gl]),
   (1 - ((zc - zo) / (zc + zo)) ^ 2)
    Exp[-gl] / (1 - ((zc - zo) / (zc + zo)) ^ 2 Exp[-2 gl])},
  {(1 - ((zc - zo) / (zc + zo)) ^ 2)
    Exp[-gl] / (1 - ((zc - zo) / (zc + zo)) ^ 2 Exp[-2 gl]),
   ((zc - zo) / (zc + zo)) (1 - Exp[-2 gl]) /
     (1 - ((zc - zo) / (zc + zo)) ^ 2 Exp[-2 gl])}}
```

■ **Transmission Line (Lossless) Note: The 2 models below do not pass the transformation test!!**
LossLineABCD[zc_,bl_]
LossLineS[rho_,bl_]

These are the models derived from Max Medley 's book

Inputs:
bl (Beta*length, βl) = Electrical length: phase constant (imaginary portion of gamma) * lenght [radians]
rho = reflection coefficient [unitless]

ABCD matrix

```
LossLineABCD[zc_, bl_] :=
 {{Cos[bl], I zc Sin[bl]}, {I Sin[bl] / zc, Cos[bl]}}
```

S matrix

```
rho = (zc - zo) / (zc + zo);
```

```
LossLineS[rho_, bl_] :=
 {{rho (1 - Exp[-2 I bl]) / (1 - rho ^ 2 Exp[-2 I bl]),
    (1 - rho ^ 2) Exp[-2 I bl] / (1 - rho ^ 2 Exp[-2 I bl])},
  {(1 - rho ^ 2) Exp[-2 I bl] / (1 - rho ^ 2 Exp[-2 I bl]),
   rho (1 - Exp[-2 I bl]) / (1 - rho ^ 2 Exp[-2 I bl])}}
```

Physical Models

Coax line model

Coax[dout, din, len, diel, cond, wr, zterm]

Equations from Seeger, chapter 2

This is often useful

```
zcoax[dout_, din_, diel_] :=
 Sqrt[Lcoax[dout, din] / Ccoax[dout, din, diel]]
```

Calculate coaxial series inductance [Henry/meter]

$$\text{Lcoax[dout\_, din\_]} := \frac{\mu o}{2\,\pi}\,\text{Log[dout / din]}$$

Calculate coaxial shunt capacitance [Farad/meter]

$$\text{Ccoax[dout\_, din\_, diel\_]} := \frac{2\,\pi\,\epsilon o\,\text{diel}}{\text{Log[dout / din]}}$$

Calculate series resistance, including skin-effect [Ohms/meter]

$$\text{Rscoax[dout\_, din\_, cond\_, wr\_]} := \sqrt{\frac{\pi\,\text{wr}}{\text{cond}}}\;2\,\pi\,(\text{dout} + \text{din})$$

Calculate the dielectric loss (This is a first order curve-fit of the cutoff frequency. See Wadell, pg 49 and notebook pg 108)

```
Losscoax[dout_, din_, diel_, freq_] :=
 Exp[-.77 freq Sqrt[diel] (dout + din) Pi / (2 light)]
```

Calculate the impedence

```
Zocoax[dout_, din_, diel_, cond_, wr_] :=
 Sqrt[((I wr Lcoax[dout, din]) /
     (I wr Ccoax[dout, din, diel]))] +
  Rscoax[dout, din, cond, wr]
```

Calculate the data vector

```
Coax[dout_, din_, len_, diel_, cond_, wr_, zterm_] :=
 {{{Abs[Rho[Zocoax[dout, din, diel, cond, wr], zterm]],
    Losscoax[dout, din, diel, freq] (1 - Abs[
        Rho[Zocoax[dout, din, diel, cond, wr], zterm]])},
   {Losscoax[dout, din, diel, freq] (1 - Abs[
        Rho[Zocoax[dout, din, diel, cond, wr], zterm]]),
    Abs[Rho[Zocoax[dout, din, diel, cond, wr], zterm]]}},
  Ccoax[dout, din, diel] len, Sqrt[diel] / light len}
```

■ Planar Pair Pogo Pin inductive model

Pin[cr_,dout, din, len, diel, wr, zterm]

len = length of exposed pin [m]
dsep = separation distance of pins [m]
dpin = diameter of pins, assumed equal [m]
assumes air dielectric

```
Lpppp[dout_, din_, len_] := len (μo / π) Log[dout / din]
```

```
Pin[cr_, dout_, din_, len_, diel_ , wr_, zterm_] :=
 {SeriesS[wr Lpppp[dout, din, len] + cr, zterm], 0,
  Sqrt[diel] / light len}
```

■ Pad model

Cappad[cap, num,wr, zterm]

This is a very simple model

```
Cappad[cap_, num_, wr_, zterm_] :=
 {ShuntSreal[1 / (wr cap num), zterm], num cap, 0}
```

■ Wire model

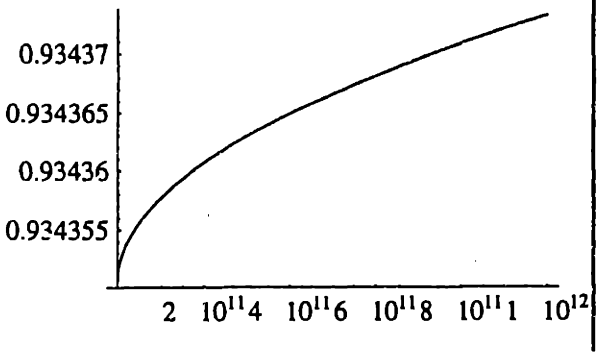
Wire[cr, din, len, cond, wr, zterm]

Equations from Wadell, pg 382

- **Clear required variables**
- **Curve-fitted data for high freq. AC effects [Grover]**
- **Calculate inductance [Henry]**

- **The above model is overkill, Tfunction is essentially 0.9344 at all but very high frequencies. Therefore the inductance model can be simplified by removing the frequency dependence as follows: (value of conductivity is for copper, wire diameter is 10 mils)**

```
Plot[Tfunction[.00025,frequency,5.183*^7],{frequency,0,10^12}]
```




```
WireInduct[din_, len_, cond_] :=
 0.002 (len 100) (Log[4 (len 100) / (din 100)] -
    1 + (din 100) / (2 (len 100)) + 0.9344 / 4)
  10^-6
```

```
Wire[cr_, din_, len_, cond_, wr_, zterm_] :=
 {SeriesS[wr WireInduct[len, din, cond] + cr, zterm],
  0, len / light}
```

- **Microstrip model**

  Microstrip[dout, din, din2, len, diel, cond, wr, zterm]

Add skin-effect resistivity (this is an approximation, need to revisit...)

```
Rsmstrip[dout_, din2_, cond_, wr_] :=
 Sqrt[(Pi wr) / cond] 4 (dout + din2)
```

This is from AIT presentation (Thomas Buck)

```
Zmstrip[dout_, din_, din2_, diel_, cond_, wr_] :=
 87 / Sqrt[diel + 1.41] Log[5.98 dout / (.8 din + din2)] +
  Rsmstrip[dout, din2, cond, wr]
```

Calculate the capacitance

```
Cmstrip[dout_, din_, diel_] := (din / dout) diel ∈o
```

Calculate the dielectric loss

This is the same loss equation used in the coax model with the addition of an exponential coefficient based on the line length. Basicall, the (15-40 len) curve-fits 2 data points: 6" of mstrip is good to 2.2 GHz and 2" is good to 3.2GHz.

```
Lossmstrip[dout_, din_, len_, diel_, freq_] := Exp[-.77
    (15 - 40 len) freq Sqrt[diel] (dout + din) Pi / (2 light)]
```

Calculate the data vector

```
Microstrip[dout_, din_, din2_,
  len_, diel_, cond_, wr_, zterm_] := {{{Abs[Rho[
      Zmstrip[dout, din, din2, diel, cond, wr], zterm]],
    Lossmstrip[dout, din, len, diel, freq]
     (1 - Abs[Rho[Zmstrip[dout,
          din, din2, diel, cond, wr], zterm]])},
   {Lossmstrip[dout, din, len, diel, freq]
     (1 - Abs[Rho[Zmstrip[dout,
          din, din2, diel, cond, wr], zterm]]),
    Abs[Rho[Zmstrip[dout,
       din, din2, diel, cond, wr], zterm]]}},
  Cmstrip[dout, din, diel] len, Sqrt[diel] / light len}
```

- Simple microstrip solver: input desired impedence, strip dimensions and dielectric constant and it will output dielectric thickness (ignores skin effect)

```
thick = Solve[87 / Sqrt[diel + 1.41]
    Log[5.98 dout / (.8 din + din2)] == 50,
  dout]
```

$$\left\{\left\{\text{dout} \to 0.00334448\ (40.\ \text{din} + 50.\ \text{din2})\ E^{\left(0.57471264367816091954022988505750 \Big/ \sqrt{1.4100000000000000000000000000000 + \text{diel}}\right)}\right\}\right\}$$

Setup Scenario

Clear parameters that will be used as variables in this scenario

```
Clear[fmag, zterm]
```

Setup Component Models

Channel Card models

Calculate parameters of individual components

```
ccpadS = Cappad[ccpadcap, ccpadnum, freq, zterm];
ccpadAB = TransSABCD[ccpadS, zterm];
```

```
ccmstripS =
 Microstrip[ccmstripdout, ccmstripdin, ccmstripdin2,
  ccmstriplen, ccmstripdiel, ccmstripcond, freq, zterm];
ccmstripAB = TransSABCD[ccmstripS, zterm];
```

```
cccoaxS = Coax[cccoaxdout, cccoaxdin,
  cccoaxlen, cccoaxdiel, cccoaxcond, freq, zterm];
cccoaxAB = TransSABCD[cccoaxS, zterm];
```

```
ccpogoS = Coax[ccpogodout, ccpogodin,
  ccpogolen, ccpogodiel, ccpogocond, freq, zterm];
ccpogoAB = TransSABCD[ccpogoS, zterm];
```

```
ccpinS = Pin[ccpincr, ccpindout,
  ccpindin, ccpinlen, ccpindiel, freq, zterm];
ccpinAB = TransSABCD[ccpinS, zterm];
```

Cascade components to generate model for entire channel card

```
cctotalAB =
 Cascade[ccpadAB, Cascade[Cascade[ccmstripAB, cccoaxAB],
   Cascade[ccpogoAB, ccpinAB]]];
cctotalS = TransABCDS[cctotalAB, zterm];
```

Convert to S matrix form

```
cctotalS = TransABCDS[cctotalAB, zterm];
```

DIB models

Calculate parameters of individual DIB components

```
dibpadS = Cappad[dibpadcap, dibpadnum, freq, zterm];
dibpadAB = TransSABCD[dibpadS, zterm];
```

```
dibmstripS = Microstrip[dibmstripdout,
  dibmstripdin, dibmstripdin2, dibmstriplen,
  dibmstripdiel, dibmstripcond, freq, zterm];
dibmstripAB = TransSABCD[dibmstripS, zterm];
```

```
dibviaS = Cappad[dibviacap, dibvianum, freq, zterm];
dibviaAB = TransSABCD[dibviaS, zterm];
```

Cascade components to generate model for entire DIB

```
dibtotalAB =
 Cascade[dibpadAB, Cascade[dibmstripAB, dibviaAB]];
```

Convert to S matrix form

```
dibtotalS = TransABCDS[dibtotalAB, zterm];
```

Pogo Tower models

Calculate parameters of individual probe tower components

```
ptpinS = Pin[ptpincr, ptpindout,
  ptpindin, ptpinlen, ptpindiel, freq, zterm];
ptpinAB = TransSABCD[ptpinS, zterm];
```

```
ptpogoS = Coax[ptpogodout, ptpogodin,
  ptpogolen, ptpogodiel, ptpogocond, freq, zterm];
ptpogoAB = TransSABCD[ptpogoS, zterm];
```

Cascade components to generate model for entire probe tower

```
pttotalAB = Cascade[ptpinAB, ptpogoAB];
```

Convert to S matrix form

```
pttotalS = TransABCDS[pttotalAB, zterm];
```

Probe Card models

Calculate parameters of individual probe tower components

```
pcpadS = Cappad[pcpadcap, pcpadnum, freq, zterm];
pcpadAB = TransSABCD[pcpadS, zterm];
```

```
pcmstripS =
 Microstrip[pcmstripdout, pcmstripdin, pcmstripdin2,
  pcmstriplen, pcmstripdiel, pcmstripcond, freq, zterm];
pcmstripAB = TransSABCD[pcmstripS, zterm];
```

```
pcneedleS = Wire[pcneedlecr, pcneedledin,
  pcneedlelen, pcneedlecond, freq, zterm];
pcneedleAB = TransSABCD[pcneedleS, zterm];
```

Cascade components to generate model of entire probe card

```
pctotalAB =
 Cascade[pcpadAB, Cascade[pcmstripAB, pcneedleAB]];
```

Convert to S matrix form

```
pctotalS = TransABCDS[pctotalAB, zterm];
```

Cascade component models

```
totalsystemAB = Cascade[Cascade[cctotalAB, dibtotalAB],
  Cascade[pttotalAB, pctotalAB]];
totalsystemS = TransABCDS[totalsystemAB, zterm];
```

Define Scenario

