UNIVERSITY OF ADELAIDE
ELECTRICAL ENGINEERING DEPARTMENT

MARCH 1971

MAN-COMPUTER GRAPHICS: CURRENT AND NEW HARDWARE IMPLEMENTATIONS

BY V. C. SOBOLEWSKI, B. E. (Hons.)

VOLUME I

INTRODUCTION AND APPENDICES

A THESIS SUBMITTED TO THE UNIVERSITY OF ADELAIDE
FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

AUTHOR'S NOTE

The length of this Thesis necessitated the work to be bound in two volumes. Volume I contains the Introductory chapters and the Appendices while Volume II discusses the proposed Interactive Graphics Console, "VIDIO-GRAPHIC", and examines its feasability, and implementation.

PREFACE

Several years ago, Ivan Sutherland, originator of the "Sketchpad" program, a milestone in Graphical Communication and Man-computer Interaction, listed "ten unsolved problems" confronting workers in the field of Computer-Graphics, of which

> "...the first and biggest unsolved problem...is to build a low-cost display device suitable for on-line graphical use" (1).

Among the requirements of such an on-line graphics unit, he stated,are:

- geometrical fidelity and display precision
- short or "instant" response to any user modification
- direct user-input capability via a "pen" or "stylus"

including - "pointing ability", by which certain features of the graphic display may be indicated to the computer for further consideration, by using the "stylus" or "pen".

The project,of which this thesis is the result, originally had as its topic a novel approach to Computer-aided-circuit-design, specifically the possibility of

using graphical methods for solving circuits with linear and non-linear elements. Immediately the problem mentioned by Sutherland arose. Without a "display device" with the aformentioned features, the project would at most have been a paper and pencil exercise.
The real dilemma was economics.
Existing commercial units with some of the above features had a starting price of around $20,000 - $30,000.
It is not surprising then that the topic of the project was changed, namely to develop a low-cost Interactive Graphic Unit with (hopefully) all of the above features.
To use existing approaches and techniques would have been merely to duplicate existing efforts and to duplicate (if not increase) existing costs. A wholly new and radical approach had to be taken.

The philosophy of the approach was to use existing components and devices and interconnect them (electrically and optically), into a new system, rather than to attempt to develop a new system from fundamentals (again ecomomics would preclude the latter approach).

The proposed Interactive Graphics Console is called "VIDIOGRAPHIC" (Video Integrated Display and Input Optical Graphic Interactive Console), the keywords being "Video" and "Optical", as TV techniques along with optical techniques are used to implement Input, Display, and Display Refresh Storage.
The major portion of this study was, quite naturally, performing feasibility studies of the proposed system; and since new concepts using existing devices were used, fundamental work in widely varying areas was done. System requirements had to be verified for satisfactory perform-

ance. Consequently a completely finished usable Graphics System was out of the question for a project such as a Ph.D. Thesis.

What resulted was very encouraging and rewarding. That a low cost Graphics System, based on a wholly new principle is feasible, is clearly evident. What is more evident is that more work needs to be done to implement this feasibility study into a useful system, or even a commercial system. This work must be left to others.

The apparent great length of the thesis is unavoidable. Since the concepts on which VIDIOGRAPHIC is based are on principles which by themselves may not be wholly new, but which when integrated to form the system, are new, a full description of these principles and their interrelation with each other, was necessary; often, since some principles or existing systemswere used for which they were not originally designed for, some trivial and some not so obvious explanations and calculations had to be done to demonstrate the feasibility of that feature in "VIDIOGRAPHIC". Thus rather than concentrate in some detail on some aspect of a hypothetical Interactive Graphics Console, the Graphics Console itself had to be examined in detail to show its feasability. Consequently paper calculations, at the expense of experimental realization, form the main body of the thesis.

The experimental work supporting it are those parts of the feasability study which could not be proved or "demonstrated" on paper, namely the very important requirement of being able to precisely measure the distortion of the Display and Graphics Input Subsystems of "VIDIOGRAPHIC". Other experimental work describes certain very linear hybrid circuits required for the implementation of "VIDIOGRAPHIC".

The results presented here, both on paper and with hardware, indicate that a truly economic (< $2000) Interactive Graphics Console is feasable.

The other major portion of this thesis is a short review of the development and capabilities of Man-Computer-Graphics (MCG), by which is meant the solution of diverse problems jointly by a user and a computer, with the mutual communication and interaction during the problem-solving phase being in some graphical form. An overview of the technology of implementing MCG is also given, by describing the requirements and implementation of the Interactive Graphics Console (IGC), which is the link between the user and the central processor where the problem is being solved.
The aim of these two chapters is to indicate where the thesis work fits into the state-of-the-art of MCG and why the realization of an economic IGC is necessary. The requirements and current technology of the IGC is given to indicate requirements for the IGC proposed here, and to avoid duplicating any previous work.

PLAN OF THESIS

The Thesis is divided into four parts:

(i) Part 1 (Chapters 1-2) introduces the field of Man-Computer Graphics (MCG) and indicates the main problem areas in this field.

(ii) Part 2 (Chapters 3-10), forming the main body of the Thesis, describes in detail the theory and feasability of the proposed Interactive Graphics Console "VIDIOGRAPHIC".

(iii) Part 3 (Chapter 11) describes some of the experimental work and circuits required to implement "VIDIOGRAPHIC".

(iv) A set of Appendices, in which expressions and concepts used in the above chapters are derived and explained, conclude the Thesis.

Specifically the contents of the chapters are are as follows:

Chapter 1 introduces the topic of Man-Computer Graphics, indicates the area of applications and the main problem areas. Appendix 1 tabularly indicates the wide range of applications of Man-Computer Graphics.

Chapter 2 describes the unit by which a user interacts with the computer during Man-Computer Graphics operation; this is the "Interactive Graphics Console" (IGC). Its requirements and performance parameters are listed and compared with existing implementations (Appendices 3-5). Main Computer - IGC interaction, as well as the software required for MGC, are briefly touched upon.

Chapter 3 introduces the proposed IGC, "VIDIOGRAPHIC", based on the novel idea of display refresh storage and user-input of graphics being implemented by optical means. The operation of "VIDIOGRAPHIC", its development, the subsystems required and the requirements for its feasability are given. Briefly the system consists of a Display Subsystem (a CRT), an Graphics Input Subsystem (a Vidicon and Light-Emitting Pen), a Storage Refresh Subsystem (the interconnected CRT-Vidicon), and a Schmidt Optical System (from existing domestic TV-Projection receivers) to implement graphics input and display refresh storage.

Chapter 4 examines the first of the major requirements for VIDIOGRAPHIC,which is that adequate luminous flux be generated within the system (at the CRT screen) to be visible by the user on a viewing screen,and to be incident on the Vidicon faceplate to generate adequate output signal for maintenance of display refresh.

Chapter 5 examines the readout process from the Vidicon,as it is usually considered that Vidicons give uneven and poor quality output signals.

Chapter 6 examines the required optical system (the Schmidt Optical System from TV-Projection receivers), which enables simultaneous CRT screen viewing,and the imaging of this onto the Vidicon faceplate, and also enables user-input of graphics with a light-emitting pen. The source for the pen is obtained from the CRT screen itself.

Chapter 7 examines the second main requirement for "VIDIOGRAPHIC", which is the ensuring of a steady display (otherwise the display would move out of the user's view within seconds, leaving a blank viewing area). This requires display points being relocated or "position corrected" each refresh cycle. The correction information to perform this location"position-correction"is obtained from Moire Patterns measurements, and then is encoded graphically in certain areas of the CRT scanned area. This graphic information is thus available each display-scan interval.

Chapter 8 describes the graphic correction decodeing circuits and the correction-implementation circuits required to result in a linear (better than 0·25%) CRT and Vidicon.

Chapter 9 examines the primary causes of CRT and Vidicon display distortion, namely pincushion and barrel distortion, derives expressions for these and the means of correcting them. The circuits in Chapter 8 corrected "remanent distortion", i.e. distortion after pincushion or barrel distortion have been minimized or eliminated.

Chapter 10 concludes the feasability study of VIDIOGRAPHIC with a general overview of the system, advantages and disadvantages and the expected cost of the system. A comparison is made between the expected performance parameters of VIDIOGRAPHIC with those required of an ideal IGC as enumerated in Chapter 2.

Chapter 11 gives the results of experimental work done towards realizing "VIDIOGRAPHIC" and the various circuits of high linearity required to implement the scanning and the distortion correction circuits.

Much of the above work is original.

The concept of "VIDIOGRAPHIC" is new , as is obviously the means of implementing it, specifically the methods of display distortion measurements (Moire patterns), the encoding of this distortion graphically as distortion correction information, and the associated means of implementing the correction.

The derivation of expressions for pincushion and barrel distortion, and defocussing in both the CRT and the Vidicon have not been seen anywhere (other than in the most approximate forms), although possibly they have been derived elsewhere, as, for example, "anti-pin-cushioning" circuits are currently available commercially.

The light emitting pen and its CRT illuminating source is new in concept.

Expressions for the photocurrent buildup and decay time contents in the Vidicon have been derived and are original.

Phosphor buildup effects requiring derivation cannot be found in the literature and are thus considered original.

The Vidicon beam discharge effects of Chapter 5, have been extended in parts, on existing work; that whole chapter (Chapter 5) forms a comprehensive resume of all possible Vidicon signal output defects.

Finally, Chapter 2, though not original in content, at least performs the function of a fairly comprehensive review of existing hardware implementation of requirements and techniques associated with Man-Computer Graphics, as a comprehensive review such as this is lacking in the literature.

A note about notation. Since the thesis ranges over a wide range of topics, consequently many quantities, units etc. are mentioned in short separate sections; the definitions and notation for these is kept within the relevent section, as it is felt that a lengthy table of notation at the beginning of the thesis is not warranted under these conditions.

. . . .

TABLE OF CONTENTS

TITLE PAGE
PREFACE including PLAN OF THESIS i
TABLE OF CONTENTS ix
SUMMARY xxxii
DECLARATION OF ORIGINALITY xxxiv
ACKNOWLEDGEMENTS xxxv

PART 1 – MAN-COMPUTER GRAPHICS : GENERAL CONCEPTS 1

CHAPTER 1. MAN-COMPUTER GRAPHICS

1.1 INTRODUCTION - THE ROLE OF MAN-COMPUTER GRAPHICS. . . .1

1.2 SHORT RESUME OF DEVELOPMENT OF MAN-COMPUTER GRAPHICS 6

1.2.1 Early History 6

1.2.2 "The Applied Approach" Design Augmented by Computer : DAC-1. 8

1.2.3 "The Fundamental Approach"-Project MAC , SKETCH-PAD etc 10

1.2.3.1 Introduction 10
1.2.3.2 Generality 11
1.2.3.3 Form and Language Describing Graphical Information. 12
1.2.3.4 Hardware. 13
1.2.3.5 Implementation - SKETCHPAD I 14
1.2.3.6 Later Developments at MIT. 15

1.2.4 Other Approaches to Man-Computer Graphics . .15

1.2.5 Summary 16

1.2.6 Development of MCG. 17

1.3 MAN- COMPUTER-CIRCUIT-DESIGN 18

1.3.1 Introduction. 18

1.3.2 Available Computer-Aided-Circuit-Design Programs. 20

1.3.3 Man-Computer Graphics Circuit Design (MCGCD) .21

1.3.3.1 Introduction 21
1.3.3.2 Characteristics of GINA. 22
1.3.3.3 Advantages and Disadvantages of GINA 23
1.3.3.4 Other MCGCD Programs 24

1.3.4 Graphical Solution of Circuits. 25

1.3.4.1 Problems of MCGCD. 25
1.3.4.2 Device Modelling 26
1.3.4.3 Graphic Solutions of Circuits. 27

1.4. CONCLUSIONS AND COMMENTS 29
Subscript 30

CHAPTER 2. THE INTERACTIVE GRAPHICS CONSOLE (IGC): REQUIREMENTS AND EXISTING IMPLEMENTATION 32

2.1 INTRODUCTION 32
2.1.1 General. 32
2.1.2 Interactive Graphics Console Functions . . . 35

2.2. USER-DISPLAY INTERFACE 36

2.2.1 Requirements 36
2.2.2 Display Parameters 37

2.2.2.1 Resolution. 38
2.2.2.2 Brightness and Contrast 39
2.2.2.3 Flicker and Display Refresh 40
2.2.2.4 Display Response Time 42
2.2.2.5 Display Area Size and Placement in Relation to User 44
2.2.2.6 Display Addressability and Repeatability. 46
2.2.2.7 Geometrical Fidelity of Display - Distortion. 48
2.2.2.8 Readability 50
2.2.2.9 Halftones and Colour. 51
2.2.2.10 Hard-Copy Capability. 52
2.2.2.11 Optional Display Device Features. . . . 53

2.3 DISPLAY GENERATION 56
2.3.1 Introduction 56
2.3.2. Data Positioning 56

2.3.2.1 Random Positioning 56
2.3.2.2 Raster-scan Positioning. 57
2.3.2.3 Typewriter or Textual-scan Positioning. 59

2.3.3. Display Generation 59

2.3.3.1 General. 59
2.3.3.2 Point Plotting Mode. 59
2.3.3.3 Stroke-Writing Mode. 60
2.3.3.4 Ramp-Writing Mode. 61
2.3.3.5 "Stencil" or "Shaped-beam" Writing Mode 63

2.3.4 Alphanumeric Generation 64

2.3.4.1. General 64
2.3.4.2 Raster Scanning Character Generation 65
2.3.4.3 Function-generated Alphanumeric Generation 66
2.3.4.4. Comparisons 66

2.3.5 Vector or Line Generation 67

2.3.5.1 General. 67
2.3.5.2 Stroke Vector Generation 67
2.3.5.3 Ramp-mode Vector Generation 68
2.3.5.4 Raster Scan Vector Generation 68
2.3.5.5 Storage Properties of Displays 69

2.3.6 Other Graphics Operations 69

2.3.6.1 General. 69
2.3.6.2 Pointing or "Picking". 70
2.3.6.3 Display Deletion or Erasure 71
2.3.6.4 Rotation 72
2.3.6.5 Scaling 73
2.3.6.6 Translation. 73
2.3.6.7 Replication. 74
2.3.6.8 Comments 74

2.3.7 Forms and Types of Available Displays . . . 74

2.3.7.1 General 74
2.3.7.2 Cathodoluminescent Displays - CRTS . 75
2.3.7.3 Electroluminescent Displays - E-L Panel Displays 76
2.3.7.4 Gas Electroluminescence - Ionization-Plasma Display Panels. 77
2.3.7.5 Injection Luminescence - Light-emitting Diodes (LED Arrays) 79
2.3.7.6 Other Forms of Displays 79

2.4. USER - INPUT INTERFACE 80

2.4.1 General. 80

2.4.2 Requirements of a User-Input Interface . . . 80

2.4.3 Input Interface Parameters 82

2.4.3.1 General 82
2.4.3.2 Resolution 84
2.4.3.3 Geometrical Linearity and Accuracy. . 84
2.4.3.4 Repeatability or Positioning Precision 85
2.4.3.5 Input Area Size 86
2.4.3.6 Input Response Time 86
2.4.3.7 Hard-Copy Capability. 87

2.4.4 Forms and Types of Available Input Devices . 87
2.4.4.1 Teletype Keyboards and Function Keyboards 87
2.4.4.2 Two-dimensional Graphic Input Pads or Tablets. 88
2.4.4.3 Light Pens. 89
2.4.4.4. Cursors, Caret or Tracking-cross Input Devices 90
2.4.4.5 Electro-mechanical x-y Digitizers . . 92
2.4.4.6 Photo-tube Input Devices. 92
2.4.4.7 Hard-Copy Scanners. 93

2.5 GRAPHIC INFORMATION and DISPLAY STORAGE 93

2.5.1 General. 93

2.5.2 Random Access Memories (RAMs) 96

2.5.3 Serial Access Memories (SAMs) or Cyclic Memories 99

2.5.3.1 General 99
2.5.3.2 Simple SAMs 99
2.5.3.3 Shift Register SAMs 100
2.5.3.4 Semi-serial SAMs. 100

2.5.4 Direct View Storage Tubes (DVSTs). 101

2.5.4.1 Storage Tubes 101
2.5.4.2 Other Analog Storage Tubes. 103

2.5.5 Miscellaneous Image Storage 104

2.6 ICC SYSTEM CONFIGURATION 105

2.6.1 General IGC Configuration. 105

2.6.2 IGC-CPU Configurations 106

2.6.2.1 General. 106
2.6.2.2 Stand-alone or Single IGC Configurations 108
2.6.2.3 Multi-station Configurations 108
2.6.2.4 Peripheral or Interface Computer Configurations 109
2.6.2.5 Multi-station Multi-CPU Configuration 110
2.6.3 Analog Graphics Systems- ARTRIX 110

2.7 HARDWARE- SOFTWARE TRADEOFFS 110
2.7.1 General 111
2.7.2 INTERGRAPHIC - Efficient Tradeoffs. 112
2.8 GRAPHICS SOFTWARE 113
2.8.1 General 113
2.8.1.1 Software Organization 113
2.8.1.2 Graphic Data Structure. 116
2.8.2 General Purpose Graphics Language 117
2.8.2.1 General. 117
2.8.2.2 Requirements of a General Purpose Language 117
2.8.3 Display Interface Software 120
2.8.3.1 General. 120
2.8.3.2 Control Commands 121
2.8.3.3 Vector Generation Commands 122
2.8.3.4 Alphanumeric or Text Commands . . . 123
2.8.3.5 Instruction Commands 123
2.8.4 User-Interaction Software 123
2.8.4.1 General 123
2.8.4.2 Interactive Languages 125
2.8.4.3 Existing Software Systems 127
2.9 OTHER GENERAL CONSIDERATIONS. 129
2.10 GENERAL CONCLUSIONS 130

PART 11 - FEASABILITY STUDY OF THE PROPOSED INTERACTIVE GRAPHICS CONSOLE-"VIDIOGRAPHIC" 132

CHAPTER 3. THE PROPOSED INTERACTIVE GRAPHICS CONSOLE: "VIDIOGRAPHIC" 132

3.1 INTRODUCTION 132

3.2. GENESIS OF "VIDIOGRAPHIC". 133
3.2.1 The Display Subsystem 133
3.2.2 The Graphic Input Subsystem 134
3.2.3 The Display Refresh Storage Subsystem . . . 135
3.2.4 The Optics Subsystem. 136
3.2.5 The Storage Process 136
3.2.6 Brief Description of System Operation . . . 137
3.2.6.1 Optical Signals 137
3.2.6.2 Electrical Signals 138
3.2.6.3 Signal Requirements. 139
3.2.7 Signal Level Requirements 141
3.2.7.1 Introduction 141
3.2.7.2 Classes of Electrical Signals. . . . 142
3.2.8 System Features 143
3.2.9 Basic Requirements for Feasability. 144

3.3 BRIEF DESCRIPTION OF SUBSYSTEMS 144
3.3.1 The Cathode Ray Tube 144
3.3.2 The Vidicon TV-Camera 146
3.3.3 The Schmidt Optical System. 149
3.3.4 The Light Pen and Projection Screen 150

3.4 OUTLINE OF FEASABILITY STUDY. 150
3.4.1 Light Energy Transfer and Requirements . . 151
3.4.2 Complete Photoconductor Scanning 151
3.4.3 Measurement and Correction of Display Distortion 151
3.4.4 Efficient Optics System. 151

CHAPTER 4. LIGHT ENERGY REQUIREMENTS IN "VIDIOGRAPHIC" . . . 153

4.1 LUMINANCE OF THE CRT PHOSPHOR 153
4.1.1 Introduction 153
4.1.2 Phosphor Luminance in CRTs 154
4.1.3 Buildup Characteristics of "W" Phosphor. . 157
4.1.4 Spectrum of "W" Phosphor 159
4.1.5 Summary. 160

4.2 ILLUMINATION INCIDENT ON VIDICON PHOTOCONDUCTIVE TARGET. 161

4.3 SIGNAL BUILDUP IN VIDICON AND RESPONSE TO "W" PHOSPHOR ILLUMINATION 163
4.3.1 Introduction. 163
4.3.2 Exponential Rise-and-Decay of Photocurrents 166
4.3.3. Photocurrent Time Constants 168
4.3.4 Photocurrent Buildup. 170
4.3.5 Vidicon Output Signals. 173
4.3.6 Response to User-Specified Inputs . . 176
4.3.7 Summary 178

CHAPTER 5. BEAM DISCHARGE OF VIDICON PHOTOCONDUCTOR 180

5.1 INTRODUCTION 180

5.2 BEAM DISCHARGE EFFECTS 182
5.2.1 Introduction. 182
5.2.2 Beam Impedance. 184
5.2.3 Beam Discharge. 185

5.3 PHOTOCONDUCTOR POTENTIAL STABILIZATION 187
5.3.1 Introduction
5.3.2 Dark Current Stabilizing Potential 187

5.3.3 Signal Buildup of New Display Information. . 188
5.3.4 Signal Lag after Erasure of Display Information 189
5.3.5 Magnitude of "Beam" Discharge or "Capacitive" Lag. 189

5.4 MODIFICATION OF PHOTOCONDUCTOR STABILIZATION POTENTIAL DUE TO DEFLECTION 190

5.5 MODIFICATION OF PHOTOCONDUCTOR STABILIZATION DUE TO ASSYMETRY OF SCANNING. 194
5.5.1 Introduction 194
5.5.2 Beam Deflection Due to Retarding Mesh-Field. 195
5.5.3 Dark Current Assymetry 199
5.5.4 "Edge Ripple" and "Flicker". 200
5.5.5 Scanned Area Edge Effects. 202
5.5.6 Non-linearities in Scanning Waveforms. . . . 204
5.5.7 Thickness Variations in Photoconductor Target 207

5.6 REMARKS 210

5.7 SIGNAL-TO-NOISE IN VIDICONS. 211
5.7.1 Introduction 211
5.7.2 Noise due to illumination. 212
5.7.3 Nett S/N at Output 213
5.7.4 Fault-free operating Interval. 214

5.8 SUMMARY 214

CHAPTER 6. THE OPTICAL SYSTEM IN "VIDIOGRAPHIC": DISPLAY, USER-INPUT AND LIGHT PEN 218

6.1 INTRODUCTION 218

6.2 THE USER-DISPLAY INTERFACE (UDI) 219
6.2.1 Requirements of the UDI. 220
6.2.2 Luminance Requirements for the UDI 220

6.2.3 Transmissive Lens Projection Optics 221

6.2.3.1 Lens Projection Optics System. . . . 221
6.2.3.2 Lens Projection Optical Efficiency . 223
6.2.3.3 Lens Projection Disadvantages. . . . 225

6.2.4 Reflective Mirror TV Projection Optics. . . 225

6.3 SCHMIDT TV PROJECTION OPTICS 227

6.3.1 Introduction 227
6.3.2 Schmidt Optics System Operation 228
6.3.3 Advantages of Schmidt Optics. 230
6.3.4 Disadvantages of Schmidt Optics. 232
6.3.5 Optical Efficiency of Schmidt Optical System 233
6.3.6 Illumination Incident on Vidicon with Schmidt Optics System 235

6.4 The USER-VIEWING DISPLAY SCREEN 236

6.4.1 Introduction 236
6.4.2 Requirement of Display Screen 238
6.4.3 Screen Performance and Transmission Gain. . 239

6.4.3.1 Introduction 239
6.4.3.2 Transmission Gain Coefficiant. . . . 240
6.4.3.3 Actual Screen Transmission Gain. . . 242

6.4.4 RCA Projection Unit Screen 243
6.4.5 Quality of Projection Screen Image due to Schmidt Optics. 245

6.5 The USER-GRAPHICS INPUT SYSTEM 245

6.5.1 Screen Requirements for User Light-Pen Input. . . 245
6.5.2 Optical Image Superposition 247
6.5.3 The CRT Screen as a Reflecting Surface. . . 248
6.5.4 CRT Screen Illumination due to Light Pen. . 249
6.5.5 Screen Reflections. 251

6.5.5.1 Internal Screen Reflections and Loss of Contrast. 251
6.5.5.2 Schmidt Optics Dimensions and Contrast Improvement. 253

6.6 THE LIGHT EMITTING PEN & LIGHT SOURCE. 256
6.6.1 The Required Light Pen User Signals 256
6.6.2 User Actions during Light Pen Operations. . 257
6.6.2.1 Angle of incidence of Pen with Respect to Screen.
6.6.2.2 Hand writing Speed 258
6.6.3 Requirements of the Light Pen 259
6.6.4 Commercially Available Illuminating Sources 261
6.6.5 Advantages of Using CRT Luminance as a Light Source. 262
6.6.6 Fibre Optics. 263
6.6.6.1 Introduction 263
6.6.6.2 Incident Illumination on Optic Fibre Entry Pupil 265
6.6.6.3 Transmission of Luminous Flux in Optic Fibre. 266
6.6.6.4 Output Luminance at Exit Pupil of Optic Fibre. 267
6.6.7 The CRT Illuminating Source 268
6.6.7.1 Introduction 268
6.6.7.2 Method of Affixing Fibres to CRT Screen 269
6.6.7.3 Finite Scanning Speed and Generation of CRT Luminance 269
6.6.7.4 Resultant Luminance Signals from CRT Illumination 272
6.6.7.5 Summary. 273
6.7 THE LIGHT PEN. 274
6.7.1 Introduction - Light Pen Requirements. . .
6.7.2 Construction of the Light Pen 280
6.7.3 The Light Pen Optics 280
6.8 RECAPITULATION . . . .
6.8.1 Recapitulation of Optics Subsystem and User Input Subsystem 282
6.8.2 Recapitulation of Chapters 4-5-6. 284

CHAPTER 7. DISPLAY DISTORTION : EFFECTS IN "VIDIOGRAPHIC", MEASUREMENT AND DISTORTION INFORMATION STORAGE. . 287

7.1 INTRODUCTION 287
7.1.1 Scope of Problem 287
7.1.2 Display Distortion Effects. 287
7.1.3 Permissable Distortion. 289
7.1.4 Causes of Distortion. 290
7.1.5 "Feed-Forward" Distortion Correction. 291

7.2 DISPLAY DISTORTION and EFFECTS 293
7.2.1 Introduction. 293
7.2.2 "Display Location Transfer Curves" of VIDIOGRAPHIC. 294
7.2.3 Modified "Display Location" Transfer Curves 296

7.3 MEASUREMENT of DISPLAY DISTORTION. 298
7.3.1 Display Distortion Measurement Requirements 298
7.3.2 Current Methods of Display Distortion Measurement 299
7.3.3 Distortion Measurement using Moire Patterns 301
7.3.4 CRT & Vidicon Distortion Contour Maps . . . 303

7.4. THEORY of DISTORTION CORRECTION in VIDIOGRAPHIC . 305
7.4.1 Requirements of Display Distortion Correction. 305
7.4.2 Methods of Distortion Correction. 310
7.4.3 Vidicon and CRT Distortion Correction . . . 312
7.4.3.1 CRT Distortion Correction. 312
7.4.3.2 Vidicon Distortion Correction. . . . 314