```
zterm = 50;
```

Test models

Define plot ranges

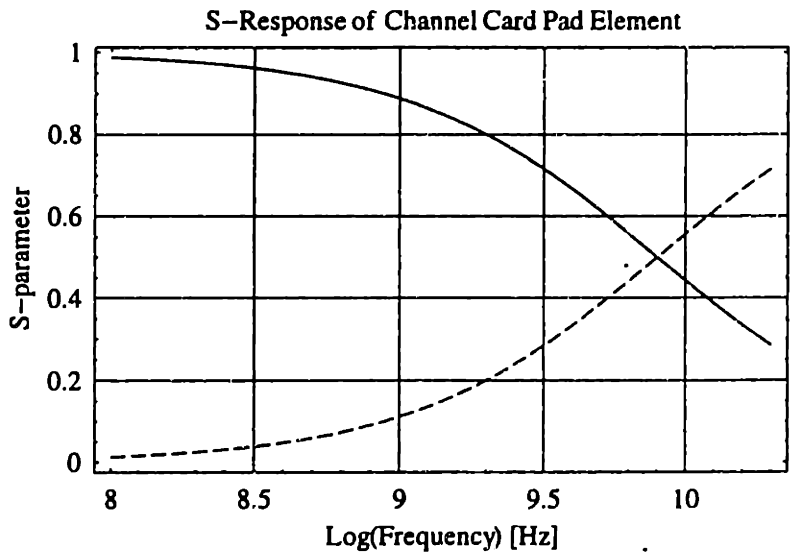
```
low = 5;     (* 100 KHz *)
middle = 8;  (* 100 MHz *)
high = 10.3; (* 20 GHz *)
```

Test Channel Card models

```
ccpadS[[2]]
ccpadS[[3]]
Plot[{ccpadS[[1, 1, 1]], ccpadS[[1, 1, 2]]},
 {fmag, middle, high},
 PlotStyle -> {Dashing[{0.02, 0.01}], Dashing[{1, 0}]},
 GridLines -> Automatic, Frame -> True,
 FrameLabel -> {"Log(Frequency) [Hz]", "S-parameter",
   "S-Response of Channel Card Pad Element", ""},
 ImageSize -> {300, 200}]
```

$8. \times 10^{-13}$

0

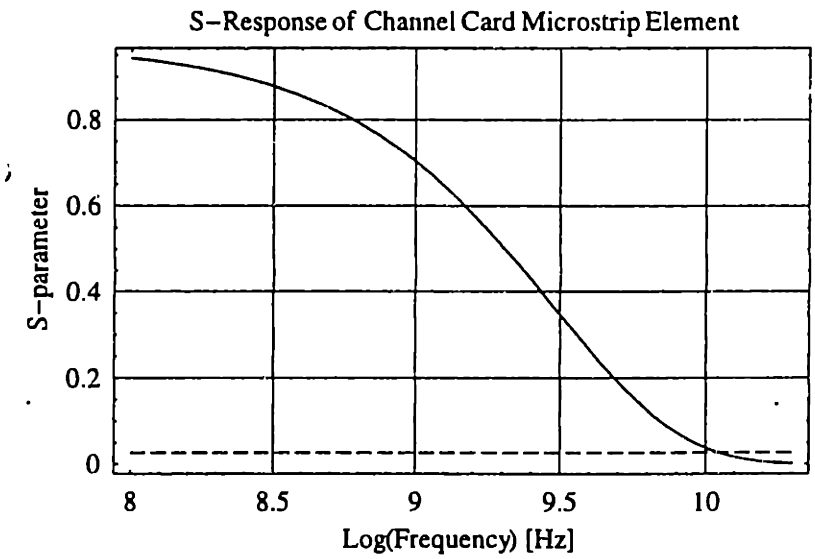


- Graphics -

```
ccmstripS[[2]]
ccmstripS[[3]]
Plot[{ccmstripS[[1, 1, 1]],
  ccmstripS[[1, 1, 2]]}, {fmag, middle, high},
 PlotStyle -> {Dashing[{0.02, 0.01}], Dashing[{1, 0}]},
 GridLines -> Automatic, Frame -> True, FrameLabel ->
  {"Log(Frequency) [Hz]", "S-parameter", "S-Response
     of Channel Card Microstrip Element", ""},
 ImageSize -> {300, 200}]
```

1.40557×10^{-12}

1.76777×10^{-10}

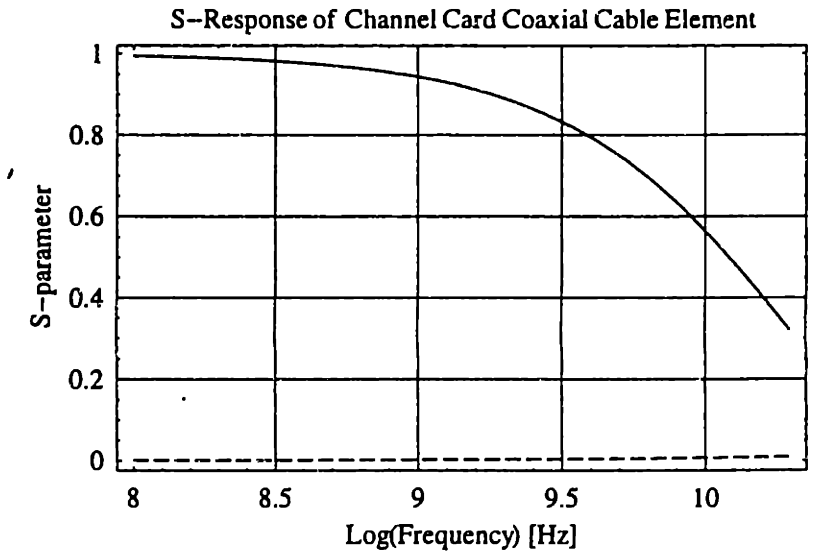


- Graphics -

```
cccoaxS[[2]]
cccoaxS[[3]]
Plot[{cccoaxS[[1, 1, 1]], cccoaxS[[1, 1, 2]]},
 {fmag, middle, high},
 PlotStyle -> {Dashing[{0.02, 0.01}], Dashing[{1, 0}]},
 GridLines -> Automatic, Frame -> True, FrameLabel ->
  {"Log(Frequency) [Hz]", "S-parameter", "S-Response
     of Channel Card Coaxial Cable Element", ""},
 ImageSize -> {300, 200}]
```

7.91552×10^{-12}

3.9525×10^{-10}

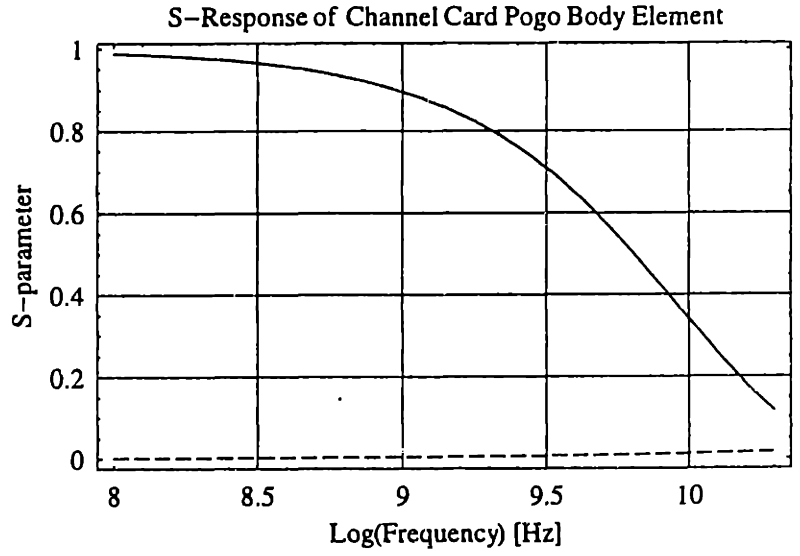


- Graphics -

```
ccpogoS[[2]]
ccpogoS[[3]]
Plot[{ccpogoS[[1, 1, 1]], ccpogoS[[1, 1, 2]]},
 {fmag, middle, high},
 PlotStyle -> {Dashing[{0.02, 0.01}], Dashing[{1, 0}]},
 GridLines -> Automatic, Frame -> True,
 FrameLabel -> {"Log(Frequency) [Hz]", "S-parameter",
   "S-Response of Channel Card Pogo Body Element", ""},
 ImageSize -> {300, 200}]
```

3.54184×10^{-12}

1.76934×10^{-10}

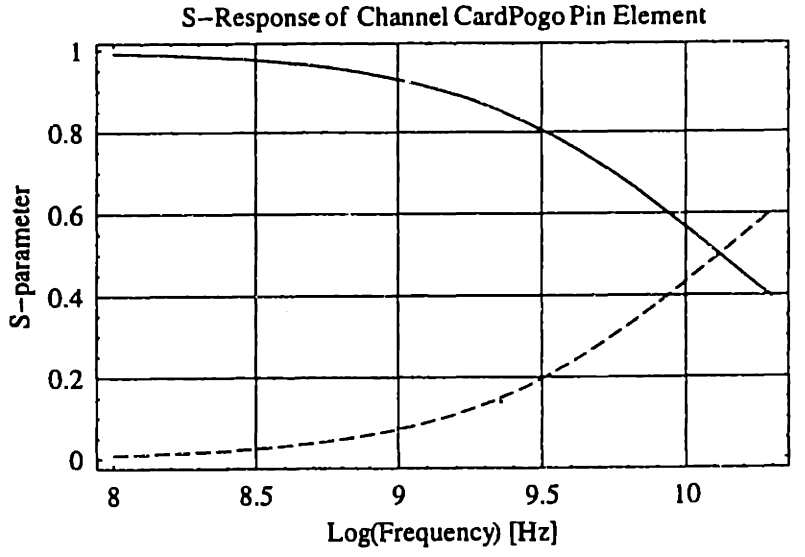


- Graphics -

```
ccpinS[[2]]
ccpinS[[3]]
Plot[{ccpinS[[1, 1, 1]], ccpinS[[1, 1, 2]]},
 {fmag, middle, high},
 PlotStyle -> {Dashing[{0.02, 0.01}], Dashing[{1, 0}]},
 GridLines -> Automatic, Frame -> True,
 FrameLabel -> {"Log(Frequency) [Hz]", "S-parameter",
   "S-Response of Channel CardPogo Pin Element", ""},
 ImageSize -> {300, 200}]

0
```

5.93333×10^{-12}

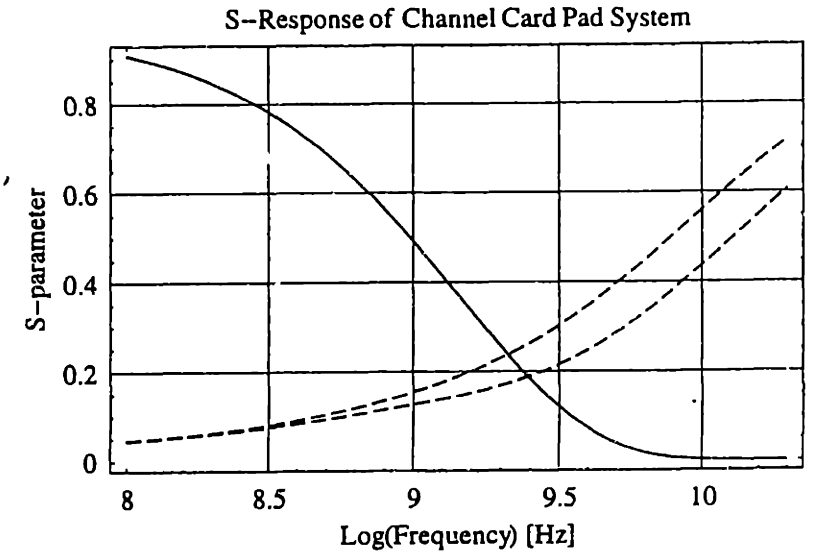


- Graphics -

```
cctotalS[[2]]
cctotalS[[3]]
Plot[{cctotalS[[1, 1, 1]], cctotalS[[1, 1, 2]],
  cctotalS[[1, 2, 2]]}, {fmag, middle, high},
 PlotStyle -> {Dashing[{0.02, 0.01}], Dashing[{1, 0}]},
 GridLines -> Automatic, Frame -> True,
 FrameLabel -> {"Log(Frequency) [Hz]", "S-parameter",
   "S-Response of Channel Card Pad System", ""},
 ImageSize -> {300, 200}]
```

1.36629×10^{-11}

7.54893×10^{-10}



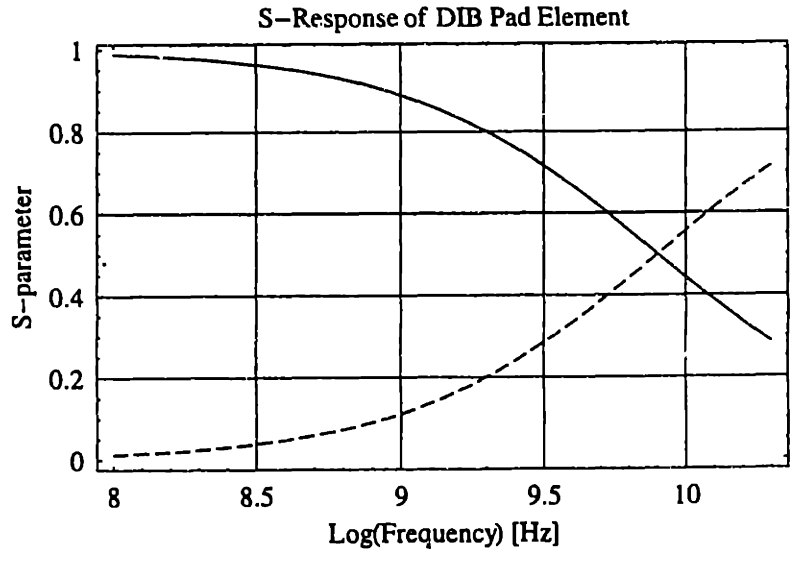
- Graphics -

DIB models