7.5 STORAGE AND FORM OF DISTORTION CORRECTION INFORMATION. 316
7.5.1 Distortion Correction Information Requirements 316
7.5.2 Standard Methods of Distortion Information Storage 318

7.5.3 Display Distortion Storage by Graphical Means. 320
7.5.4 Requirements of Graphical Information Storage. 321
7.5.5 Horizontal Correction Graphical Storage. . 323
7.5.6 Dimensions of Graphics Storage Area. 325
7.5.7 Vertical Correction Graphical Storage. . . 326
7.5.8 Physical Implementation of Graphical Storage 331
7.5.9 Graphic Storage for Larger Capacity Displays 331

7.6 SUMMARY 332

CHAPTER 8. DISPLAY DISTORTION : CORRECTION IN "VIDIOGRAPHIC" . 336

8.1 VERTICAL SCANNING WAVEFORM GENERATION, STORAGE & CORRECTION. 336
8.1.1 Requirements 336
8.1.2 Implementation 340
8.1.2.1 Vidicon Implementation. 340
8.1.2.2 CRT Implementation. 342

8.2. HORIZONTAL SCANNING WAVEFORM GENERATION, STORAGE & CORRECTION - GENERAL 343
8.2.1 Introduction 343
8.2.1.1 General 343
8.2.1.2 CRT-Horizontal Coordinate Generation 344
8.2.1.3 Vidicon-Horizontal Coordinate Generation 345
8.2.1.4 Differences between H- and V-Correction. 349

8.3. H- SCANNING WAVEFORM GENERATION and CORRECTION for VIDICON. 349
8.3.1 Interpretation of Graphic H-Distortion Correction Information 349
8.3.2 Variable Delay Shift Registers 359
8.3.3 H- Correction Implementation 354

8.3.3.1 Generation of H-Correction Control Signals 354
8.3.3.2 Variable Delay Shift Registers 359
8.3.3.3 Accuracy of Correction System. 363

8.4 H-SCANNING WAVEFORM GENERATION and CORRECTION for CRT. 365
8.4.1 Introduction. 365
8.4.2 Interpretation of Graphic H-Distortion Correction Information. 366
8.4.2.1 General Interpretation 366
8.4.2.2 Approximations 368
8.4.3 H-Correction Implementation 370
8.4.3.1 Introduction 370
8.4.3.2 Correction by Voltage controlled Astable Multivabrators (VCA) 371
8.4.3.3 Requirements of Correction Voltages for the VCA 373
8.4.3.4 Generation of Correction Voltages for the VCA. 375
8.4.3.5 Distortion Correction within "Central" Distortion Region. 383
8.4.3.5.1 Change of sign in Distortion . 383
8.4.3.5.2 Change of Distortion slope in Central Distortion Region. . . 384
8.4.3.6 "Unsymmetrical" Distortion Corrections 389
8.4.3.7 Variable Delay Shift Registers . . . 391
8.4.3.8 Synchronizing VCA pulses with H-System Timing. 394
8.4.3.9 Summary of CRT- H-correction System. 397

8.5 JOINT REQUIREMENTS FOR CRT AND VIDICON DISTORTION CORRECTION 402
8.5.1 Actual Form of H-coordinate Input for Correction. 402
8.5.2 CRT - Vidicon Alignment 402
8.5.3 Finite vs Zero Distortion at Centres of Active Display Areas. 404

8.6 GENERAL COMMENTS ON V and H- COORDINATE CORRECTION.405

8.7 SUMMARY 408

CHAPTER 9. DISPLAY DISTORTION : CAUSES AND CORRECTION OF PRIMARY SOURCES OF DISTORTION 412

9.1 INTRODUCTION. 412

9.1.1 General Requirements 412

9.1.2 Electron Optics. 414

9.2 GENERAL CONSIDERATION. 415

9.2.1 "Distortion" and "Pincushion" or "Barrel" Distortion 415

9.2.2 General Concepts of Bean Generation, Focussing and Deflection. 415

9.2.3 Electron Beam Emission 418

9.2.3.1 The Electron Beam Emission System . 418

9.2.3.2 Emission System Aberrations 419

9.2.4 Comparison between Electric and Magnetic Deflection and Focussing 421

9.3 FOCUSSING DEFLECTION and FOCUSSING. 423

9.3.1 Introduction 423

9.3.2 Requirements of CRT and Vidicon Focussing. 423

9.3.2.1 CRT Requirements. 423

9.3.2.2 Vidicon Requirements. 425

9.3.3 Methods of Focussing. 426

9.3.4 CRT Focussing and Correction 427

9.3.4.1 Defocussing due to Beam Length Variation 427

9.3.4.2 Defocussing due to Non-normal landing of Beam 431

9.3.4.3 Conclusions 432

9.3.5 Vidicon Focussing Corrections. 432

9.3.5.1 Vidicon Focussing. 432

9.3.5.2 Defocussing due to Beam Length Variations. 433

9.3.6 Vidicon Beam Landing Correction. 438

9.3.7 Focussing Aberrations. 438

9.3.7.1 3-rd Order Aberrations. 438
9.3.7.2 Other Focussing Aberrations 438

9.3.8 Summary. 442

9.4 DEFLECTION OF ELECTRON BEAMS AND PINCUSHION & BARREL DISTORTION CORRECTION

9.4.1 Introduction 444

9.4.2 CRT Distortion and Correction. 447

9.4.2.1 CRT Distortion. 447
9.4.2.2 Distortion Correction 449

9.4.3 Vidicon Distortion and Correction. 452

9.5 COMMENTS. 458

9.6 SUMMARY 458

CHAPTER 10 "VIDIOGRAPHIC": SYSTEM OVERVIEW AND CONCLUSIONS 461

10.1 INTRODUCTION .461

10.1.1 General Remarks 461

10.1.2 "VIDIOGRAPHIC" Subsystems Not Treated in Feasability Study. 462

10.2 "VIDIOGRAPHIC" AS AN IGC 464

10.2.1 "VIDIOGRAPHIC" Objectives 464

10.2.2 "VIDIOGRAPHIC" Implementations of Objectives 464

10.2.3 Additional Advantages of VIDIOGRAPHIC . . 467

10.2.4 The Realization of "VIDIOGRAPHIC" 468

10.2.5 Physical Implementation of "VIDIOGRAPHIC" 473

10.2.5.1 Required Subsystems. 473
10.2.5.2 Operating characteristics. 473

10.3 "VIDIOGRAPHIC" PERFORMANCE 475

10.3.1 Introduction. 475

10.3.2 Display Parameters. 475

10.3.2.1 Resolution 475
10.3.2.2 Addressability 475

10.3.2.2 Addressability 475
10.3.2.3 Brightness 476
10.3.2.4 Contrast 476
10.3.2.5 Geometrical Linearity. 477
10.3.2.6 Repeatability. 477
10.3.2.7 Display Response Time. 477
10.3.2.8 Size and Layout of Display Screen. 478
10.3.2.9 Readability. 478
10.3.2.10 Half-tones and Colour. 479
10.3.2.11 Hard-copy Capability 479
10.3.2.12 Optional Features. 479

10.3.3 System User-Input and User-Display Interaction. 481
10.3.3.1 Introduction 481
10.3.3.2 Resolution 481
10.3.3.3 Addressability 481
10.3.3.4 Geometrical Linearity. 482
10.3.3.5 Repeatability. 482
10.3.3.6 Input Response Time. 482
10.3.3.7 Input Area Size. 482
10.3.3.8 Hard-copy Capability 482
10.3.3.9 General Input Subsystem Considerations 482

10.3.4 The Display Refresh Storage 484
10.3.4.1 Display Refresh Storage. 484
10.3.4.2 Distortion Correction Information Storage 487

10.4 COSTING OF "VIDIOGRAPHIC". 488

10.5 POSSIBLE FUTURE DEVELOPMENT OF "VIDIOGRAPHIC". 491

10.6 GENERAL COMMENTS AND CONCLUSIONS 494

PART 3 EXPERIMENTAL WORK

CHAPTER 11 EXPERIMENTAL WORK AND SELECTED CIRCUITS

11.1 INTRODUCTION 497

11.1.1 General Remarks 497
11.1.2 Available Hardware and Test Equipment. . 499

11.1.3 Policy of Experimental Work. 501

11.2 EXPERIMENTAL VALIDATION OF PHOTOMETRIC QUANTITIES 502

11.2.1 CRT Luminance 502
11.2.2 Vidicon Photocurrent and Electron Beam Landing Validation. 505
11.2.3 Schmidt Projection Optics and Light Pen . . 508

11.3 EXPERIMENTAL WORK ON DISPLAY LINEARITY MEASUREMENT 511

11.3.1 General Remarks 511
11.3.2 Experimental Setup for"VIDIOGRAPHIC". . . . 511
11.3.3 Vertical Bar Generator. 514

11.4 HORIZONTAL AND VERTICAL SCANNING CIRCUITS. 516

11.4.1 General 516
11.4.2 Scanning Circuit Requirements and Design. . 520
11.4.2.1 Scanning Circuit Requirements. . . . 520
11.4.2.2 Design Considerations. 522
11.4.2.3 Scanning Current Linearity Measurement. 524
11.4.3 Vidicon V-scanning Waveform Generator . . . 526
11.4.3.1 Requirements 526
11.4.3.2 Implementation 527
11.4.4 Vidicon H-scanning Waveform Generator . . . 537
11.4.3.1 Requirements.. 537
11.4.3.2 Implementation 538

11.5 FUNCTION CIRCUIT BLOCKS. 548

11.5.1 General Remarks 548
11.5.2 Constant Stabilized Level Output Circuit (CSL) 549
11.5.3 Pulse Integrator (P.I). 550
11.5.4 Discharge Circuit (D) 550
11.5.5 Voltage Controlled Monostable (VCM) 551
11.5.6 Voltage Controlled Astable (VCA). 551
11.5.7 Leading and Lagging Pulse Edge Indicator (LEI) 552

11.5.8 Schmitt Trigger 553
11.5.9 Current Amplifier 554
11.5.10 Voltage Supplies. 554

11.6 TIMING AND CONTROL 556

11.6.1 Requirements and Design 556
11.6.1.1 Fundamental Timing Signals 556
11.6.1.2 Other Timing Signals 558
11.6.1.3 Display Location Coordinate Generation 559
11.6.1.4 CPU I/O. 560
11.6.1.5 General Remarks. 561
11.6.2 Implementation. 562
11.6.2.1 The H- and V- counters 562
11.6.2.2 The Timing System Crystal Oscillator 563
11.6.3 Temperature Compensation. 565

11.7 CONCLUDING REMARKS 571

PART 4 –APPENDICES A.1

APPENDIX 1. APPLICATIONS OF MAN–COMPUTER GRAPHICS A.2
TABLES. APPLICATIONS OF MAN-COMPUTER GRAPHICS A.3

APPENDIX 2. AVAILABLE CIRCUIT ANALYSIS COMPUTER PROGRAMS . . A.5

AVAILABLE COMPUTER AIDED-CIRCUIT ANALYSIS PROGRAMS. A.6
TABLE 1. COMPARISON OF SOME AVAILABLE CIRCUIT ANALYSIS PROGRAMS. A.9
TABLE 2. SOME PRESENT MCG CIRCUIT DESIGN PROGRAMS A.11

APPENDIX 3. COMPARISON OF MAJOR CLASSES OF DISPLAYS 12

TABLE - COMPARISON OF MAJOR CLASSES OF DISPLAYS A.13

APPENDIX 4. COMPARISON OF COMMERCIALLY AVAILABLE CRT DISPLAYS AND IGCs. A.14

TABLE 1. REPRESENTATIVE SELECTION OF "LOW-COST" CRT GRAPHICS TERMINALS. A.15

TABLE 2. REPRESENTATIVE SELECTION OF COMMERCIALLY AVAILABLE IGCs A.16

APPENDIX 5. GRAPHICS INPUT AND POINTING DEVICES A.17

A.5.1 GENERAL. A.18

A.5.2 BRIEF DESCRIPTION OF GRAPHIC INPUT DEVICES. A.18

A.5.2.1 GRAFACON. A.18
A.5.2.2 SYLVANIA Data Tablet. A.18
A.5.2.3 GRAF-PEN Graphic Tablet A.19
A.5.2.4 ECRICON Graphic Tablet. A.19
A.5.2.5 Voltage Pen (Rose). A.19
A.5.2.6 Voltage Pen 2 (Resistive Input Tablet). A.20
A.5.2.7 Magnetically-coupled Pen and Tablet A.20
A.5.2.8 "Touch Sensitive" x - y Positioner A.20
A.5.2.9 Surface-Wave Detection Pen . . A.21
A.5.2.10 "Pressure Paper" Device - Analog A.21
A.5.2.11 "Pressure Paper" Device -Digital A.21
A.5.2.12 "Touch-wire" Device. A.21
A.5.2.13 Light-Pen or Light-Gun A.22
A.5.2.14 Light Pen for Plasma Display Panel. A.22
A.5.2.15 Beam Pen A.22
A.5.2.16 Lincoln Wand A.23
A.5.2.17 S.R.I "Mouse". A.23
A.5.2.18 "Joystick" or "Trackball". . . . A.23
A.5.2.19 MIT "Crystal Ball" A.23
A.5.2.20 "Data-coders", x - y Digitizers A.24
A.5.2.21 "Rho-Theta" Transducer A.24

TABLES - GRAPHICS INPUT AND POINTING DEVICES A.25

APPENDIX 6 CRT AND VIDICON BEAM DEFLECTION A.29

A.6.1 DERIVATION OF "δL", THE ELECTRON BEAM LENGTH INCREASE DUE TO BEAM DEFLECTION. . A.29

A.6.2 FEASABILITY OF RETAINING VOLTAGE DRIVEN VERTICAL DEFLECTION COILS IN CRTs A.31

A.6.3 DERIVATION OR CRT BEAM DEFLECTION AND DISPLAY DISTORTION - SUBMITTED PAPER. . . A.35

A.6.4 DERIVATION OF VIDICON BEAM DEFLECTION AND DISPLAY DISTORTION. A.56

APPENDIX 7. BEAM DISCHARGE IN VIDICONS A.67

A.7.1 NATURE OF ELECTRON SCANNING BEAM. A.68

A.7.2 BEAM IMPEDANCE AND BEAM DISCHARGE TIME CONSTANT. A.70

A.7.3 VALIDITY OF USING A STATIONARY BEAM WITH DWELL TIME "T_d" WITH THE SCANNING BEAM. . A.72

A.7.4 EVALUATION OF BEAM DISCHARGE EXPRESSIONS. A.74

A.7.5 BEAM BENDING DUE TO MESH RETARDING FIELD. A.75

A.7.6 PREDICTION OF EDGE-FLICKER FROM STABILIZATION CURVES. A.78

A.7.7 DERIVATION OF TIME INTERVAL BETWEEN ERROR SIGNALS A.82

APPENDIX 8. PHOTOMETRY AND OPTICS A.88

A.8.1 PHOTOMETRY. A.89

A.8.1.1 Photometric Quantities - An Introduction A.89
A.8.1.2 Luminous Flux. A.89
A.8.1.3 Luminous Intensity A.91
A.8.1.4 Luminance, Brightness. A.91
A.8.1.5 Luminosity A.92
A.8.1.6 Illumination A.92

A.8.2 VIDICON SENSITIVITY A.93

A.8.2.1 Introduction A.93

A.8.2.2. Sensitivity of Vidicon Photo-Conductor to "W" Phosphor Spectral Illumination. A.93
A.8.2.3 Number of Photons per foot-candle of Illumination. A.99

A.8.3 OPTICS RELATIONSHIPS. A.102
A.8.3.1 On-Axis Illumination from Sources A.102
A.8.3.2 Off-Axis Illumination from Sources. A.104
A.8.3.3 Lens Image Illumination. A.105
A.8.3.4 Vidicon Lens Depth of Field. . . A.106

APPENDIX 9. PHOSPHORS AND CATHODOLUMINESCENCE A.110

A.9.1 PHOSPHORS AND CATHODOLUMINESCENCE. A.111
A.9.1.1 Introduction A.111
A.9.1.2 Phosphors. A.111
A.9.1.3 Phosphorescence. A.113
A.9.1.4 Phosphorescence Decay. A.114
A.9.1.5 CRT Phosphors. A.115

A.9.2 PRACTICAL CRT LUMINANCE RELATIONSHIPS . . A.116
A.9.2.1 CRT Output Luminance A.116
A.9.2.2 CRT Input Output Transfer Curves A.119

A.9.3 CRT LUMINANCE DECAY. A.121
A.9.3.1 General Form of Luminance Decay A.121
A.9.3.2 Light Pen Illuminating Peak Luminance A.123

A.9.4 PHOSPHOR LUMINANCE BUILD UP. A.124

APPENDIX 10. PHOTOCONDUCTIVITY A.128

A.10.1 PHOTOCONDUCTORS AND PHOTOCONDUCTIVITY . A.129
A.10.1.1 Introduction. A.129
A.10.1.2 Photoconductivity A.131
A.10.1.3 Photoconductor Properties . . . A.134

A.10.1.3.1 Dark Current. A.134
A.10.1.3.2 Spectral Response A.135
A.10.1.3.3. Speed of Response . . . A.136
A.10.1.3.4 Photosensitivity. . . . A.138

A.10.2 VIDICON REQUIREMENTS FOR PHOTOCONDUCTORS A.140

A.10.2.1 Introduction A.140
A.10.2.2. Signal Current Considerations. A.140
A.10.2.3 Phototarget Capacitance. . . . A.141
A.10.2.4 Phototarget Thickness. A.142
A.10.2.5 Maximum Capacitance of Phototarget.. . . . . A.144
A.10.2.6 Phototarget Resistance A.145
A.10.2.7 Phototarget Contacts A.146

A.10.3 EXPRESSIONS FOR PHOTOCURRENTS, DARK CURRENT, GAIN ETC.. . . . . A.147

A.10.3.1 Dark Current A.147
A.10.3.2 Electron Mobility. A.148
A.10.3.3 Photoconductor Gain,"G". . . . A.149

A.10.4 DERIVATION OF PHOTOCURRENTS BUILDUP AND DECAY TIME CONSTANTS. A.152

A.10.5 EVALUATION OF PHOTOCURRENT BUILDUP AND DECAY A.157

APPENDIX 11. MEASUREMENT OF DISPLAY DISTORTION AND VIDICON SCANNED AREA DISTORTION : MOIRE FRINGES . . . A.160

A.11.1 INTRODUCTION. A.160

A.11.2 DISTORTION CONCEPTS A.162

A.11.2.1 "Location-to-Location" Distortion A.163
A.11.2.2 "Relative" or "Cumulative" Distortion. A.164

A.11.3 MOIRE PATTERN METHODS OF MEASURING DISTORTION. A.165

A.11.4 GRAPHIC DISTORTION CORRECTION INFORMATION FROM MOIRE PATTERN MEASUREMENTS A.167

A.11.4.1 Horizontal Correction Information A.167
A.11.4.2 Vertical Correction Information A.168
A.11.4.3 Conclusions. A.172

A.11.5 REPRINT ON MOIRE FRINGES. A.173

APPENDIX 12 TYPICAL TV-CRT AND VIDICON CAMERA SYSTEMS A. 175

A.12.1 INTRODUCTION AND REQUIREMENTS A.175

A.12.2 TV-CRT SYSTEM. A.180

A.12.2.1 VIDEO Amplifier Subsystem. . . A.181
A.12.2.1 Deflection Subsystem. A.181

A.12.3 VIDICON CAMERA SYSTEM. A.183

A.12.3.1 VIDEO Amplifier Subsystem . . A.184
A.12.3.2 Deflection Subsystem A.184

APPENDIX 13 - REPRINTS OF PAPERS AND MISCELLANEOUS A.186

A.13.1 REPRINT OF PAPER "COMPUTER DISPLAY DISTORTIONS: CAUSES, MEASUREMENT & CORRECTION" A.187

A.13.2 REPRINT OF ITEM IN "COMPUTER GRAPHICS" WINTER, 1969 - 70. . . . . A.199

A.13.3 INVITATION TO PARTICIPATE AS A PANEL MEMBER TO DISCUSS "COMPUTER GRAPHICS" AT THE ACM CONFERENCE, NEW YORK. A.202

A.13.4 COPIES OF DESIGN NOTES ACCEPTED FOR PUBLICATION. A.204

A.13.4.1 "Linear Voltage Controlled Astable and Monostable Using the uL 914" A.206
A.13.4.2 "Pulse Leading and Lagging-edge Indicator ". A.218

A.13.5 COPY OF SHORT DESIGN NOTE ACCEPTED FOR PUBLICATION. A.220

"Low Voltage-Rail Voltage Controlled Monostable". A.222

A.13.6 MISCELLANEOUS LETTERS. A.226

REFERENCES R.1

General Bibliography. R.2
References. R.3

SUMMARY

"Man-Computer Graphics" may be defined as the solution of problems by a computer where, during some stage of data input, processing or output of results, data is manipulated or presented in graphic form; in addition, the user interacts with the computer on-line during one or more stages of data input, processing, or data output. The Man-Computer Interface enabling this is an "Interactive Graphics Console" or IGC.

One of the major current problems in Man-Computer Graphics is the absence of an economical and highly interactive IGC. By "economical" is meant low capital outlay for the IGC and low CPU tie-up time during use; and by "highly interactive" is meant the ability by the user to input graphic data easily, as say with a "pen" or "stylus", with very short response time, and the ability to interact with displayed graphics information.

A wholly new means of implementing such an IGC is described and its feasability explored.

The name of the proposed system, "VIDIOGRAPHIC", (for Video Integrated Display and Input Optical GRAPHic Interactive Console) indicates the means used to implement this. TV techniques and commercially available equipment are used; optical methods are used to achieve user-graphics input, user-display interaction, and display refresh storage.

Basically the system consists of a TV Vidicon

camera looking back at its own output displayed on a CRT. Under certain conditions of CRT and Vidicon linearity and CRT luminance, a CRT display is self-maintaining. Graphics input and user-display interaction is achieved with a simple light-emitting "pen."

Calculations and feasability studies indicate the practical feasability of "VIDIOGRAPHIC". Means of implementing the proposed system and new circuit designs to achieve this are given.

A general overview of the Man-Computer Graphics field along with comprehensive hardware review of existing techniques are also given; in the poorly documented Man-Computer Graphics field these in themselves form a useful function. In this case they are also useful in comparing how "VIDIOGRAPHIC" meets the requirements of an economical and highly interactive IGC.

. . . .

DECLARATION OF ORIGINALITY

This Thesis contains no material which has been presented or accepted for the award of a degree or diploma in any University, and to the best of my knowledge, the Thesis contains no material previously published or written by anyone else, except where due reference is made.

V.C. SOBOLEWSKI

ACKNOWLEDGEMENTS

The author wishes to make known his gratitude to his supervisor, Mr.D. Pucknell, who took over the supervision of his Ph.D and was always there to guide and help as the occasion demanded; to Prof. E.O. Willoughby who earlier on helped to keep the author's feet on the ground, and to Prof. J.L. Woodward, Head, Electrical Engineering Dept., who kept the Administrative hounds at bay and the financial coffers open.

Last, but by no means least, to Mrs. J. Grevins, whose perseverance and typing ability transformed the illegible into the legible and presentable, and finally to Miss Izabella Milczewska (the soon-to-be Mrs.V. Sobolewski), for her patience, her pattern recognition powers and artistic ability to change the author's graffiti into recognizable figures and illustrations.

Expressions of thanks must also go to the Commonwealth Department of Education and Science for the Commonwealth Postgraduate Scholarship during the years 1965 - 1968, and to the University of Adelaide for a University Research Grant, during a period of 1969, which enabled the author to carry out this work.

MARCH 1971

. . . .

PART 1

MAN-COMPUTER GRAPHICS: GENERAL CONCEPTS

CHAPTER 1

MAN — COMPUTER GRAPHICS

1.1. INTRODUCTION - THE ROLE OF MAN - COMPUTER GRAPHICS

The "usefulness" or "power" of a computer as judged by a user depends on the "user" output presented to him by the computer, compared with the input he has presented to the computer, and on how fast the output appears after the input. Clearly the usefulness then is dependent on the preliminary work the user does in preparing a program: changing data into suitable computer formats, programming and coding the problem etc. Minimizing this work would greatly increase the "usefulness". So, if the input into a computer is in a form most natural to the user, such as flow diagrams, block diagrams, circuit diagrams, graphs, waveshapes and the like, "computer power" would be significantly increased. Similarly, if computer outputs are again in a "natural" user format, such as graphs, curves etc, rather than the usual numerical printouts which need be further manually processed, the usefulness again would be enhanced. If Computer Input and Output Units capable of accepting and displaying data in the above formats are available, a great increase in computer usefulness would result. The Input/Output System of computers and computer usefulness are thus intimately interrelated.

The second factor which increases usefulness of computers is shortening the response time of the computer (or rather the computing facility), between presenting the input and obtaining the output or results. This leads to "real-time", or "on-line", or direct man-computer interaction. The increase in usefulness results when the often irksome delays due to closed-shop, batch processing of programs are eliminated. Turn-around times of a day, or hours at the very least, cause loss of trend of thought and user dissatisfaction. A single error in punching data cards or book-keeping error of the program cards may cause a frustrating delay of a day. Modifying the I/O System of a computer, together with systems programming extensions to make on-line operation possible (and economical), increases computer usefulness.

There is no simple answer as to what form the input data and information and output results must take to be considered "most natural" to the user. The range of problems solved on computers is very large and varied. Sometimes purely alpha-numerical results are adequate or desirable; at other times purely graphical results make sense. However there are few engineers or scientists that do not at some stage, within a given problem, make use of sketches, flow-or block-diagrams, or plot one variable against another. Such examples indicate that graphical "language" is very common and natural; in some stages, particularly the "design" stage of a problem, "graphical" language seems to be exclusive.