```
dibpadS[[2]]
dibpadS[[3]]
Plot[{dibpadS[[1, 1, 1]], dibpadS[[1, 1, 2]]},
 {fmag, middle, high},
 PlotStyle -> {Dashing[{0.02, 0.01}], Dashing[{1, 0}]},
 GridLines -> Automatic,
 Frame -> True, FrameLabel -> {"Log(Frequency) [Hz]",
   "S-parameter", "S-Response of DIB Pad Element", ""},
 ImageSize -> {300, 200}]
```

$8. \times 10^{-13}$

0

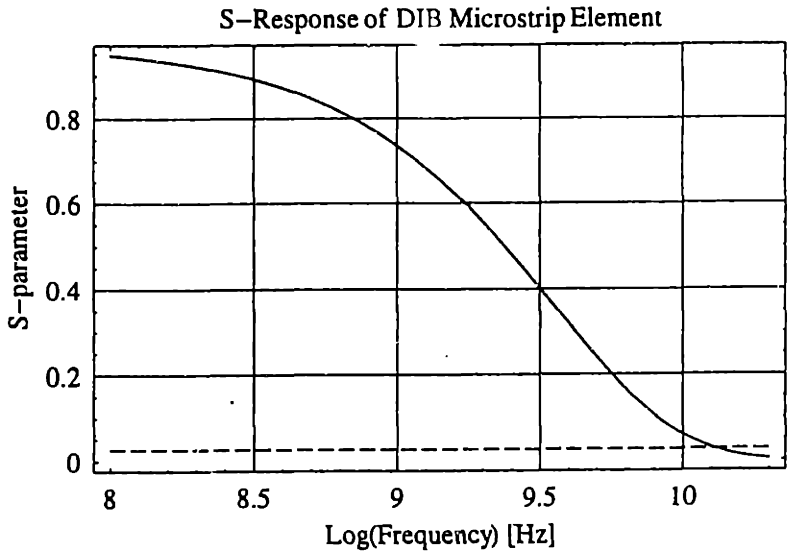


- Graphics -

```
dibmstripS[[2]]
dibmstripS[[3]]
Plot[{dibmstripS[[1, 1, 1]],
  dibmstripS[[1, 1, 2]]}, {fmag, middle, high},
 PlotStyle -> {Dashing[{0.02, 0.01}], Dashing[{1, 0}]},
 GridLines -> Automatic, Frame -> True,
 FrameLabel -> {"Log(Frequency) [Hz]", "S-parameter",
   "S-Response of DIB Microstrip Element", ""},
 ImageSize -> {300, 200}]
```

4.21672×10^{-12}

5.3033×10^{-10}

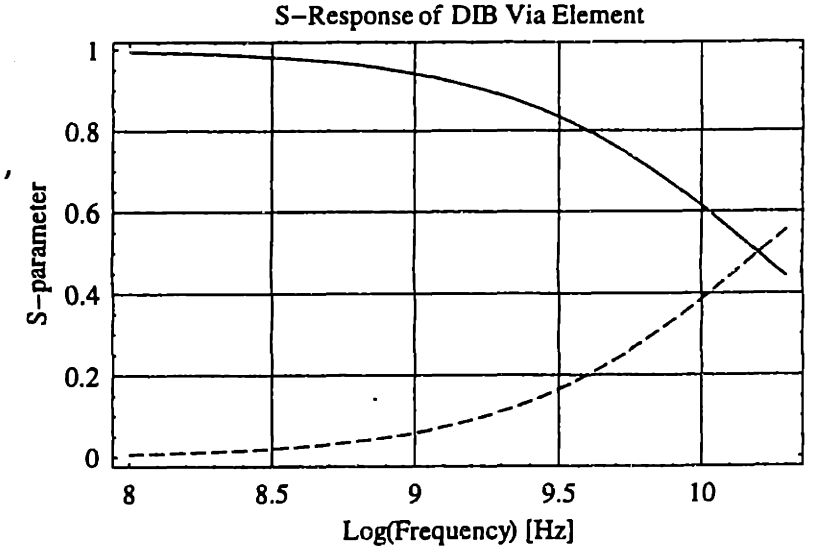


- Graphics -

```
dibviaS[[2]]
dibviaS[[3]]
Plot[{dibviaS[[1, 1, 1]], dibviaS[[1, 1, 2]]},
 {fmag, middle, high},
 PlotStyle -> {Dashing[{0.02, 0.01}], Dashing[{1, 0}]},
 GridLines -> Automatic,
 Frame -> True, FrameLabel -> {"Log(Frequency) [Hz]",
   "S-parameter", "S-Response of DIB Via Element", ""},
 ImageSize -> {300, 200}]
```

$4. \times 10^{-13}$

0

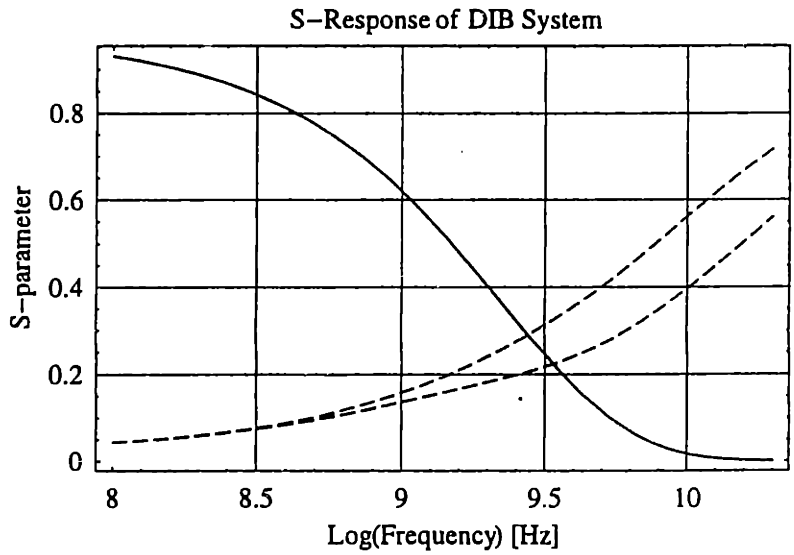


- Graphics -

```
dibtotalS[[2]]
dibtotalS[[3]]
Plot[{dibtotalS[[1, 1, 1]], dibtotalS[[1, 1, 2]],
  dibtotalS[[1, 2, 2]]}, {fmag, middle, high},
 PlotStyle -> {Dashing[{0.02, 0.01}], Dashing[{1, 0}]},
 GridLines -> Automatic,
 Frame -> True, FrameLabel -> {"Log(Frequency) [Hz]",
   "S-parameter", "S-Response of DIB System", ""},
 ImageSize -> {300, 200}]
```

5.41672×10^{-12}

5.3033×10^{-10}



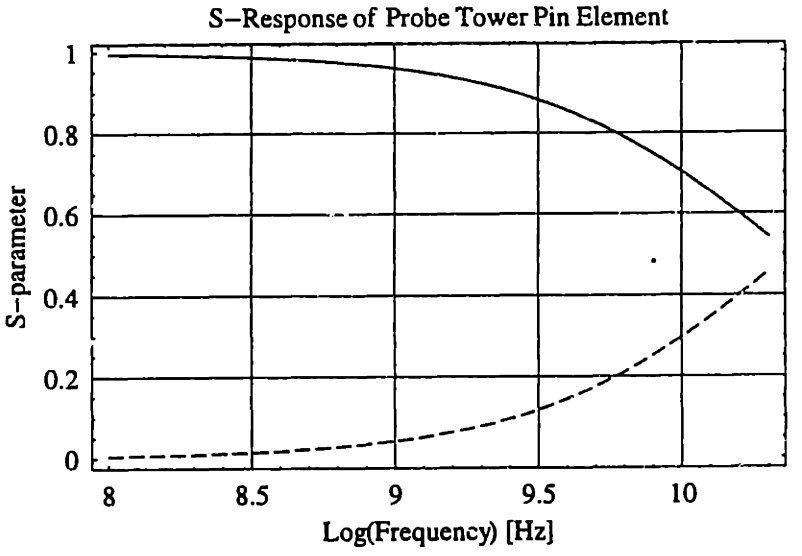
- Graphics -

Probe Tower models

```
ptpinS[[2]]
ptpinS[[3]]
Plot[{ptpinS[[1, 1, 1]], ptpinS[[1, 1, 2]]},
 {fmag, middle, high},
 PlotStyle -> {Dashing[{0.02, 0.01}], Dashing[{1, 0}]},
 GridLines -> Automatic, Frame -> True,
 FrameLabel -> {"Log(Frequency) [Hz]", "S-parameter",
   "S-Response of Probe Tower Pin Element", ""},
 ImageSize -> {300, 200}]
```

0

8.33333×10^{-12}

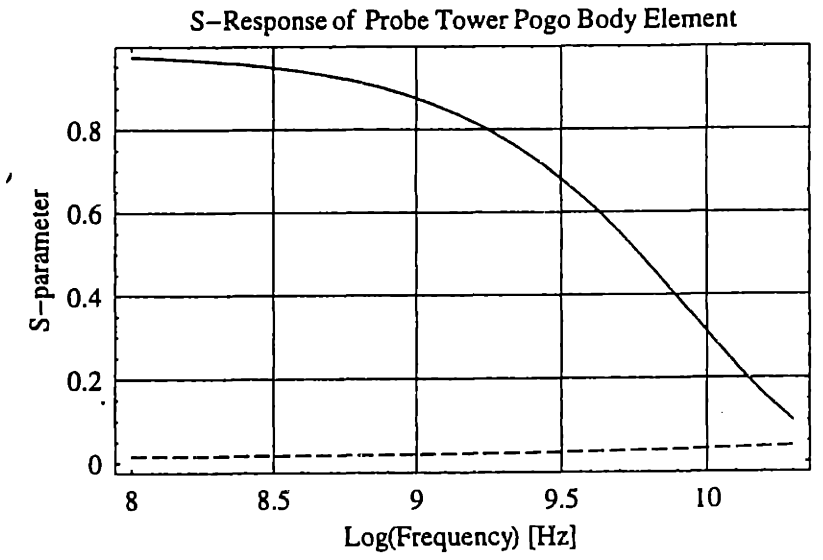


- Graphics -

```
ptpogoS[[2]]
ptpogoS[[3]]
Plot[{ptpogoS[[1, 1, 1]], ptpogoS[[1, 1, 2]]},
 {fmag, middle, high},
 PlotStyle -> {Dashing[{0.02, 0.01}], Dashing[{1, 0}]},
 GridLines -> Automatic, Frame -> True,
 FrameLabel -> {"Log(Frequency) [Hz]", "S-parameter",
   "S-Response of Probe Tower Pogo Body Element", ""},
 ImageSize -> {300, 200}]
```

2.76854×10^{-12}

1.42361×10^{-10}

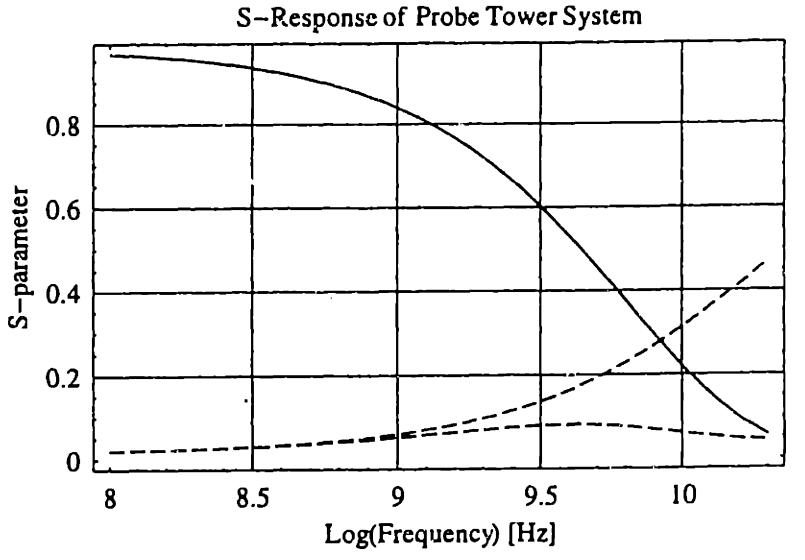


- Graphics -

```
pttotalS[[2]]
pttotalS[[3]]
Plot[{pttotalS[[1, 1, 1]], pttotalS[[1, 1, 2]],
  pttotalS[[1, 2, 2]]}, {fmag, middle, high},
 PlotStyle -> {Dashing[{0.02, 0.01}], Dashing[{1, 0}]},
 GridLines -> Automatic, Frame -> True,
 FrameLabel -> {"Log(Frequency) [Hz]", "S-parameter",
   "S-Response of Probe Tower System", ""},
 ImageSize -> {300, 200}]
```

2.76854×10^{-12}

1.50694×10^{-10}



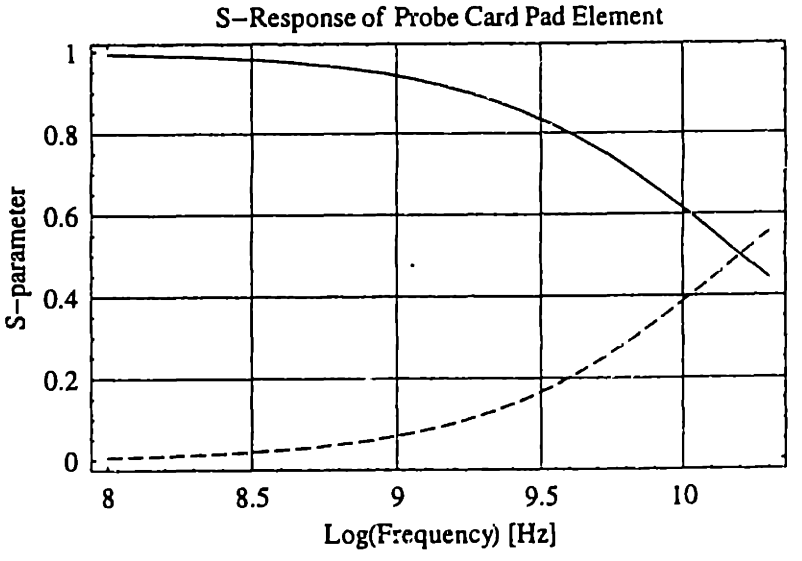
- Graphics -

Probe Card models