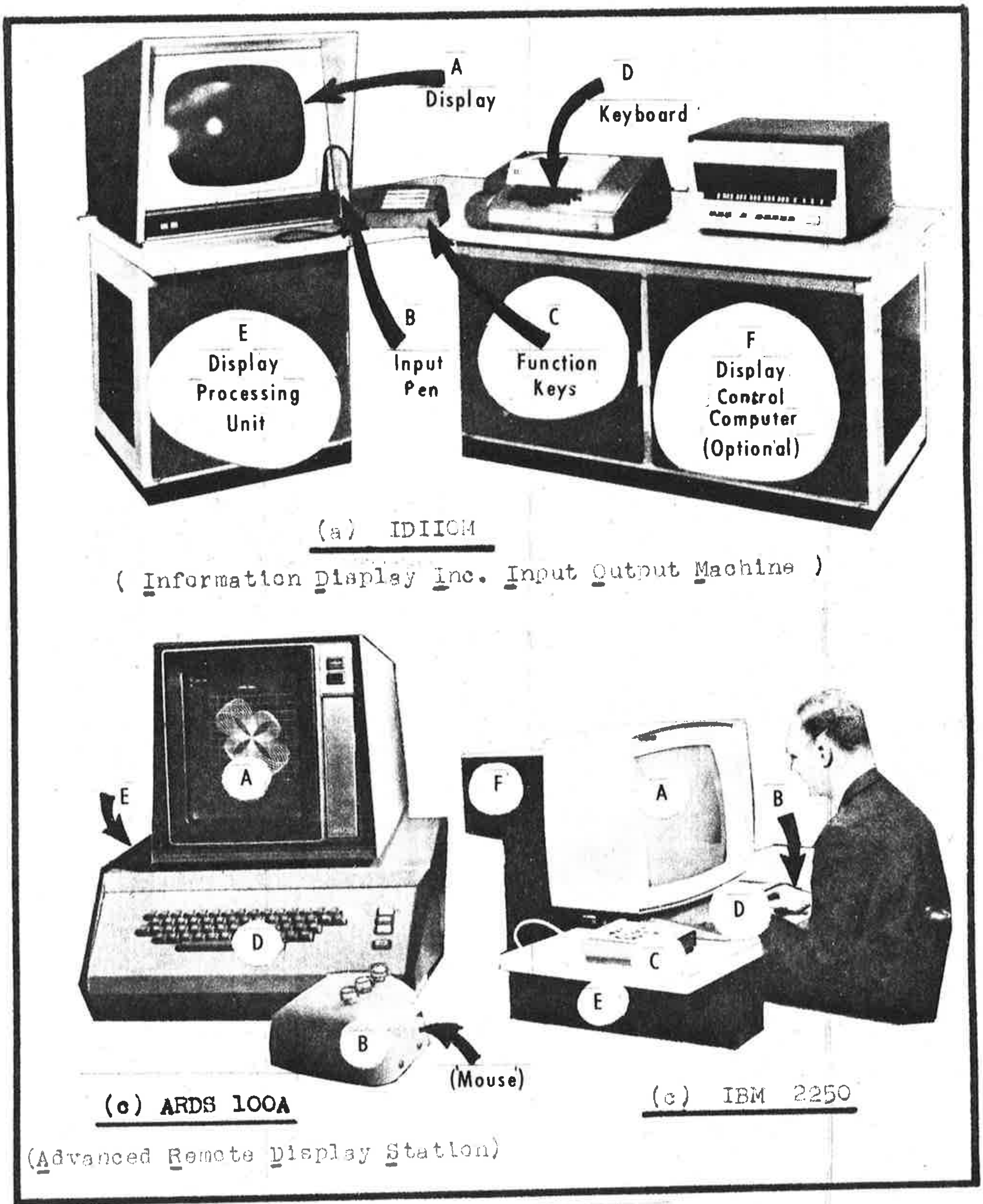
The statement "a picture is worth a thousand words" has often been made. For data and numerical results, the advantages of graphic representations are obvious as trends and relationships otherwise not apparent from tabular representation become evident. Sometimes results can be represented only sensibly in graphic form such as in the case of printed circuit board layouts. At other times, particularly when explicit mathematical relationships are not known, such as in non-linear problems, solutions are often carried out graphically; often this may be the only way of obtaining a solution. When formulating a problem, sketches such as flow diagrams, block or circuit diagrams are indispensable. It can be seen that "pictures" or graphic representations enter all stages of the problem-solving procedure,from the early innovative stages of designing the problem to the presentation of results in an efficient and informative format. Fig. 8,showing a typical circuit design cycle,indicates that most stages of the cycle include some form of graphic representation or evaluation. Thus improvements in the Input-Output System of the computer, so that problem information and solutions can be carried out in the above graphical forms, greatly help the user and increase the usefulness of a computer as judged by a user.

A second and perhaps more important improvement can result. Creative scientists and engineers when presented with the facilities of solving problems in natural and graphic form and able to see the results of their efforts immediately, tend to be more adventurous and try more design modifications (for optimization) or even different design approaches (2). The result is a more optimum design with a better understanding of the problem by the designer. More profitable time utilization of the designer results to the benefit of management and the user himself.

Speed increase of 8 - 10 fold have been reported (3) over a complete design process when graphics are used. In the generation of artwork for multilayer printed circuits and I-C's speed increases of 10^5 over manual techniques have been achieved (78).

The subject of this Thesis is, as stated in the Preface, an economic Input/Output Graphics Unit, or Interactive Graphics Console (IGC). This immediately implies "Computer Graphics" and particularly "Man-Computer- Graphics (MCG). Computer Graphics (CG) refers to that aspect of computer activity which involves partially, or in whole, the accepting, processing, generating and outputting of data in graphical form. Early in the development of the field, output displays such as Cathode-Ray-Tubes (CRT), pen-recorders, x - y plotters etc and their implementation comprised Computer Graphics. Lately it has come to include not only the computer manipulation and generation of graphic images, but also those aspects of on-line, Man-Computer interactions performed during the process of computer solution of problems. More properly this is called "Man-Computer Graphics" (4).

Some typical installations with Man-Computer Graphics in process and Interactive Graphics Consoles are shown in Fig.1. "Computer -Aided-Design", "Computer-Augmented Design" are some other synonyms which have been used with MCG. These descriptors are not preferred as the concept of on-line, Man-Computer interaction, so implicit in MCG, is not present. Indeed "design problems" such as say the A.C. analysis of a circuit, solved with the designer off-line, and with results such as node voltages etc. numerically tabulated, still come under the name of "Computer-Aided-Design".



(a) IDIIOM

(Information Display Inc. Input Output Machine)

(c) ARDS 100A

(Advanced Remote Display Station)

(c) IBM 2250

Figure 1. Typical Interactive **Graphics** Consoles Showing major **Subsystems.**

A look at the development of "CG" and "MCG" will indicate that it has been motivated to increase computer usefulness and to utilize the user-designer's time more effectively. This may seem an obvious point to make, but there are still many people who consider MCG or even CG to be an expensive luxury. This negative view-point has (had?) some justification. Previously (and to some extent, still at present) I/O systems were designed with the computer and not the user in mind. To get data in and out, computations had to be stopped and this cost money. Such delays, when data was input or output, were minimized if they were under computer and not user control. However the concept of "time-sharing" (5,6) and its implementation in the early Sixties has overcome the economic objections.

> "Time shared computer usage is absolutely necessary to bring the full potential of computer power to each computer user..." (4)

In "time -sharing", the computer, rather than attend to one problem at a time, simultaneously attends to several problems by allocating "slices" of its time (of say 0.2 seconds each) to each problem in sequence. The allocation in sequence and in length of time-slice can be controlled (and thus the sequence broken) by so-called "program-interrupts" which efficiently perform this allocation between the various simultaneous programs; the priority of service being usually based on speed of the equipment demanding attention. The breaking of the normal sequence of time-sharing can be due to, say, entering of new data by the operator. It is not the purpose here to go into details of time-sharing but merely to indicate it making MCG an economic feasability. Faster I/O equipment, reduction in computing cost and equipment cost, may have made MCG possible, but without time-sharing it would have remained the luxury that the critics of MCG say it is.

This argument of course holds for large computing facilities. The other development which has helped the growth of MCG has been the introduction of small-scale economic computers. Small capital outlay and low running costs makes it possible for a graphics console to be wholly serviced and controlled by the one computer even with large CPU downtime during I/O operations. Limited memory, however, precluded all but the relatively small graphics problems to be solved on these. As a consequence this led to the use of such small computers as "buffer-computers" or "graphics interface" computers between a large central computer and the graphic console (s) with the main computer doing the computing and graphics manipulation, while the smaller computer does the graphics generation and graphic console servicing.

A short resume of possible system configurations with graphics consoles is given in section 2.6.2.

1.2. SHORT RESUME OF DEVELOPMENT OF MAN-COMPUTER GRAPHICS

1.2.1. Early History

Suffice to say that by the end of the Fifties simple computer-controlled displays in the form of Cathode Ray Tubes (CRT's), X-Y electromechanical plotters and the like were common. Graphic input devices such as electro-mechanical digitizers, flying-spot scanners were also implemented. Quite excellent reviews of the development and implementation of such graphic I/O equipment are found in (7,8,9). Applications included graphic program-debugging, representation of motions of objects in time sequence, direct plotting of results and so on. With the Graphic Input equipment, control tapes for numerically controlled machines could be generated from engineering drawings while automatic map-contour reading and direct inputting of graphical results obtained from previous calculations were occasionally implemented. Such graphic data was digitized and processed numerically.

It only remained to solve the problem of computer manipulation of graphic data.

Two different approaches were taken, outwardly with two different motives in mind. These will be examined in some detail for within these two approaches lies the heart of Man-Computer Graphics. Their implementation and development has guided MCG ever since.

(a) "Applied Approach"

One approach, developed by several automotive and aero-space firms concerned itself with the engineering design of complex wing sections, body parts, structures etc with the help of computers. Since the input and output for these problems consisted of engineering diagrams of mechanical assemblies, it was felt that the intermediate computer processing and manipulation should also be done in graphical form.

(b) "Fundamental Approach"

The other approach, taken up by several research establishments and universities, mainly MIT and TRW, had as its aim the increasing of the scope of problems capable of being solved on digital computers, particularly those problems which were, and still are, difficult to define explicitly. This led to the development of new forms of languages and communications, in particular graphical languages and graphical communication.

1.2.2. "The Applied Approach" - Design Augmented by Computer "DAC-1" (10,11,12,13,14,15,16)

Developed by General-Motors, who provided the requirements and underlying philosophy, and IBM, who developed the required hardware,this system was initiated in 1959 and the results publicly released in 1964. The business of GM was automotive design, a highly competitive field with new, sucessful designs required in the shortest possible time. "Computer augmented" design was the only answer to keep the design cycle period as short as possible. The philosophy was based on the conclusion that all stages in the design process can be described by drawings - an initial sketch, a proposed assembly drawing, a final engineering plan. Particularly,

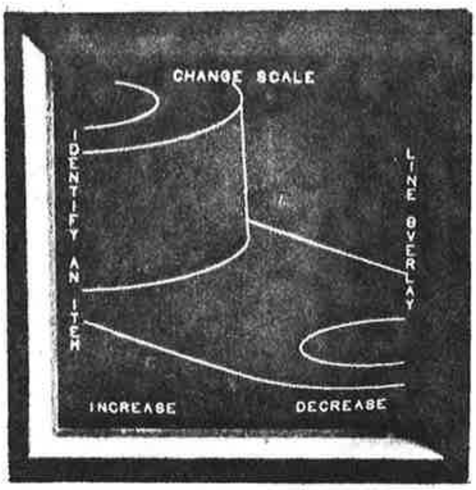
> "...Dependence on graphical techniques ...is fundamental to the design process. Graphics serve as a language of communication among design personnel and as a mechanism for design evolution" (10)

The system which resulted had one or two engineers sitting in front of a graphics console, consisting of a large working CRT (17" diameter) in front of them, an alpha-numeric keyboard, and a keyboard with program function keys, such as "rotate", "scale up", "delete". A hand-held indicating pen, based on direct voltage-pickup (see Appendix 5) was used to locate or point to any feature on the display. This "pointing ability", effectively an addressing function to any feature of the display, is a fundamental feature of graphic communication. Excellent illustrations of the above system can be found in (10,11) and in Fig. 2.

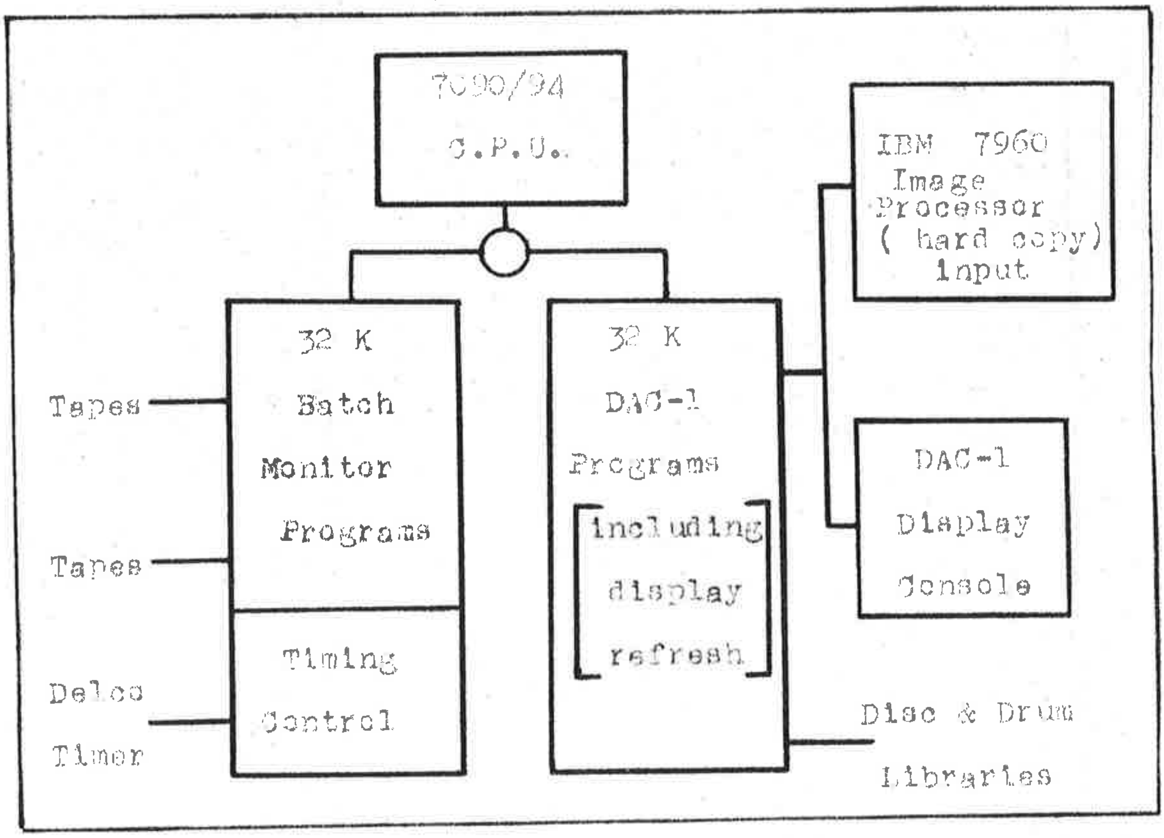
There was also a graphics processor (IBM 7960 Image Processor) which generated hard-copy of drawings etc on 35mm film and which was used as a graphics store and for graphics I/O to and from the computer.