```
pcpadS[[2]]
pcpadS[[3]]
Plot[{pcpadS[[1, 1, 1]], pcpadS[[1, 1, 2]]},
 {fmag, middle, high},
 PlotStyle -> {Dashing[{0.02, 0.01}], Dashing[{1, 0}]},
 GridLines -> Automatic, Frame -> True,
 FrameLabel -> {"Log(Frequency) [Hz]", "S-parameter",
   "S-Response of Probe Card Pad Element", ""},
 ImageSize -> {300, 200}]
```

$4. \times 10^{-13}$

0

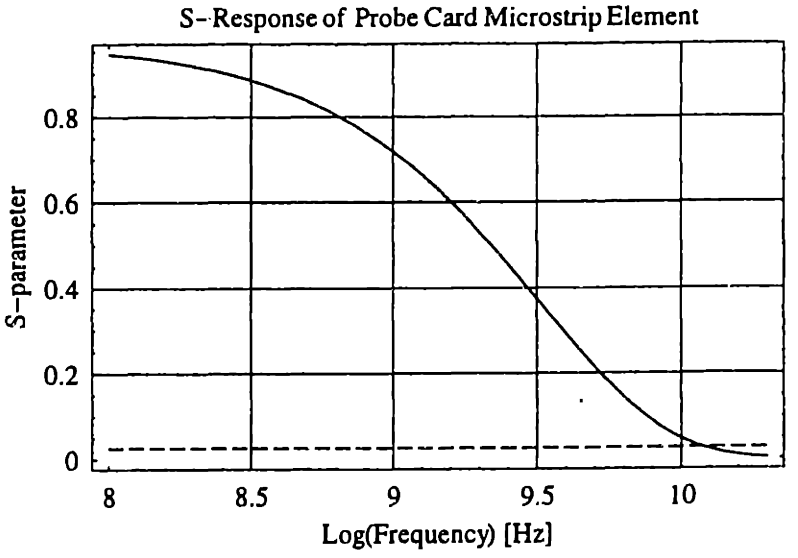


- Graphics -

```
pcmstripS[[2]]
pcmstripS[[3]]
Plot[{pcmstripS[[1, 1, 1]],
  pcmstripS[[1, 1, 2]]}, {fmag, middle, high},
 PlotStyle -> {Dashing[{0.02, 0.01}], Dashing[{1, 0}]},
 GridLines -> Automatic, Frame -> True,
 FrameLabel -> {"Log(Frequency) [Hz]", "S-parameter",
   "S-Response of Probe Card Microstrip Element", ""},
 ImageSize -> {300, 200}]
```

2.81115×10^{-12}

3.53553×10^{-10}

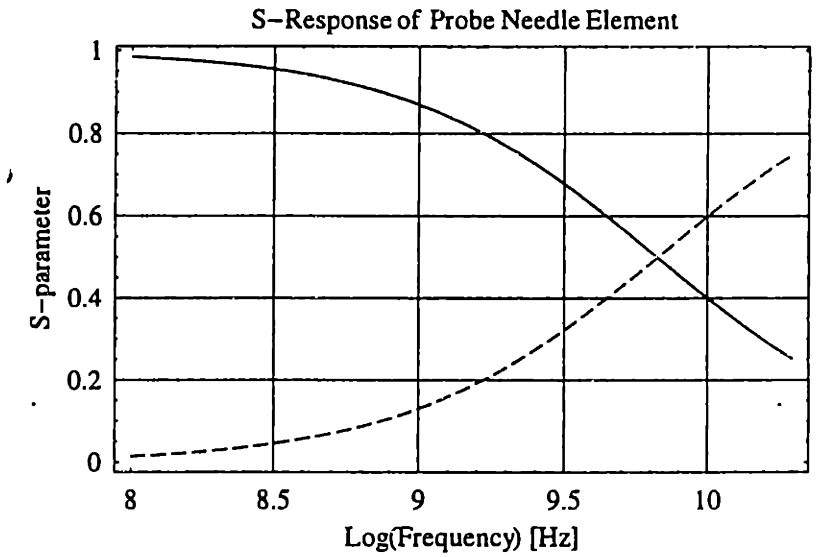


- Graphics -

```
pcneedleS[[2]]
pcneedleS[[3]]
Plot[{pcneedleS[[1, 1, 1]],
  pcneedleS[[1, 1, 2]]}, {fmag, middle, high},
 PlotStyle -> {Dashing[{0.02, 0.01}], Dashing[{1, 0}]},
 GridLines -> Automatic, Frame -> True,
 FrameLabel -> {"Log(Frequency) [Hz]", "S-parameter",
   "S-Response of Probe Needle Element", ""},
 ImageSize -> {300, 200}]
```

0

8.33333×10^{-11}

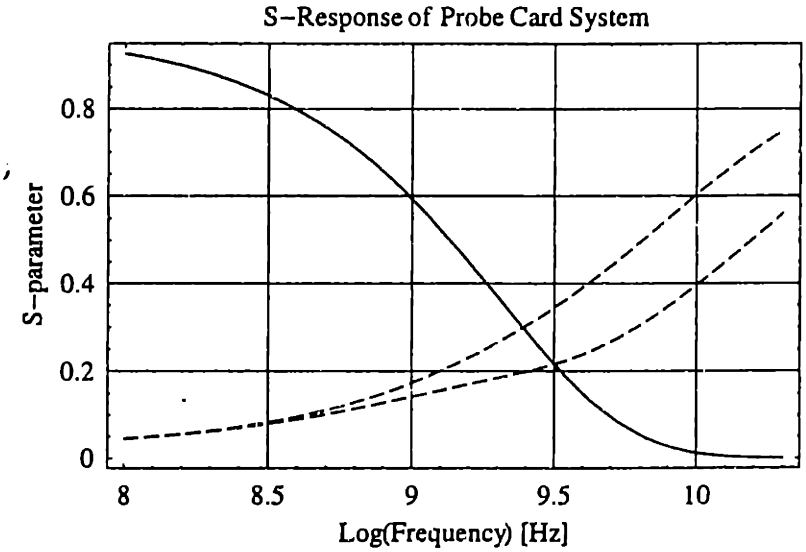


- Graphics -

```
pctotalS[[2]]
pctotalS[[3]]
Plot[{pctotalS[[1, 1, 1]], pctotalS[[1, 1, 2]],
  pctotalS[[1, 2, 2]]}, {fmag, middle, high},
 PlotStyle -> {Dashing[{0.02, 0.01}], Dashing[{1, 0}]},
 GridLines -> Automatic, Frame -> True,
 FrameLabel -> {"Log(Frequency) [Hz]", "S-parameter",
   "S-Response of Probe Card System", ""},
 ImageSize -> {300, 200}]
```

3.21114×10^{-12}

4.36887×10^{-10}



- Graphics -

System Model

System Results

```
fmag = 8.5;  (* 300 MHz *)
totalsystemS[[1]]
fmag = 9;    (* 1 GHz *)
totalsystemS[[1]]
fmag = 9.5; (* 3 GHz *)
totalsystemS[[1]]
fmag = 10;  (* 10 GHz *)
totalsystemS[[1]]
```

```
{{0.172206 , 0.522096 }, {0.522096 , 0.182138 }}
```

```
{{0.210068 , 0.161569 }, {0.161569 , 0.240365 }}
```

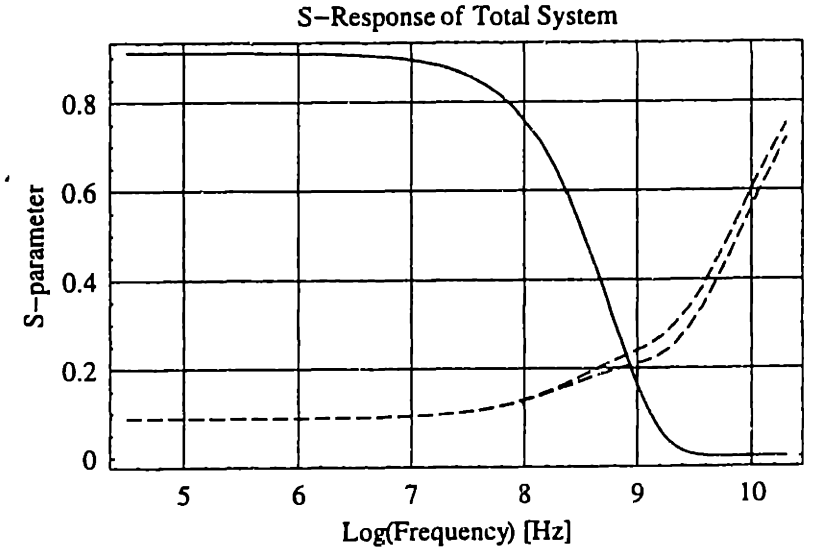
```
{{0.307815 , 0.00436839 }, {0.00436839 , 0.352584 }}
```