(a) DAC-1 in use.

(From [11])



(b) Typical Output Display with available Cues clearly shown.



(c) DAC-1 System Configuration (memory-oriented).

Figure 2. DAC-1 (Design Augmented by Computer)

Typically an engineer(s) may call some drawing section from store and see it displayed on the CRT; his immediate actions, which he can take, are presented to him alpha-numerically on the CRT as cues such as "rotate" and "specify angle of rotation", "modify", "scale" etc (see Fig. 2(b)). Using the keyboard (or the pointing "voltage pen"), he specifies his action or request. The computer, an IBM 7094, does the rest. Over 700,000 operating instructions were necessary!

The proportion of operator-time to CPU time was about 10:1 i.e. each 10 minutes of console operation required about 1 minute of CPU time. For CPU efficiency and economy, the other 9 minutes were utilized on some other part of the program or a different program. This was achieved through "multiprogramming", which is a technique whereby a supervisory program allows the concurrent processing of several routines or programs, in particular when I/O operations are involved.

It can be seen that the above effort involved hardware development such as linear CRT's, pen-position detection hardware, image processors via fast film development etc and software development incorporating multi-programming, disc libraries of programs and the subroutines for graphic manipulation.

The system not only displayed and accepted graphic information, but the operator conversed in a natural graphical form with the computer by pointing to features of a drawing with his pen, and specifying the operation to be done by a key-board or indicating to the required cue, if it were a displayed option; all this to the user appeared to be in "real-time" or at least, with insignificant delays (say several seconds).

Similar schemes were initiated independently by other companies (17,18,19) particularly Lockheed - Georgia. The scope of design problems in the latter Company was expanded to include printed-circuit board and Integrated-Circuit lay-outs, mechanical member design, successive orientations of free-falling objects in space, pilot visibility from cock-pits and others, all essentially graphic-geometrical problems. It has been stated by one of the authorities in this field that

> "... the potentialities of MCG in this area have only begun to be exploited..." (4)

1.2.3. "The Fundamental Approach" - Project MAC, SKETCHPAD etc (20 - 30)

1.2.3.1. Introduction

A more basic approach was taken at MIT, with the work starting around 1958-59. The aim was to explore the basic problems of graphic-information processing: the manipulation and generations of drawings, the computer specification of graphic information, subroutine specification, the operator design-process and the role of graphics therein, the basic features of graphic information etc amongst others. The experimentation and implementation of the above problems was made possible by the development, in early 1960-61, of a time-sharing system (6) - (time-sharing itself was proposed barely two years earlier by Strachey (5)).

The task, immense as it was, could be divided into several main problem areas:

(a) the system had to have generality - unlike the specialized tasks for which DAC-1 was designed.

(b) the <u>nature of graphical information</u> and the associated language whereby it and the design process could be described had to be determined.

(c) <u>hardware</u> problems.

(d) <u>implementation</u> - resulting in SKETCHPAD I and other systems.

<u>1.2.3.2. Generality</u> (20)

The system had to cater for many forms of design - and not merely automotive design as in DAC-1. Thus the user design process had to be analysed and features common to each design process recognized. It was concluded that for a computer to work in partnership with a designer, the computer

> "... must be able to accept, interpret and remember shape-descriptive information introduced graphically..." (20).

In addition the usual computational facilities to perform the mathematics relevent to a particular design process, such as say circuit analysis, stress analysis etc need be provided.

To achieve this, two different strategies were possible.

(a) One (that chosen by the DAC-1 workers with the disc libraries of programs)

"would be to imbed in the computer a large set of special purpose packaged procedures, each designed to perform some special task" (20)

This clearly was impossible as it had been found that a true design process is unpredictable - it is impossible (or at least unprofitable) to "build in" unpredictability into the system!

(b) The other alternative was to imbed in the computer

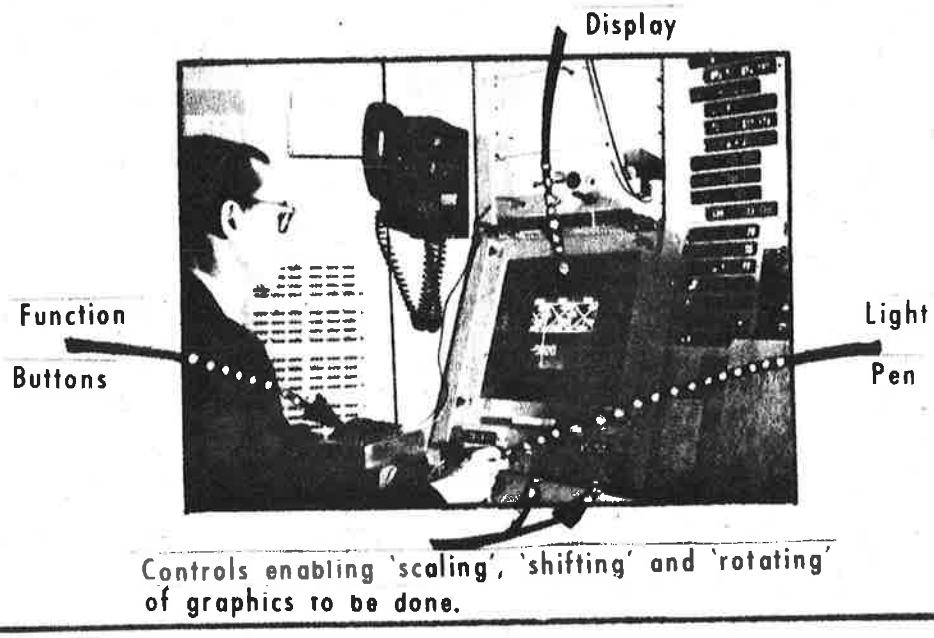
"..a few...routines of the utmost generality, so designed that it permits its own modification by the designer, using its own natural language forms, including...the graphical form" (20).

1.2.3.3. Form and Language Describing Graphical Information (21,22,23)

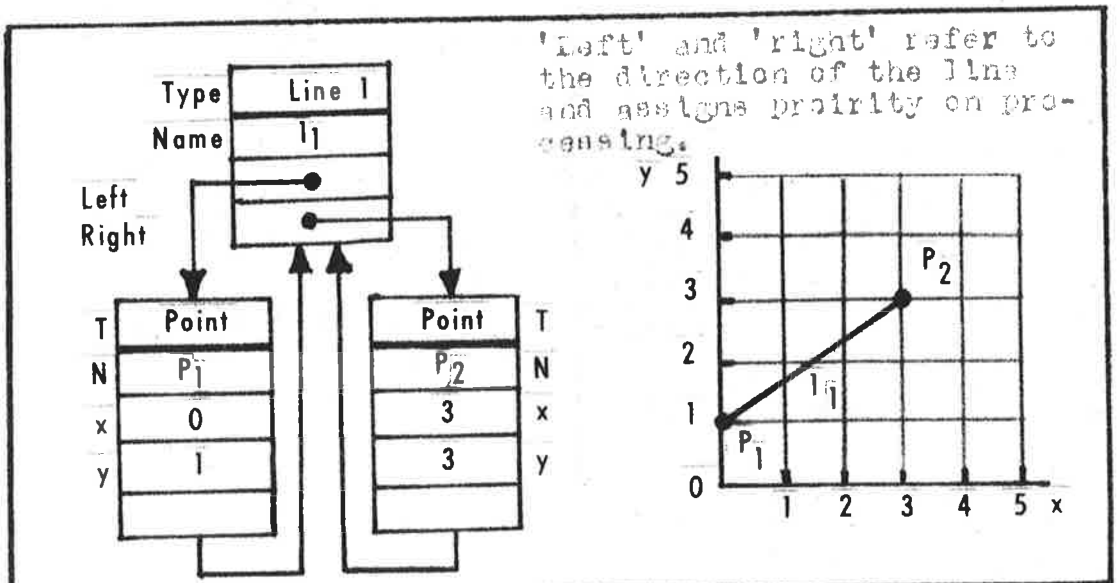
The above conclusions led to the necessity of a new descriptive and executive language capable of describing graphic data forms and implementing the desired manipulation of such data. If this could be implemented then, because all the necessary information about a problem and its method of solution can be described, the problem itself can be solved.

Graphics information was modelled using "n-component" and "plex" structures which are akin to multi-dimensional list structures. An "n-component" element is a single unit of information about a problem; each component in it specifies some particular attribute or property of the elemtnt: "n" is arbitrary and is different for each element; each of these attributes may be independently accessed. Several such n-components form a "modelling plex" with the various attributes interconnected.

In Fig. 3(b), 3 n-components, each having the descriptors "line", "point" and "point" respectively, when interconnected,describe a line. By joining different attributes, different properties may be obtained such as length, slope of line etc. The accessability of the separate line properties is reflected in the accessability by the user to corresponding features in the graphic image, by pointing. "Plex-structures" are thus the natural data structures for graphics data and images.



(a) SKETCHPAD implemented on TX-2 computer.



(b) Graphical data representation ('modelling n-component plex structure') of line "l_1" shown on right.

Figure 3. SKETCHPAD (from [26])

Plexes are sucessively processed through so-called "parsing algorithms" which rearrange the plexes in a certain order for computation. The language in which plexes and processing routines were written in was a language specifically developed for this, "AED-0" (Algol Extended for Design or Augmented Engineering Design)(22). It is beyond the scope of this short review to go into any detail of this language; the above references (21,22) cover this aspect comprehensively.

1.2.3.4. Hardware (24,25)

The graphics console and associated hardware provide the interface between user and computer. The computer-controlled CRT, with a "working area" of about 10"x10", has an indicating stylus, the "light pen" with pointing ability (see Appendix 5). Also, function generators are provided, capable of generating on the display straight lines and second-order curves. A hardware unit, generating a "rotation matrix" which rotates simple curves and figures from simple views (e.g. from a plan view) to produce complex 3-D projected views or complex curves, is also provided. Thus complex conic sections, say, can be generated from a simple circle appropriately rotated; this economizes on CPU time. The light pen (described in (25)) is essentially a light detector which, when pointed to at the CRT screen, responds to the initial flash of the phosphor on the CRT (to indicate precisely the instantaneous position of the electron beam). With some sophisticated subroutines and keyboard controls, lines can be generated by pointing at the two end points, circles can be drawn by indicating the centre and typing-in the radius etc. Other control knobs, giving rotation result in complex curves being generated. These features are shown in Fig.3(a).

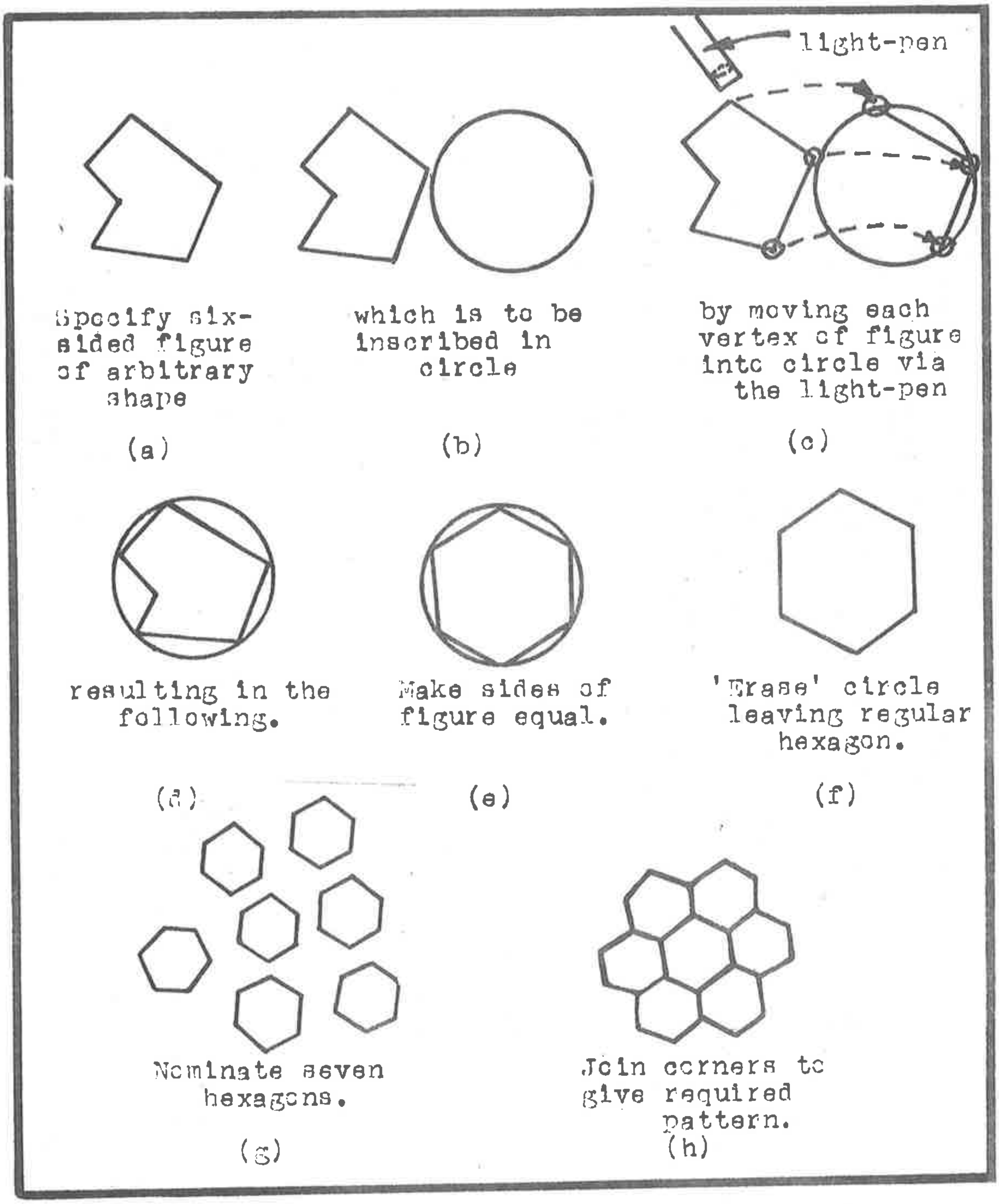


Figure 4. Example of SKETCHPAD in use.

Construction of complex figures by simple user-actions and simple SKETCHPAD commands.

(from [26])

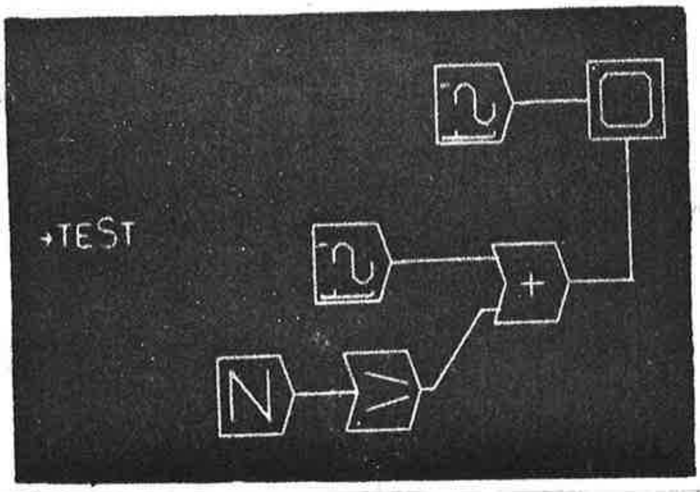
1.2.3.5. Implementation - Sketchpad I (26)

The above concepts systems and language software and hardware were spectacularly (!) synthesized by I. Sutherland in his Sketchpad I program.

A user program is run without typed statements (except for legends and descriptors) and with only line drawings being used. The execution of commands (and hence solutions) is based on the fact that geometrical (and hence graphical) constraints need be obeyed - e.g. triangles must always close, lines under bending, rotation etc must remain constant in length and so on. (The actual geometrical constraints are computed numerically). Geometrical constructions are made by pressing appropriate function keys. An "erase" function is also present.

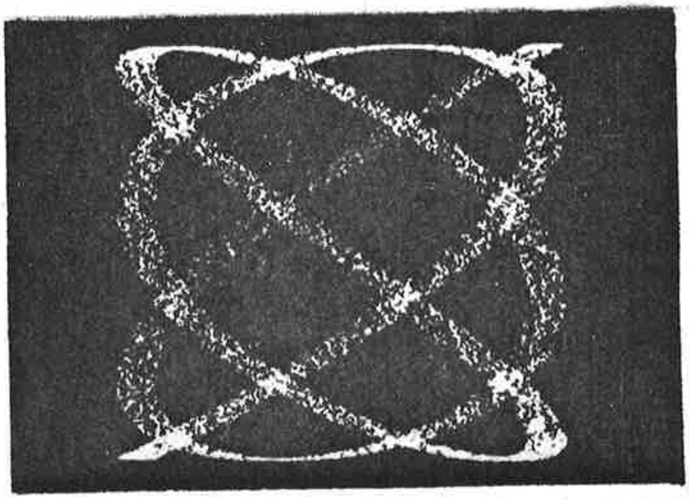
As this survey is on computer graphics, the advantages of graphics should be shown rather than described. An illustrative example showing the versatility of Sketchpad I with simple user orders is given in Fig. 4; it is self-explanatory.

The superiority of this approach over conventional programs using, say, FORTRAN can best be demonstrated by the author of an article who stated that within a space of several hours,five separate design problems were initiated and solved (20); a cubic polynomial from which the argument could be found by specifying the ordinate; deflections of loaded bridge members; the motion of connected links of a kinematic linkage; fluid flow within a certain configuration; and the construction of a general second degree curve. The last four problems are still quite often solved graphically by hand.



Simulation of 'noisy' Lissajous patterns by retrieving from store function blocks represented graphically, interconnecting them and observing the resultant display.

(a) Via the light pen the function blocks are positioned and interconnected. "∿" are varying frequency sinusoid generators, "N" is a noise generator "⊳" is an attenuator and "□" a quadrature summer.



(b) The resultant displayed 'noisy' Lissajous figure.

Figure 5. Graphical Programming. (from [30])

1.2.3.6 Later Developments at MIT

Later developments led to Sketchpad III, a program for drawing and manipulating in 3 - dimensions (27,28). W. Sutherland (brother of I.) went one step further by graphically describing computer procedures (29). By drawing flow diagrams consisting of "blocks" representing some function, and interconnecting these blocks, various arithmetical procedures could be executed e.g. a block with a "+" inside it, with two inputs, specifies the addition of the inputs at the block output. A much more "graphic" demonstration is given in (30) and reproduced in Fig. 5 - the captions make it self explanatory.

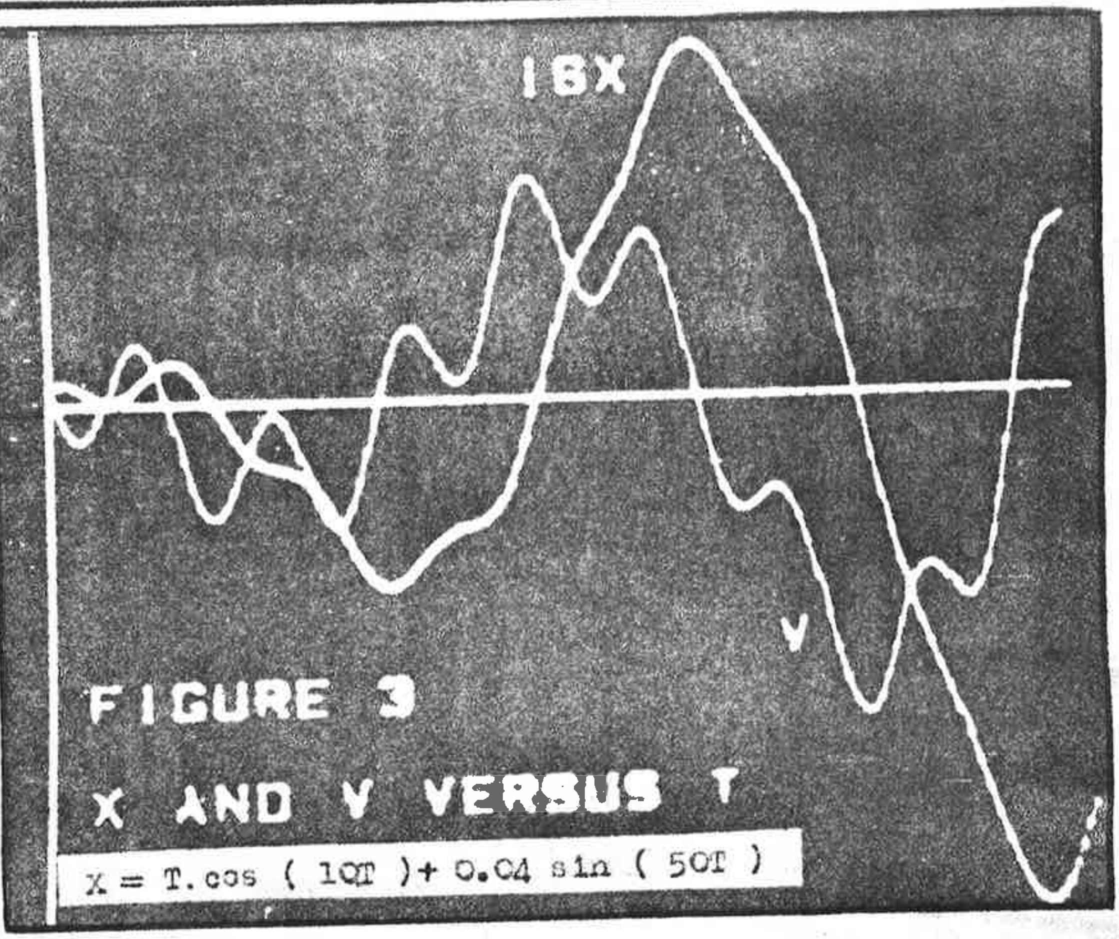
1.2.4. Other Approaches To Man-Computer Graphics

A third development concurrent to DAC-1, SKETCHPAD and Project MAC must be mentioned. This was the TRW Two-Station On-line Scientific Computer, developed by Culler and Fried (31,32). Again the motivation was as a result of ...

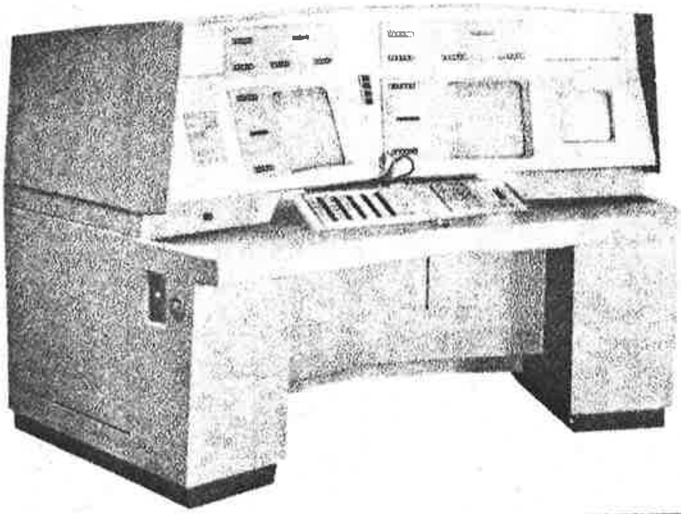
> "... the fundamental problem of communication... includes language and equipment requirements, as well as the basic matter of means of interaction between the problem solver and the problem structure....." (32).

the same problems which confronted the other workers in the field. The difference lay in the application - the main goal was in "scientific work", particularly classical mathematical analysis.

The characteristics of the resultant system were essentially based on on-line programming, with the programming units or "words" representing mathematical analysis subroutines, including any subroutines developed during a problem run, rather than primitive functions as in say FORTRAN. Intermediate and final results could be displayed graphically immediately on computation; erroneous trends



(a) Actual photo of displayed function and its derivative obtained by a sequence of simple keyed-in functions. (from [32])



(b) View of 'Display and Analysis Console' (DAC).

Figure 6. Example and Layout of TRW Operator's Console

in calculations could be easily detected and thus a particular approach aborted. The value of this system was demonstrated by successfully solving non-linear partial differential equations with non-simple coefficients.

It is natural that the originator, or someone wholly familiar with such a problem, should be at the console and it is possible that such a person would not be a good or an expert programmer; a good programmer would be a necessity to attempt to cover all possibilities of solution in such a complex problem, if off-line problem solving were used. As was stated at the time,

> "...the work... was motivated by the troubles which have been commonly encountered in using a computer to solve research problems whose structure is for the most part unknown and frequently suprising." (31)

Figure 6 shows the layout of the TRW system together with a typical display.

1.2.5. Summary.

To recapitulate - the motivation and steps leading to Man-Computer Graphics were:

(i) the need for man-computer interaction in ill-defined problems.

(ii) ill-defined problems and the design process have much in common.

(iii) much of the information transfer and communication in ill-defined problems is in graphic form.

(iv) to execute and solve such ill-defined problems expressed in graphic form, computer manipulation of graphics is necessary.

1.2.6. Development of MCG

From the explosive impact of the results of the above research released in 1963-64, the fallout is still continuing. The main developments were in widening the applications, the introduction of multiple graphics consoles serviced by a single central processor (a natural outcome of time-sharing) and in more efficient software. The main problems which always cropped up were expense and lack of low-cost graphic consoles; this in turn implied that many people could not enter this field and consequently offer their talents to (hopefully) solve some of these very problems which prevented them entering the field in the first place!

How the field of MCG or CG has grown can be seen by the literature growth since that time. In 1961, in the Proc. IEEE "Special Issue on Computers" (33), under "Advanced Computer Techniques" (34), MCG was not even hinted at; graphics I/O equipment was still "primitive" (35). By 1966, the "Special Computer Issue" (36) contained several good review articles on the subject. In 1967, a "Special Issue" was devoted to Computer Aided Circuit Design (37) with several on-line and display applications described.

In 1963 the Society for Information Display was formed, the words "information display" implying electronic information display.

By 1966 there were already 25,000 graphic display consoles, (actually alphanumeric displays) built or on order in the U.S.; by 1970 (estimate made in 1965) it was estimated 100,000 consoles would be produced yearly (38). By 1973, it is estimated, 13% of the cost of each computer installation installed in a manufacturing company will be for display and graphic terminals.

These forecasts were somewhat optimistic. In 1967, only about 70,000 alphanumeric displays were installed (39), much less manufactured. By 1974, it is estimated that over 250,000 would be installed. Of these, interactive graphics consoles would, at best, number only in the several thousands.