```
{{0.562176 , 1.24584 ×10^-7}, {1.24584 ×10^-7, 0.604074 }}
```

```
totalsystemS[[2]]
totalsystemS[[3]]
Plot[{totalsystemS[[1, 1, 1]], totalsystemS[[1, 1, 2]],
  totalsystemS[[1, 2, 2]]}, {fmag, 4.5, high},
 PlotStyle -> {Dashing[{0.02, 0.01}], Dashing[{1, 0}]},
 GridLines -> Automatic,
 Frame -> True, FrameLabel -> {"Log(Frequency) [Hz]",
   "S-parameter", "S-Response of Total System", ""},
 ImageSize -> {300, 200}]
```

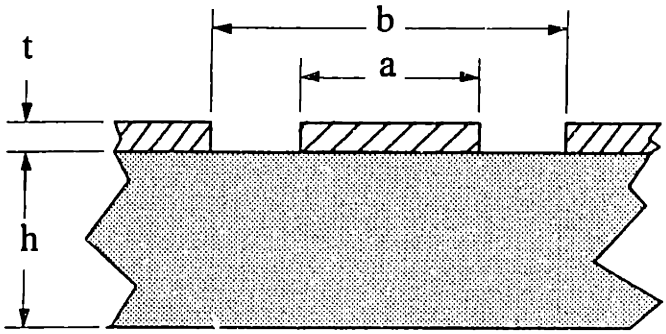
2.50593×10^{-11}

1.8728×10^{-9}



- Graphics -

9.1.1 Impedance Calculations for Coplanar Waveguides



The following equation can be used for calculating the nominal impedance of a coplanar waveguide[47].

$$Z_o = \frac{30.0\pi}{\sqrt{\varepsilon_{eff,t}}} \frac{K(k_t')}{K(k_t)}$$

The K-factors correct this equation for trace thickness and can be calculated as follows:

$$\varepsilon_{eff,t} = \varepsilon_{eff} - \frac{\varepsilon_{eff} - 1}{\frac{(b-a)}{1.4t}\frac{K(k)}{K'(k)} + 1} \qquad \varepsilon_{eff} = 1 + \frac{\varepsilon_r - 1}{2}\frac{K(k')K(k_1)}{K(k)K(k_1)}$$

$$k_t = \frac{a_t}{b_t} \qquad k = \frac{a}{b}$$

$$k_t' = \sqrt{1 - k_t^2} \qquad k' = \sqrt{1 - k^2}$$

$$k_1 = \frac{\sinh(\frac{\pi a_t}{4h})}{\sinh(\frac{\pi b_t}{4h})} \qquad k_1' = \sqrt{1 - k_1^2}$$

$$a_t = a + \frac{5t}{4\pi}\left[1 + \ln\left(\frac{4\pi a}{t}\right)\right] \qquad b_t = b - \frac{5t}{4\pi}\left[1 - \ln\left(\frac{4\pi a}{t}\right)\right]$$

Dielectric, conductor and radiation losses can also be calculated for coplanar waveguides. These equations can be found in the cited reference.

[47] Wadell, 1991

9.2 Autoplanarization Data

This appendix contains data from the prototype module development including autoplanarization concept and prototype development as well as equations and recommendations for the implementation of this design.

9.2.1 Proof of Concept and Autoplanarization Development[48]

In an effort to verify the operation and performance of the kinematic coupling, a proof-of-concept prototype was built consisting of an aluminum weldment to simulate a testhead structure and a 3/8" plate to simulate the structure of a prober. Three prototype kinematic modules that had the basic design functionality (ball, lock and pull-down mechanism to groove, motorized Z-adjust) were mounted to the weldment and three aluminum grooves were mounted to the plate. The preload mechanism uses a T-bolt in conjunction with an air cylinder to pull the V-groove and the ball assembly together. A small air cylinder, supplied with 60 PSI provided the preload force.

The kinematic module was designed to have adjustable z-height. The height adjustment system is computer controlled and consists of a DC motor, a belt drive system used to drive the leadscrew, and a linear bearing which incorporates recirculating balls on a round shaft (e.g., a Ball Bushing from Thompson Industries). The leadscrew nut is epoxied to the inside of the round shaft which is an extension of the steel ball. This assembly is shown in Figure 9.1 below.

[48] Some sections excerpted from Alagheband, 1996

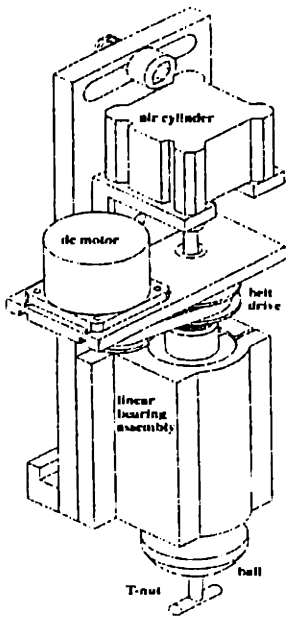


Figure 9.1: Proof of Concept Prototype Ball-Module

9.2.2 Preliminary Repeatability Results

Previous experiments have shown that kinematic arrangements such as the three ball/three V-grooves can provide repeatability of better than 10 micrometers. A brief set of experiments using this testbed demonstrated repeatability on the order of 0.001". Upon closer examination, excessive wear in the aluminum grooves was observed. It would be possible to achieve an order of magnitude improvement in the results by using hardened stainless steel balls and grooves and a preload mechanism capable of exerting higher magnitude forces.

9.2.3 Autoplanarization Proof of Concept

Currently, it is common for the probe card and the DIB to reside on the prober before the testhead is docked. In the future, situations might arise in which the probe card is attached directly to the testhead. In the former case, it is desirable that the pogo pins exert a distributed force uniformly across the DIB board once the testhead is docked. This will minimize deflections of the probe card. This requires that all the pogo pins lie in a horizontal plane. If the probe card is attached to the testhead, it is desirable that all the probe needles be planar to the wafer.

In either scenario, the desired objectives are attainable through a routine known as *autoplanarization.* The objective of autoplanarization is to measure the error in position between the testhead interface plane (defined by the plane of the probe tips) and the plane of the wafer. A algorithm based on the geometry of the testhead is then used to determine the Z-adjustment of each module required to bring these planes into alignment. For example, it is conceivable to determine the x,y, and z coordinates of three probe needles and calculate the equation of a plane. The appropriate combination of z-height adjustments of the KDS modules will ensure that the three probe needles lie in the horizontal plane and are planar to the wafer, this is shown in Figure 9.2 below.

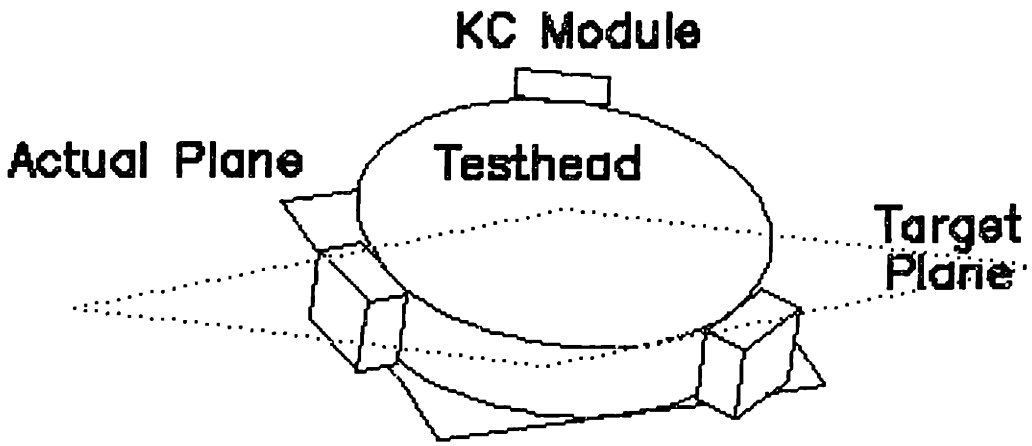


Figure 9.2: Autoplanarization Geometry

An autoplanarization algorithm was implemented using a PC-based closed-loop servo system represented by the schematic shown in Figure 9.3.

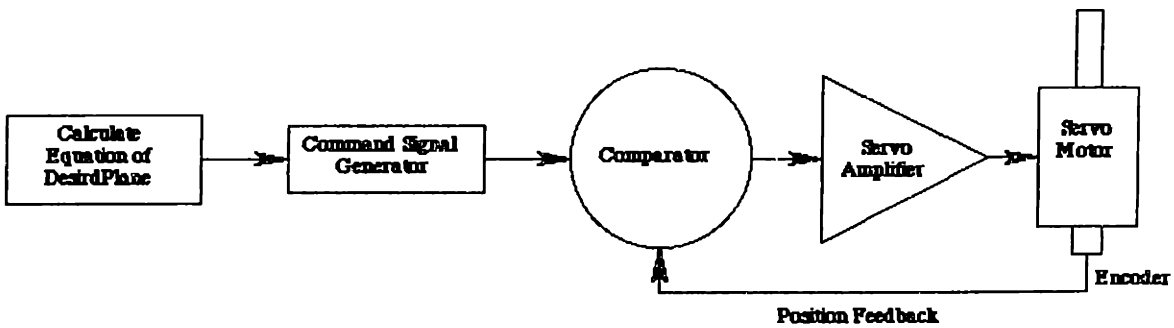


Figure 9.3: Autoplanarization Servo-Loop

The servo system was implemented using a universal motion interface board (UMI 4A) form NuLogic, Inc. and servo amplifiers (MC3X series) from Advanced Motion Controls.

Nulogic provides a user interface based on LabView, a PC-based virtual instrument software package. The servo motors and the corresponding encoders were supplied by Maxon, Inc. The complete control system is shown in Figure 9.4.

Figure 9.4: Autoplanarization Servo Control System

The equation of the desired plane was calculated using three data points acquired from the three Sony Magnescale linear encoders mounted near the center of the testhead weldment. These sensors have an accuracy of 0.002 mm. The encoders indicate the position of the testhead datum relative to simulated wafer chuck mounted at the center of the top plate as shown in Figure 9.5.

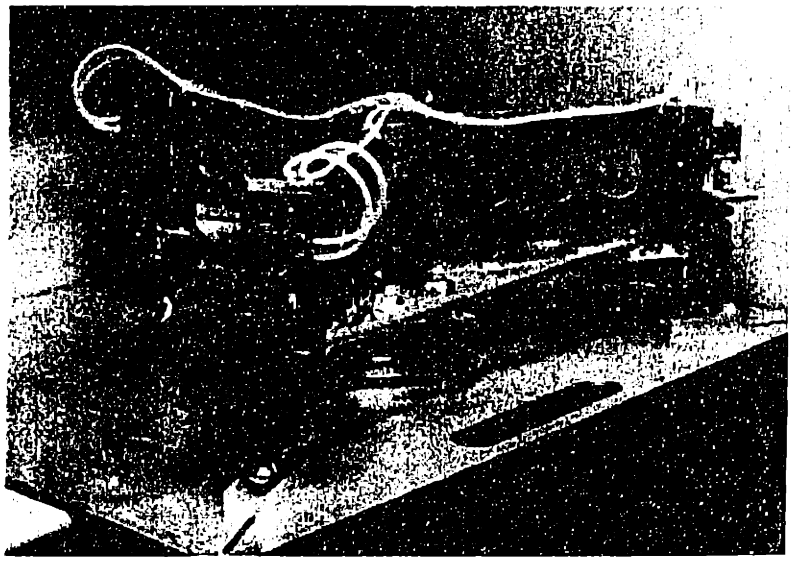


Figure 9.5: Testhead Weldment with Proof-of Concept Modules

The encoders essentially mimic the role of three probe needles on the probe card. In order to planarize the testhead the following algorithm was implemented: The geometry for this system is shown in Figure 9.6 below.

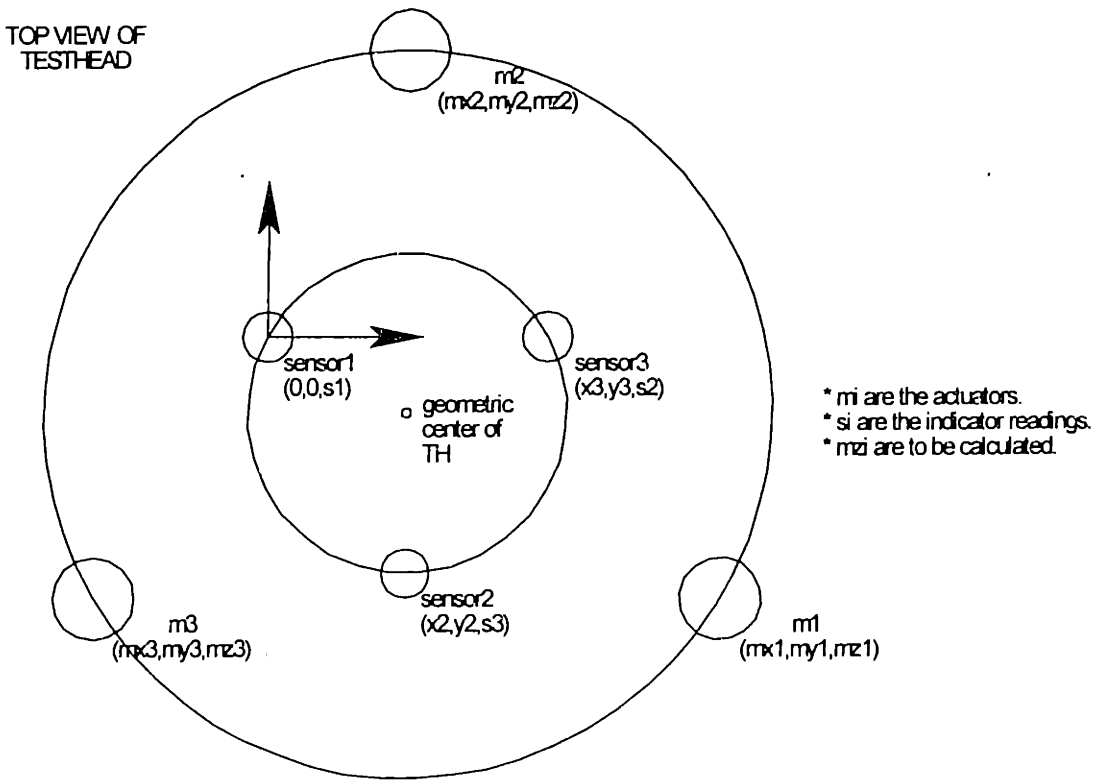


Figure 9.6: Proof of Concept Measurement Geometry

The coordinates of the three sensor contact points can be written as follows

$$P_1 = (0,0,s_1)$$
$$P_2 = (x_2, y_2, s_2)$$
$$P_3 = (x_3, y_3, s_3)$$

Where s_1,s_2 and s_3 are the encoder measurements. These points form a plane of the form Ax+By+Cz+D=0.

Using the points P_i, we write:

$$s_1.C + D = 0$$
$$x_2.A + y_2.B + s_2.C + D = 0$$
$$x_3.A + y_3.B + s_3.C + D = 0$$

If we let C=1, then D=$-s_1$ and we have to solve for A and B.

$$\begin{bmatrix} x_2 & y_2 \\ x_3 & y_3 \end{bmatrix} \cdot \begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} s_1 - s_2 \\ s_1 - s_3 \end{bmatrix}$$

Therefore,

$$\begin{bmatrix} A \\ B \end{bmatrix} = \begin{bmatrix} x_2 & y_2 \\ x_3 & y_3 \end{bmatrix}^{-1} \cdot \begin{bmatrix} s_1 - s_2 \\ s_1 - s_3 \end{bmatrix}$$

$$\begin{bmatrix} A \\ B \end{bmatrix} = \frac{1}{x_2 y_3 - y_2 x_3} \begin{bmatrix} y_3 & -y_2 \\ -x_3 & x_2 \end{bmatrix} \cdot \begin{bmatrix} s_1 - s_2 \\ s_1 - s_3 \end{bmatrix}$$

This determines the coefficients of the plane equation Ax+By+Cz+D=0. In order to calculate the desired z-height of the kinematic modules, the (x,y) coordinates of the actuators are used and the target position is calculated for each module as follows:

z=(Ax+By+D)/C

9.2.4 The Alpha Prototype Development

The next step in the implementation of the kinematic coupling concept was to assemble and evaluate an improved, "alpha" prototype that incorporated some more robust design features. The alpha coupling consists of a hardened steel ball and groove to address the wear issues discussed above and has an automatic latch mechanism as well as a z-height adjustment mechanism. This prototype is shown in Figure 9.7

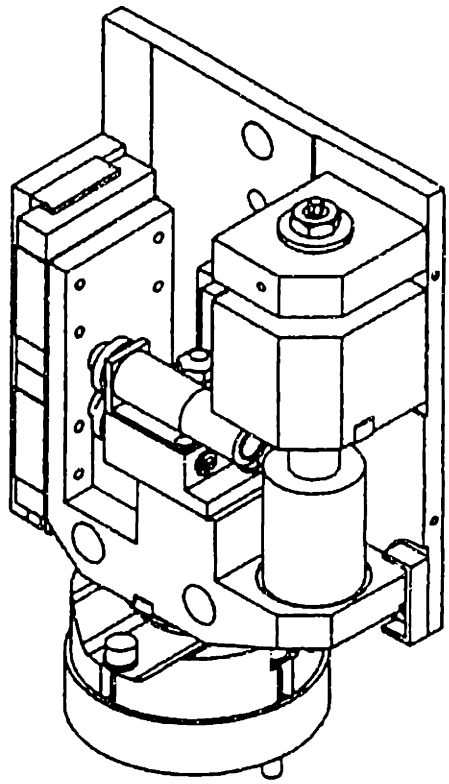


Figure 9.7: Sketch of Alpha Prototype

Preparation of the experimental setup involved the completion of the control electronics for latch/preload sensing, building an improved mockup prober/handler structure, and mounting the kinematic modules on a real testhead structure.

For these experiments, a pre-production testhead casting was used from a testhead that was currently under development.. Six positions were identified as possible mounting locations for the kinematic coupling modules. The couplings are represented by the blocks in Figure E8. The openings on the perimeter of the testhead are necessary for airflow in order to cool the electronics. Note that the couplings located in the front and the back would block the airways if mounted directly on to the testhead. In order to resolve the airflow issue, the modules were mounted on flanges designed to provide structural support for the couplings and to allow airflow in to the testhead. The V-grooves were mounted on the top plate and a mockup cable bundle was attached to the testhead to simulate actual test conditions.

Figure 9.8 shows the mockup prober structure, testhead, cable bundle, one of the kinematic coupling modules and its corresponding V-groove, and digital indicators used for repeatability measurements.

Figure 9.8: Alpha Measurement Setup

The Sony encoders were used to obtain planarity measurement. Figure 9.9shows an encoder mounted on the testhead. The three positions not occupied by the modules were used to mount the sensors.

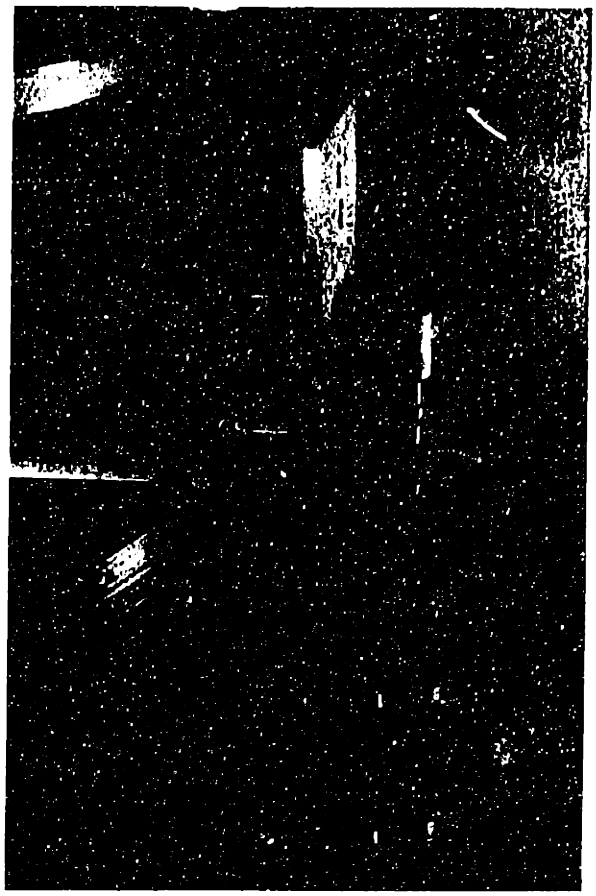


Figure 9.9: Z-Axis Encoder Fixturing

The sensor readings can be utilized to determine the repeatability of the interface in z, tumble (rotation about the x-axis), and twist (rotation about the y-axis). As described above, the sensor readings can be used to determine the equation of the plane the testhead lies in. Once the equation of the plane is known, the repeatability in z height can be measured at the testhead's geometric center (gc). Using the engineering drawings, the exact x-y coordinates of the gc is determined. Using the plane equation, the z-height of the gc relative to the horizontal plane can be calculated. We denote this as z_0.

The repeatability in tumble can be calculated using the following formula:

$$\theta_{tumble} = \tan^{-1} \frac{z_{\#3} - z_0}{y_{\#3}}$$

Where $z_{\#3}$ is the z-coordinate of a point (A) along the tumble axis and is calculated using the plane equation. $y_{\#3}$ is the distance from the gc to point (A).

The repeatability in twist can be calculated similarly as follows:

$$\theta_{twist} = \tan^{-1} \frac{z_j - z_0}{y_j}$$

Where y_j is the distance from the gc to a point along the tumble axis.

In addition to repeatability measurements in planarity, digital indicators mounted on the top plate were utilized to evaluate repeatability of the kinematic interface in x, y, and theta. Figure xx shows one of three digital indicators used. These indicators have an accuracy of 0.0002" and a repeatability of 0.0001". In order to avoid inflicting any damage to the indicator's spindle, a retraction mechanism was designed and mounted on the indicator. The retraction mechanism consists of a retractor lever and a cylindrical Alnico magnet used to keep the spindle retracted while the testhead is docked or undocked. Once the testhead is docked, the spindle is released manually and allowed to make contact with a steel "paddle" bolted to the testhead. The paddle was designed so as to deflect less than 0.0005" when in contact with the indicator's spindle. The spindle exerts 0.43 lbs. of contact force.

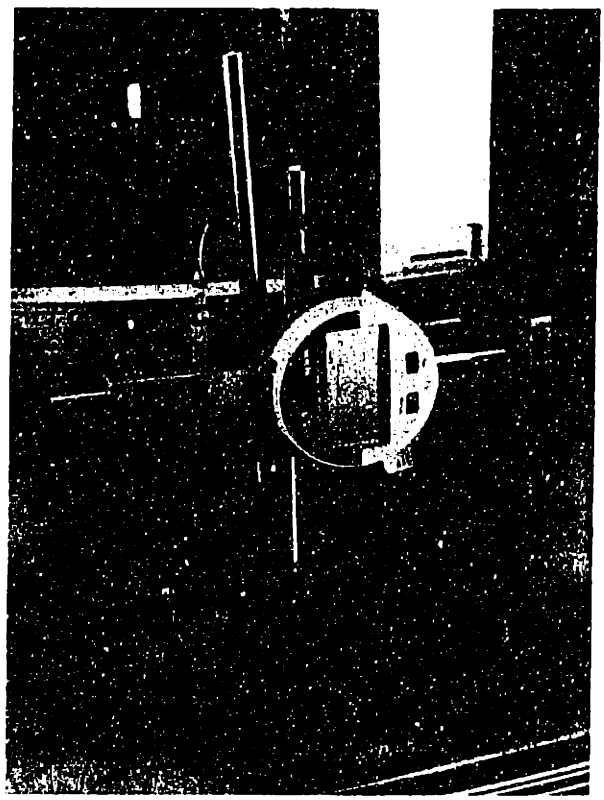


Figure 9.10: X-Y and Yaw Measurement Setup

The digital indicators were mounted on magnetic stands and arranged as shown in Figure 9.10. Repeatability in x, y, and theta can be measured using the following formulas.

From sensor1,

$$(r + \Delta x)\theta - \Delta y = y_1 \quad (A)$$

From sensor2,

$$(r - \Delta x)\theta + \Delta y = y_2 \quad (B)$$

From sensor3,

$$(r_3 - \Delta y)\theta - \Delta x = x_3 \quad (C)$$

Where y_1, y_2, y_3 are the digital indicator readings.

We want to calculate $\Delta x, \Delta y, \theta$.

From (C), we write

$$\Delta x = (r_3 - \Delta y)\theta - x_3$$

Substituting into (A) and (B) and adding, we get

$$\theta = \frac{1}{2r}(y_1 + y_2)$$

Substituting for x, θ in (A), we get

$$\Delta y = \frac{\frac{1}{2}(y_1 + y_2) + (r_3 - x_3)\theta - y_1}{(1 + \theta^2)}$$

9.2.5 Autoplanarization Using the Alpha Prototype

The alpha module was designed with a z-height adjustment mechanism. As discussed in above, this capability allows the operator to either planarize the testhead relative to the DIB or planarize the probe needles to the wafer in cases where the probe card might be directly mounted on the testhead. Recall that the height adjustment mechanism of the original proof of concept module lies along the same axis as the preload mechanism. This was not achievable for consequent design iterations because an automatic latch/preload mechanism was implemented. Vertical motion is achieved by using a rolling element linear motion bearing (linear guide) and a preload nut was used to eliminate any backlash in the leadscrew. The leadscrew is supported by a thrust bearing and z adjustment is achieved by manually turning the leadscrew.

The first kinematic interface design would allow the operators to manually planarize the testhead. The next generation of the modules would incorporate motorized z-height adjustment for autoplanarization. In order to implement autoplanarization on the alpha prototypes, they were modified as shown in Figure 9.11. A Maxon DC servomotor was mounted on the module and a helical flexible coupling was used to drive the leadscrew. The use of the flexible coupling allows for greater radial and angular misalignment of the motor spindle with the leadscrew.

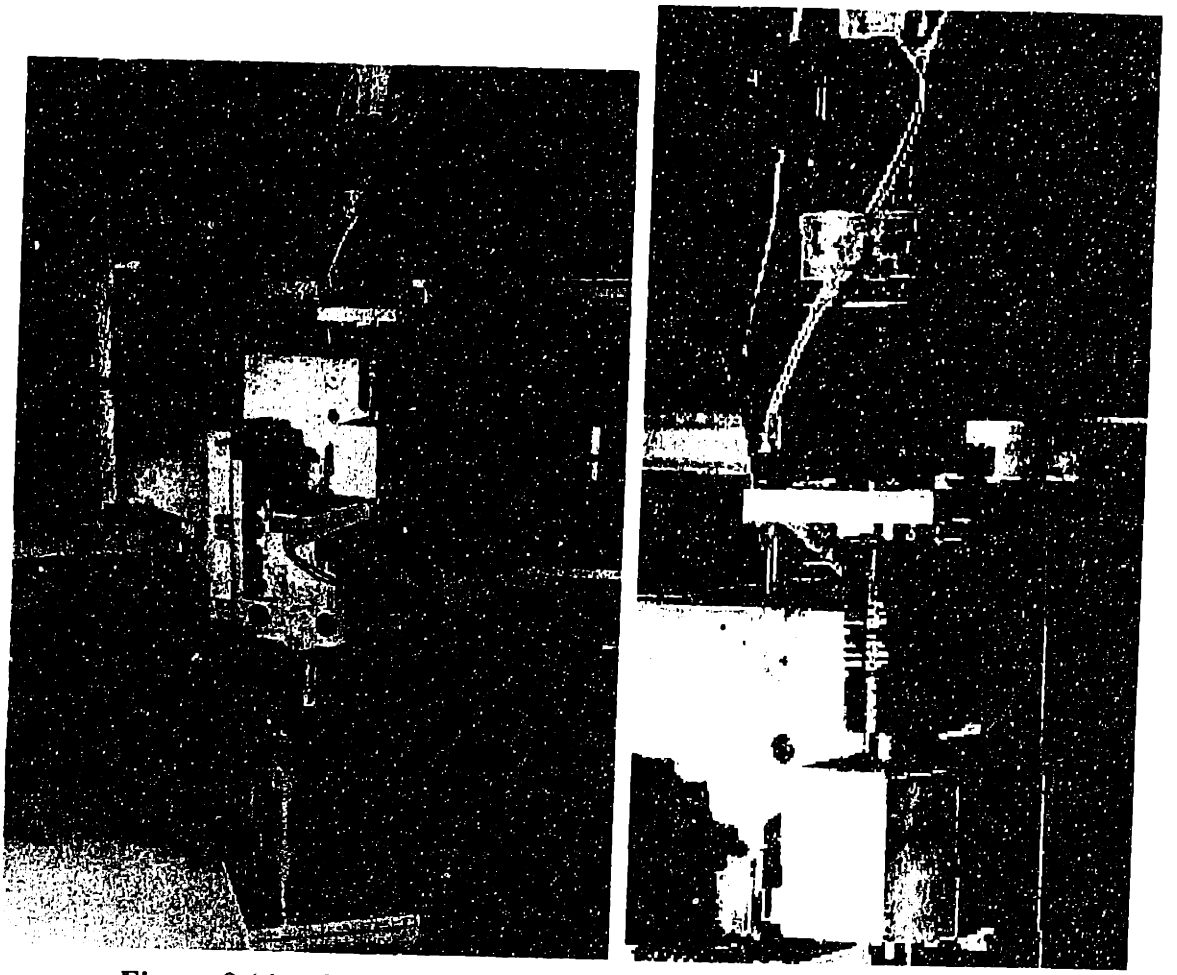


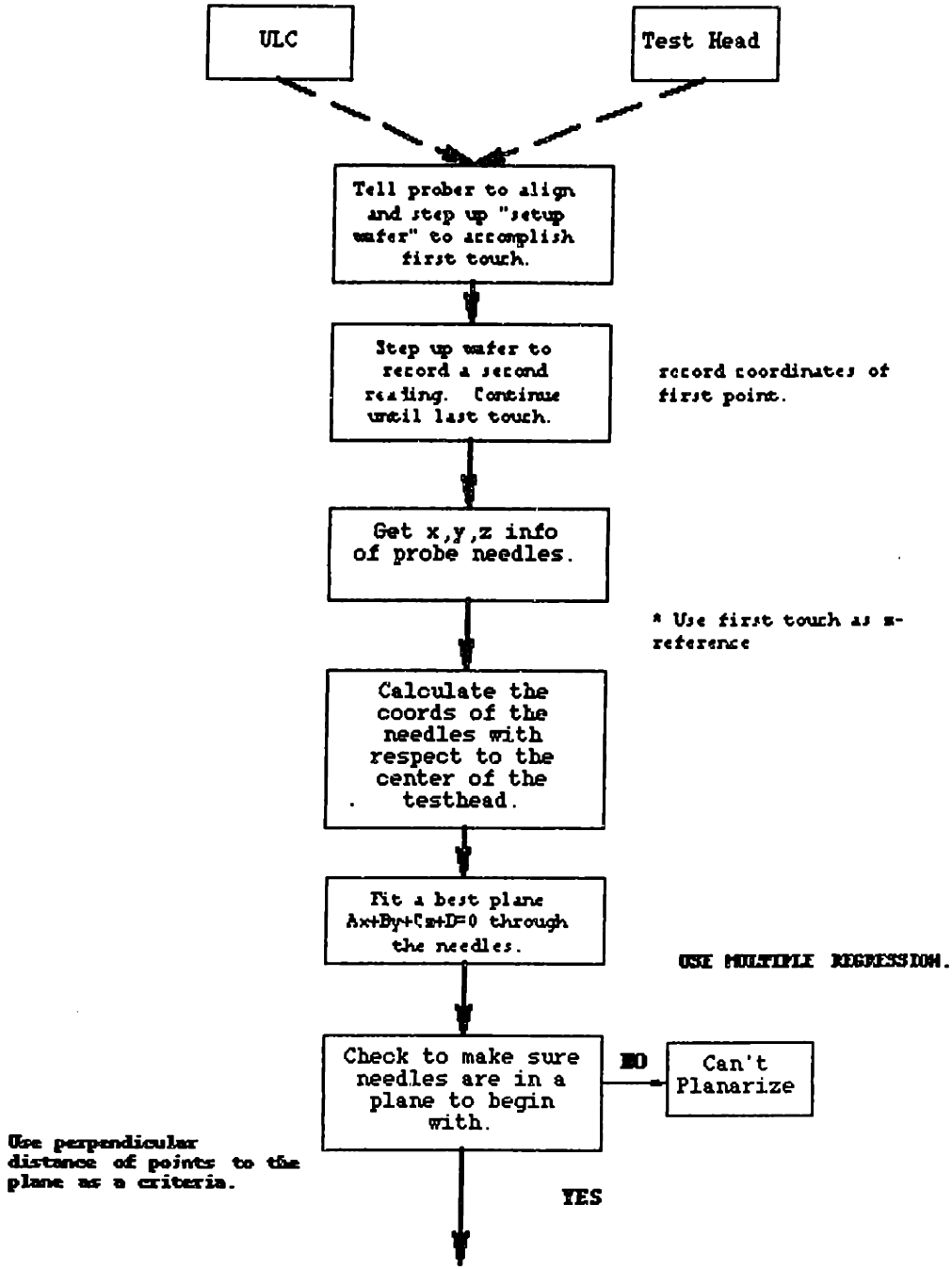
Figure 9.11: Alpha Prototype Autoplanarization Modifications

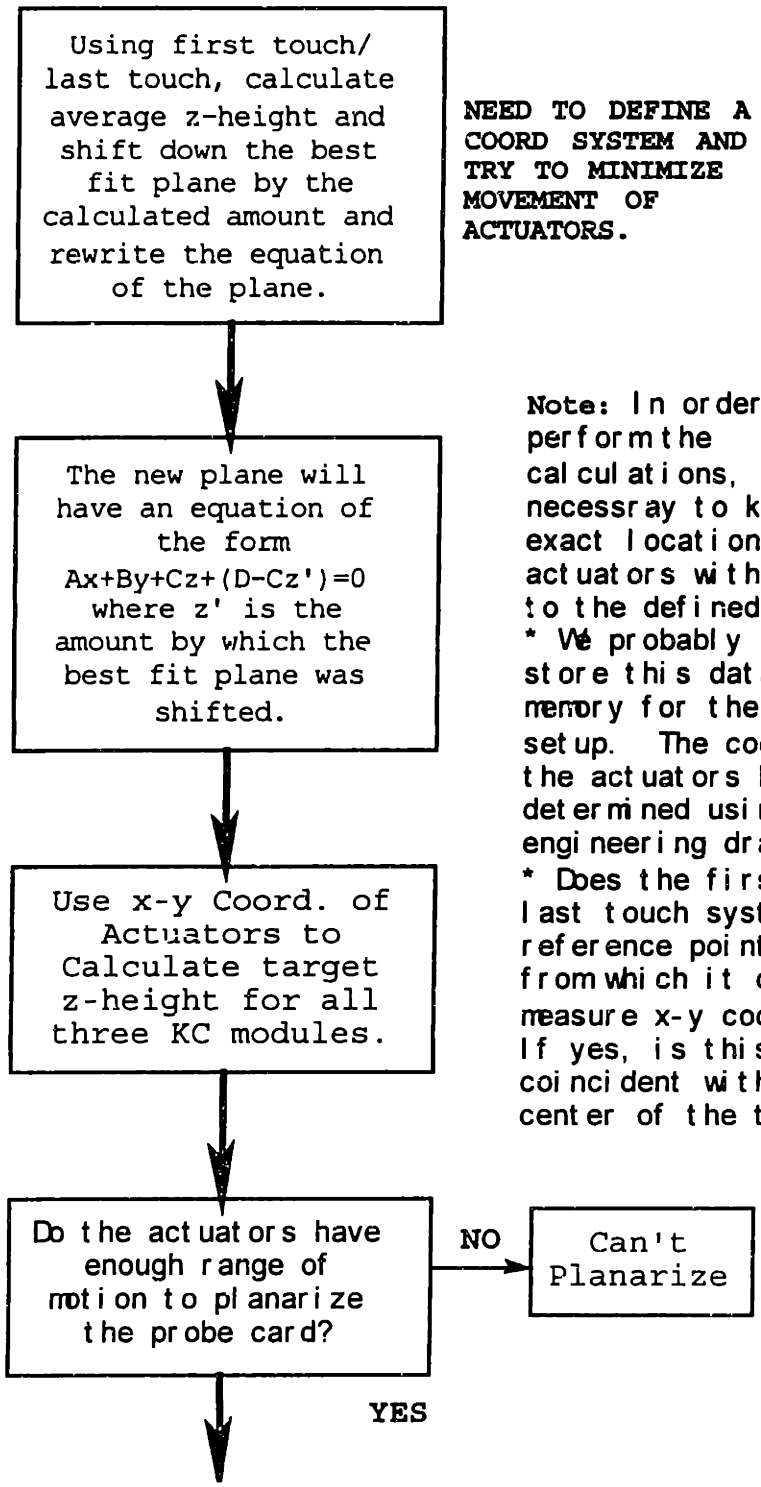
Due to space constraints, an actual implementation of autoplanarization does not allow the motor assembly to protrude beyond the edges of the mount plate. The autoplanarization algorithm implemented on the spacestation was utilized and proved to be successful.

9.2.6 Future Considerations

Implementation of the autoplanarization feature on the kinematic interface requires far more sophisticated software than the PC-based data acquisition and motion control system used to evaluate the concept. In order to implement autoplanarization on customer test floors, the tester must be able to exchange data with the probe. For example, in cases where the probe card might be mounted on the testhead, the x,y, and z coordinates of the probe needles are required for autoplanarization. This information can be acquired from the prober. This exchange of information can be accomplished by software developed in joint effort between the tester and the prober companies. The charts shown in Figure 9.12 shown on the next three pages demonstrate one approach to this task.

Software Requirements for Planarization Using the KC Module




Using first touch/ last touch, calculate average z-height and shift down the best fit plane by the calculated amount and rewrite the equation of the plane.
NEED TO DEFINE A COORD SYSTEM AND TRY TO MINIMIZE MOVEMENT OF ACTUATORS.
The new plane will have an equation of the form Ax+By+Cz+(D-Cz')=0 where z' is the amount by which the best fit plane was shifted.
Note: In order to perform the calculations, it is necessray to know the exact location of the actuators with respect to the defined origin.
* We probably have to store this data in memory for the specific setup. The coords of the actuators have to be determined using engineering drawings.
* Does the first touch/ last touch system have a reference point/origin from which it can measure x-y coordinates? If yes, is this origin coincident with the center of the test head?
Use x-y Coord. of Actuators to Calculate target z-height for all three KC modules.
Do the actuators have enough range of motion to planarize the probe card?
NO
Can't Planarize
YES

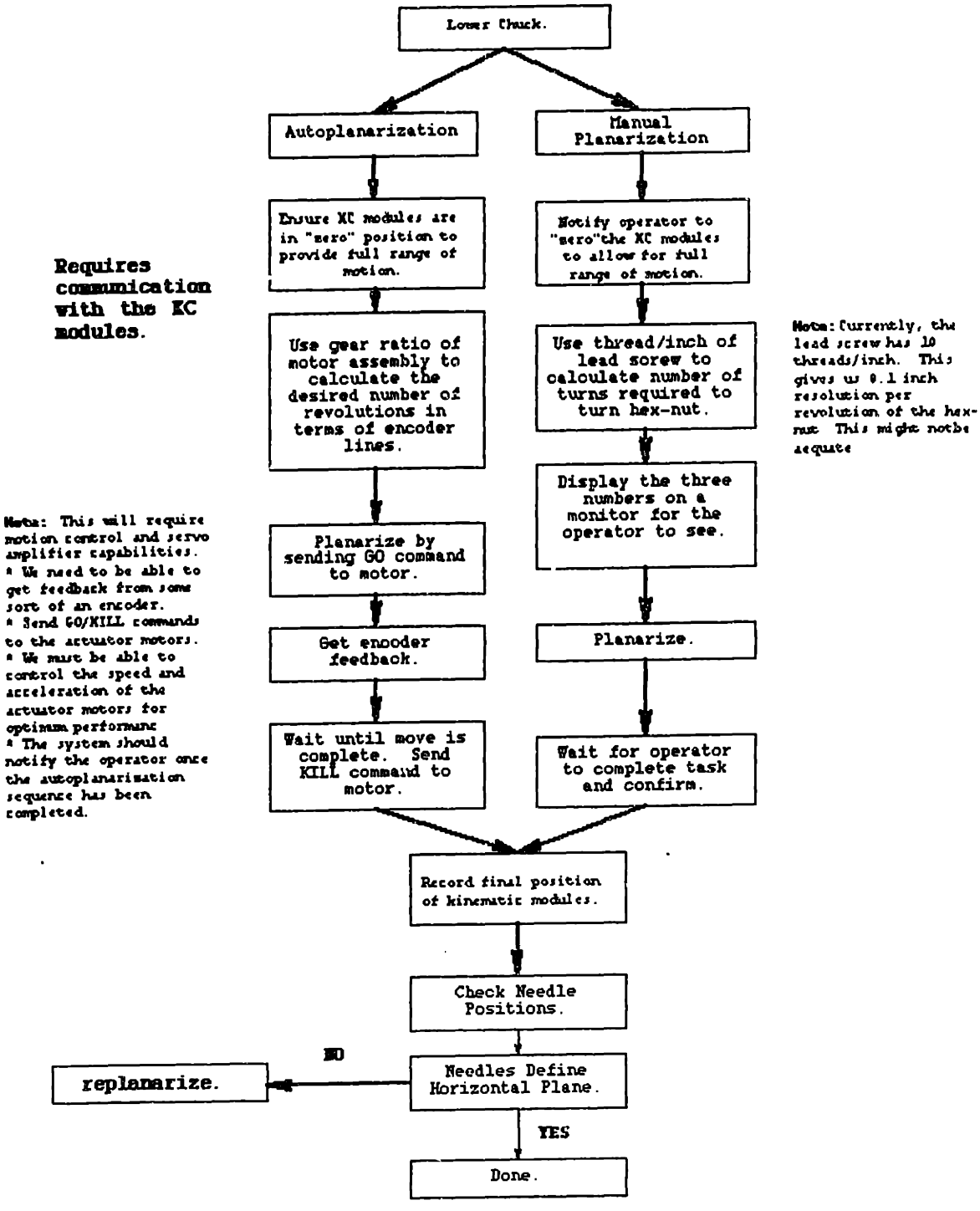


Figure 9.12: Autoplanarization Process Flowchart

In addition to planarization adjustment, the position of the V-grooves can also be adjusted to accommodate planar positioning errors. Instead of using an setup fixture, an adjustable V-groove in the x-y plane could be used to allow for precise alignment of pogo pins with the DIB bond pads. The adjustment mechanism would allow the

operator to move the V-groove perpendicular to the direction of the groove. Since the balls can slide in the grooves, the testhead can be translated and rotated as necessary for proper alignment of the pogo pins with the bond pads. The schematics shown in Figure 9.13 derive the effective translation and rotation which result by adjusting the V-grooves. L is the amount of desired translation and ζ is the desired angle of rotation. The angle between the y-axis and the adjustment axis (direction perpendicular to the V-groove) is denoted by θ. The grooves must be adjusted by an amount τ+G to obtain proper alignment.

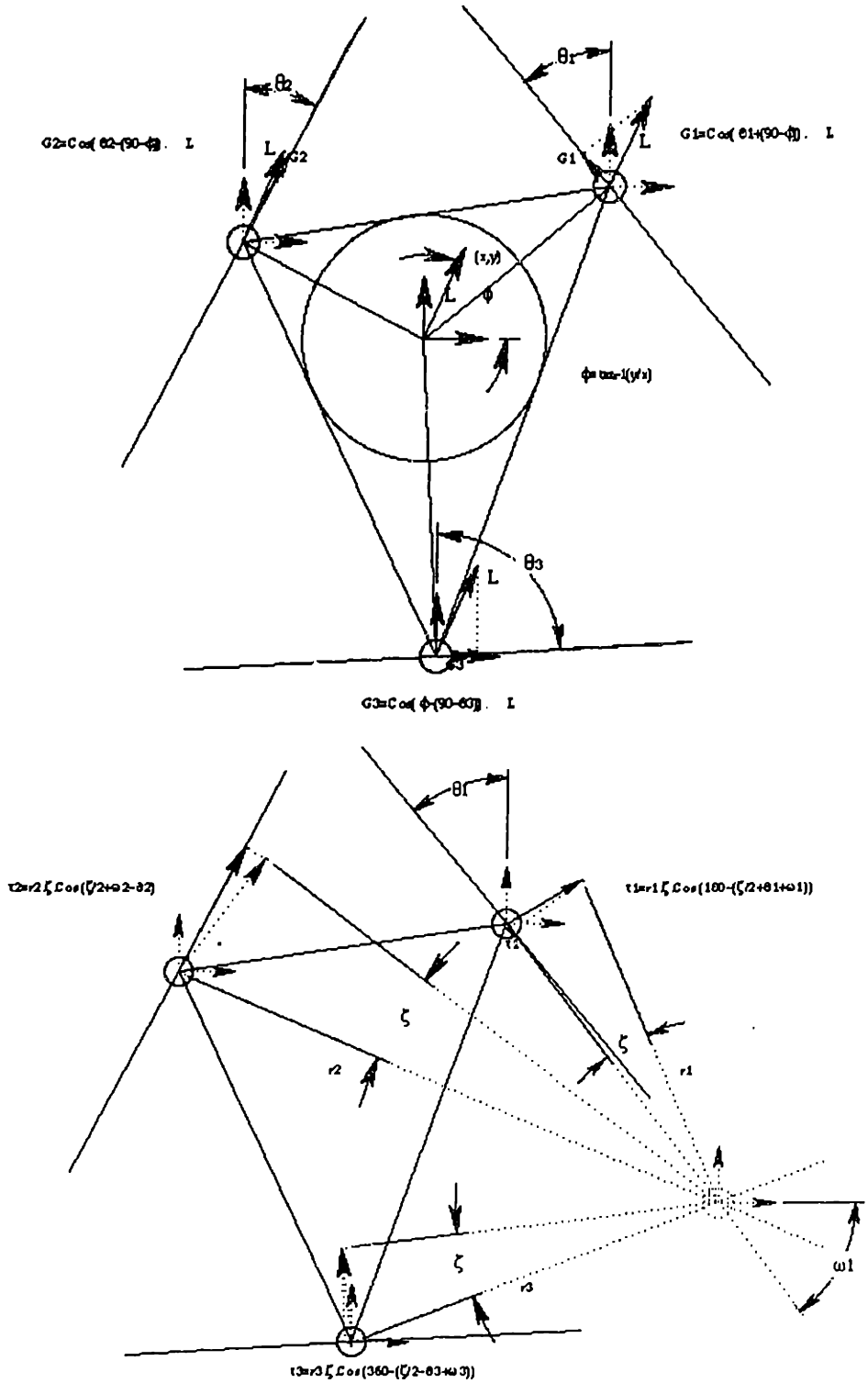


Figure 9.13: Groove Adjustment Geometry

9.3 Glossary

ADC	Analog to Digital Converter
ASIC	Application Specific Integrated Circuit
ATE	Automatic Test Equipment
BGA	Ball Grid Array
BIST	Built-In Self Test
CMOS	Complimentary Metal-Oxide Semiconductor
CSP	Chip Scale Package
DAC	Digital to Analog Converter
DC	Direct Current
DFM	Design for Manufacte
DFT	Design for Test
DIB	Device Interface Board
DIP	Dual Inline Package
DUT	Device Under Test
EM	ElectroMagnetic
GHz	GigaHertz (10^9 cycles/second)
HIB	Handler Interface Board
HSD	High-Speed Digital
KDS	Kinematic Docking System
KGD	Known Good Die
MCM	Multi-Chip Module
MHz	MegaHertz (10^6 cycles/second)
PGA	Pin Grid Array
PIB	Probe Interface Board
PLC	Programmable Logic Controller
QFP	Quad Flat Pack
SIA	Semiconductor Industry Association
SMT	Surface Mount
SOC	System on a Chip
SOP	Small Outline Package
TEM	Transverse ElectroMagnetic
TSOP	Tape Small Outline Package
TQM	Total Quality Management
VLSI	Very Large Scale Integration
ZIF	Zero-Insertion Force

9.4 Bibliography

Alagheband, Ali. The Engineering Case Study as a Basis for the Master of Science Thesis." Master's Thesis. Massachusetts Institute of Technology. 1996.

Alden, John. "The Development and Characterization of Kinematic Interfaces for the Semiconductor Industry." Master's Thesis. Massachusetts Institute of Technology. 1997.

Atwater, H. A. Introduction to Microwave Theory. Malabar, FL: Robert E. Krieger Publishing, 1981.

Baglia, John. "Final Packaging: Onshore or Offshore?" Semiconductor International. April, 1997: 61-64.

Berins, Michael L, ed. Plastics Engineering Handbook. Society of the Plastics Industry. New York, NY: Chapman & Hall, 1991.

Bogatin, Eric. Roadmaps of Packaging Technology. Scottsdale, AZ: Integrated Circuit Engineering, 1997.

Buck, Thomas. "High Speed Digital Interconnection Seminar", Islip, NY: Advanced Interconnection Technology, 1997.

Canright, Robert E. James C. Chamblin, Willam L. Pattison, and James A. Root., "Packaging of High-Speed Digital Electronic Systems." Electronics Packaging and Interconnection Handbook. Ed. Charles Harper. New York, NY: McGraw Hill, 1991.

Chiu, Michael A., Godfrey, Jim, and Tivnan, John. *Characterizing the Teradyne A50/KLA 1201 Test Cell.* Teradyne Users Group Proceedings, 1996

Choma, J, Jr. "Signal Flow Analysis of Feedback Networks." IEEE Transactions on Circuits and Systems. Vol. 37, No. 4 (1990): 455-463.

Clayton, Antony P. ed. Handbook of Electronic Connectors. Ayr, Scotland: Electrochemical Publications Limited, 1982.

Collin, Robert E. Foundations for Microwave Engineering. New York, NY: McGraw-Hill, 1966.

Corwith, Bill. Test Fixture Design Techniques to Maximize Signal Integrity. Application note #102. Gilbert, AZ: Fresh Test Technology, 1993.

Cutherell, David. Product Architecture, The PDMA Handbook of New Product Development, New York, NY: John Wiley and Sons, 1996. 217-236.

Davis, Brendon. Economics of Automatic Testing. New York, NY: McGraw-Hill, 1982.

Elliott, David. Integrated Circuit Fabrication Technology. New York, NY: McGraw-Hill, 1982.

Elliott, David. Microlithography: Process Technology for IC Fabrication. New York, NY: McGraw-Hill, 1986.

Foster, R. "The S-curve: A New Forecasting Tool." Innovation, The Attacker's Advantage. Summit Books, New York, NY:.Simon and Schuster, 88-111.

Genco, Robert M., Jack Pitts, and Gary Gallager. "Control Microcontaminants with Wafer Cassette Purging." Semiconductor International April, 1997: 91-94.

Goldberg, Joel M. "The Future is Self Test." Test & Measurement World. February, 1996: 22-26.

Grover, F.W. Inductance Calculations. Princeton, NJ: Van Nostrand, 1946.

Harper, Charles A, ed. Electronics Packaging and Interconnection Handbook. New York, NY: McGraw-Hill, 1991.

Hawkins, G. A, and Max Jakob. Elements of Heat Transfer. New York, NY: John Wiley and Sons, 1957.

Hutchenson, G. Dan, and Jerry D. Hutchenson. "Technology and Economics in the Semiconductor Industry." Scientific American January, 1996: 54-62.

Hocken, Robert, et al. "Three Dimensional Metrology." Annals of the CIRP. Vol. 26, No. 2. (1977): 402-408.

Horowitz, Paul, and Winfred Hill. The Art of Electronics. 2nd ed. Cambridge, England: Cambridge University Press, 1993.

Itoh, Tatsuo, ed. Numerical Techniques for Microwave and Millimeter wave Passive Structure. New York, NY: John Wiley & Sons, 1989.

Johnson, Howard J, Graham, Martin. High-Speed Digital Design: A Handbook of Black Magic, Englewood Cliffs, NJ: Prentice Hall, 1993.

Lienhard, John H. A Survey of Thermal Design Practices for Cooling PC Board Electronics. Cambridge, MA: Study sponsored by Teradyne, Inc, 1995.

Mai, Phill. "JEM Vertical Contact Probe Card." IEEE Southwest Test Conference. San Diego, CA, April 1997.

Mansingh, Vivek. Real-time solutions for your thermal problems. National Electronics Cooling Course and Workshop. 1993.

Medley, Max. Microwave and RF Circuits: Analysis, Synthesis and Design. Boston, MA: Artech House, 1993.

Milo, Paul. "Why Bother to Test: Editorial." EE-Evaluation Engineering. March, 1996: 8.

Minnis, Brian J. Designing Microwave Circuits by Exact Synthesis. Boston, MA: Artech House, 1996.

Ogata, Katsuhiko. Modern Control Engineering. 2nd ed. Englewood Cliffs, NJ: Prentice-Hall, 1990.

Owyang, Gilbert H. Foundations for Microwave Circuits. New York, NY: Springer-Verlag, 1989.

Rao, Navin. "ASICs at the Crossroads: Systems on a Chip." Portable Design July, 1997: 49-51.

Rosenau, Milton D., Griffin, Abbie, Castellion, George A., and Anschuetz, Ned F. The PDMA Handbook of New Product Development. New York, NY: John Wiley & Sons, 1996.

Seeger, John A. Microwave Theory, Components, and Devices. Englewood Cliffs, NJ: Prentice-Hall, 1986.

Semiconductor Industry Association (SIA). The National Technology Roadmap for Semiconductors. San Jose, CA: Semiconductor Industry Association, 1994.

Semiconductor Industry Association (SIA). The National Technology Roadmap for Semiconductors. Preliminary Copy. San Jose, CA: Semiconductor Industry Association, 1997.

Shen, Liang Chi, and Jin Au Kong. Applied Electromagnetism. Monterey, CA: Brooks/Cole Engineering, 1983.

Singer, Peter. "1996: A New Focus on Equipment Effectiveness." Semiconductor International January 1996: 70-74.

Singer, Peter. "This is your 300 mm Wake Up Call." Semiconductor International January 1997: 15.

Skilling, Hugh Hildreth. Fundamentals of Electric Waves. New York, NY: John Wiley & Sons, 1942.

Slocum, Alex. Precision Machine Design. Englewood Cliffs, NJ: Prentice Hall, 1992.

Slocum, Alex. "Design of Three-Groove Kinematic Couplings." Precision Engineering: Journal of the ASPE January, 1992.

Slocum, Alex. "Kinematic Couplings for Precision Fixturing – Part I: Formulation of Design Parameters." Precision Engineering: Journal of the ASPE Vol. 10, No. 2. April 1988: 85-91.

Slocum, Alex, Donmez, Alkan. "Kinematic Couplings for Precision Fixturing – Part II: Experimental Determiniation of Repeatability and Stiffness." Precision Engineering: Journal of the ASPE Vol. 10, No. 3. July 1988: 115-122.

Teradyne, Inc. Annual Report. Boston, MA: Teradyne, Inc., 1996.

Terman, Frederick Emmons. Electronic and Radio Engineering. 4th ed. New York, NY: McGraw-Hill, 1955.

Timoshenko, S., and S. Woinowsky-Krieger. Theory of Plates and Shells. New York, NY: McGraw-Hill, 1959.

Tipler, Paul. Physics. New York, NY: Worth Publishers, 1982.

Valkenburg, Mac E, ed. Reference Data for Engineers, Radio, Electronics, Computer and Communications. Carmel, IN: SAMS Prentice Hall, 1993.

Van Doren, Matthew. "Precision Machine Design for the Semiconductor Equipment Manufacturing Industry." Dissertation. Massachusetts Institute of Technology. 1995.

von Hippel, Eric. The Sources of Innovation. New York, NY: Oxford University Press, 1988.

Wadell, Brian. Transmission Line Design Handbook. Boston, MA: Artech House, 1991.

Walsh, Patrick. "ASICs at the Crossroads: Systems-on-a-Chip." Portable Design Magazine July 1997

Weck, Manfred. Handbook of Machine Tools: Volume 4, Metrological Analysis and Performance Tests. Trans. John Wiley and Sons, 1984. (translated from German)

Wolfram, Stephen. Mathematica: A System for Doing Mathematics by Computer. Redwood City, CA: Addison-Wesley, 1988.

9.5 Vita

Michael A. Chiu was born on April 9, 1968 in La Crosse, Wisconsin. He graduated from La Crosse Central High School in May of 1986. Michael began his college studies in architecture at The University of Minnesota in Minneapolis. He eventually switched to Mechanical Engineering and received his Bachelor's of Science, Summa Cum Laude, in 1991. He continued his studies at the Massachusetts Institute of Technology where he received a Master's of Science, also in Mechanical Engineering, while working in MIT's Precision Machine Design Lab.

Michael has held various coop and intern engineering positions at Ford Motor Company, Pacific Northwest Labs and the National Institute of Standards and Technology. He currently is working at Teradyne, in Boston, Massachusetts and is building a woodworking shop designed to make even his advisor jealous.

Address: 17 Kidder Avenue
Somerville, MA 02144

This thesis was written in Microsoft Word using the Times New Roman font.