1.3. MAN-COMPUTER - CIRCUIT - DESIGN

1.3.1. Introduction.

Most of the development of "graphic console" hard-ware and its testing out was carried out by electronic engineers. By nature designers, their common hardware design language is in the large, block diagrams and in detail, circuit diagrams, made up of standard symbols indicating resistors, capacitors, inductors, transistors, voltage and current sources and the like.

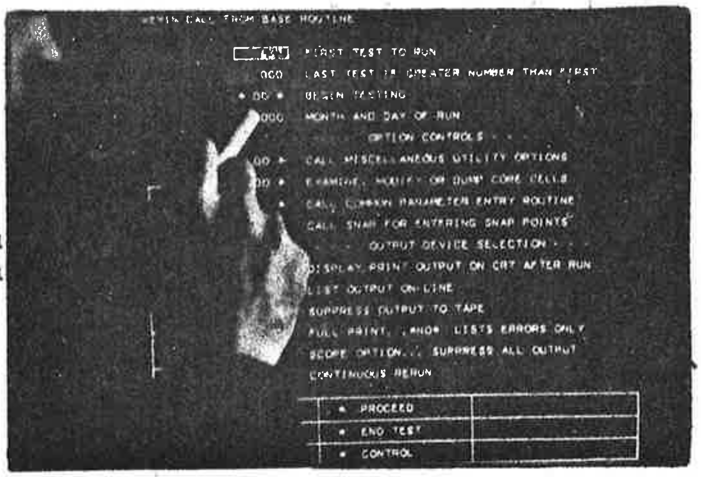
Military electronics and consumer circuits demand high reliability and thus require sensitivity analysis i.e. the variation in the performance of the circuit when one or more components vary from the nominal value or are degraded by envioronment, usage and time, need be determined. Such calculations are time consuming and repetitive; results for comparison, often need be in graphical form. Computer-aided-circuit design and graphics seemed to provide the answer.

Further, in the design of circuits themselves, the design cycle is, in the main, identical to the design processes mentioned previously. Man-computer-graphics again suggested itself.

What was envisaged (40,41,42,43,44) was an engineer sitting at a graphic console: with a light pen

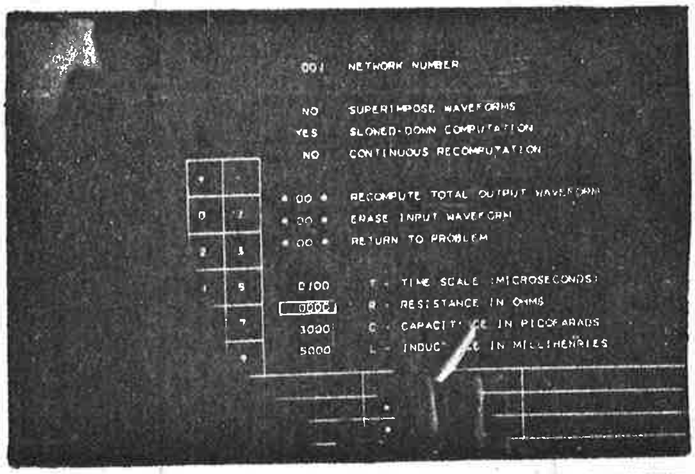
The user selects from a 'PROGRAM Menu' the sequence of steps to be executed as his program(see section 2.8.4.1). A light pen is used for selection

(a)



A selected subprogram is displayed enabling a required circuit ,in this case a low pass RC circuit, to be retrieved from store.

(b)



The driving waveform has been drawn by the user and the component values(which may be varied by touching the variable arrows) nominated.
The resultant output waveform is displayed and evaluated.

(c)

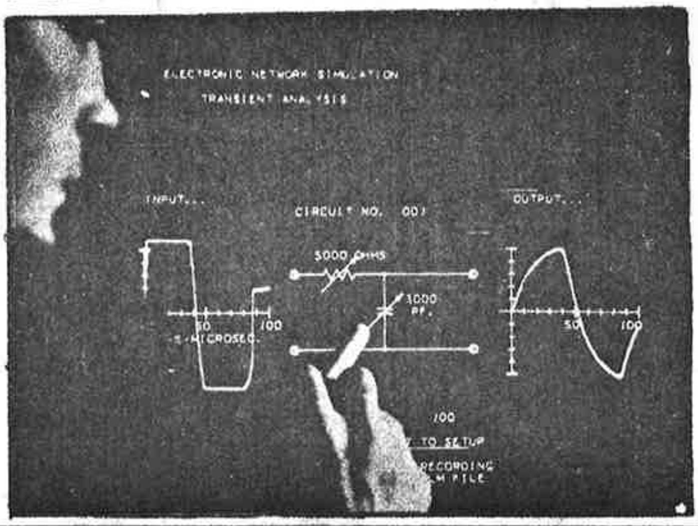


Figure 7. Man-Computer-Graphic-Circuit-Design. (from [21])

he points to a spot on the CRT face and presses a button labelled "transistor" say. A symbol of a transistor appears on the location indicated. Successively indicating locations and pressing keys marked "resistor", "current source" etc he draws a complete circuit of say 20 components, active and passive. He specifies an input waveshape, a sinusoid, its frequency and amplitude, indicates a circuit node at which he wishes to examine the voltage waveform when the circuit is excited with the specified input, and pushes a "compute and display" key. Almost "instantaneously" the required voltage waveform appears; it is not what he has expected. He inserts a decoupling capacitor, changes the value of a resistor by pushing appropriate keys, and again demands a result. The circuit this time is satisfactory. He then specifies component size (either directly or by type number) and demands a printed layout and connections, having specified the position of the input and output. A display of the layout appears and a hard-copy together with dimensions etc and component values is made. In less that 15 minutes a complete circuit has been designed, and laid out: sensitivity analyses may then be made (pushing the "sensitivity" key which varies each component value by ±5% from its nominal value, computes the output values and displayes the results). The Development Lab. mocks it up and the Production Dept. initiates a production run.

The above scenario is already a reality (see Fig. 7) but with limited success and spread of usage, mainly due to higher than anticipated hardware costs and lack of development of efficient applications programs (circuit design, sensitivity analysis etc).

The resultant extensive program libraries required can only be implemented on the largest of existing computers to perform the above form of circuit design. Indeed where it has been commercially implemented (45) (other than in purely research institutions), it is on a semi-closed-shop basis, with the clients being smaller manufacturers of electronic circuits.

How optimistic were the forecasts for MCGCD, it is only necessary to quote the following statement made in 1966:

> "The moral of this advance (in MCGCD) should wallop design engineers on the head - i.e. 99.99 percent of engineers who are designing circuits without on-line graphical-language facilities are, in one sense, already "living in the past". (46)

By all accounts, four years later (or double the life length of MCGCD as it was in 1966), 99% of circuit design engineers (excluding I-C design engineers) are still "living in the past"!

1.3.2. Available Computer-Aided-Circuit-Design Programs

A vast number of so-called "computer-aided-circuit design" programs are available and comprehensively documented. Appendix 2 is a short overview of such available programs.

These so called "circuit-design programs", however, are circuit analysis programs, applicable to a-priori defined circuit configurations, that is, the component values and circuit topology are presented as input for processing; the program output is a print-out of currents, voltages, powers, current gains etc, at specified nodes or branches. Such programs are equally applicable to batch processing and indeed most of them have been written and designed for this purpose. Complex in structure and large in size (some requiring 32-48K of core store), they perform only the "analysis" section at the circuit design process (Fig. 8). For true Man-Computer design interaction, a number of other programs are necessary, namely those enabling the user (who may be no programmer) to input tentative circuit configurations and component values in the manner described above; other programs enabling the user to delete components and generally break into a computation or analysis cycle; programs enabling circuit optimization to be performed; purely graphic design programs enabling optimum circuit lay-out to be performed and so on, and supervisory programs to control the interaction of these programs between user and computer; the problems, complexity and size providing these interactive programs are touched upon in section 2.8.

To satisfy some of the requirements on existing computers and display hardware, and to allow interaction during circuit analysis, the circuit analysis program has to be "cut to size", with the result that only small circuits thus far have been designed and solved interactively. Another development resulted in the increasing use of small peripheral computers (section 2.6.2.) to take over some of the burden off the CPU.

Appendix 2 concerns itself mainly with circuit analysis programs and the problems encountered with them.

1.3.3. Man-Computer-Graphics Circuit Design (MCGCD)

1.3.3.1. Introduction

To illustrate a typical successful on-line graphic circuit design system, GINA (Graphical Input for Network Analysis) is briefly described (47); other on-line curcuit programs are very similar (40,42,48,49,50,51,52) and are tabulated in Table 2 in Appendix 2.

1.3.3.2 Characteristics of GINA

GINA features the following:

(a) Only frequency response of circuits is computed using CALAHAN's program (197). "Gain"in dB vs frequency" is the presented result. (Such a lack of generality in analysis is a common feature of current MCGCD programs).

(b) The user requires a basic knowledge of circuit synthesis and of the system but no detailed programming knowledge.

(c) The operator uses a console and a light pen. The allowable components (R,L,C and current source) and several simple functions such as "Select node", "Select branch","Analyse", "Exit" etc are available as "light buttons", which are displayed on the CRT face, and are activated by pointing the light pen at the appropriate "light button". They replace a keyboard, and have been found to be more effective, user-comfort-wise.

(d) Simple circuits can only be designed i.e. having limited types of elements and with a limited number of components. Active circuits can be modelled by simple current-source models (transistors at least anyway). 20 elements would seem to be the size limit of the circuit, due to display area size and available store limitations. Some 20,000 words of store are required for the generation of display, analysis files, and programs. The display unit is a commercial CDC-280 while the computer is an IBM 7094.

(e) The sequence of design is roughly as described in the ideal case previously (see section 1,3,1).

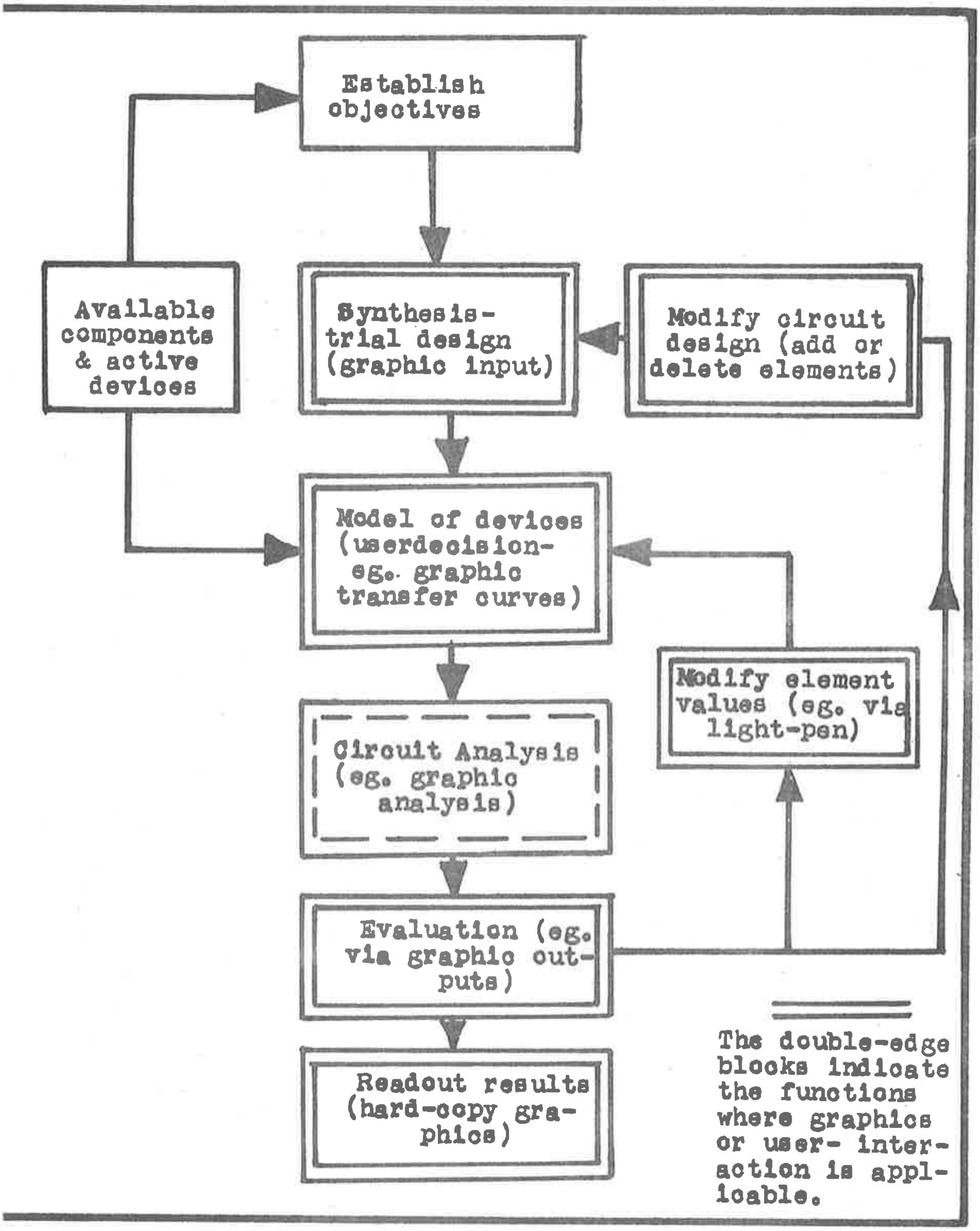


Figure 8. Simplified Design Cycle in Circuit Design.

The options or "moves" available to the user are limited by the light buttons available. After each action, say nominating a node (via the "Select node" button), which cues the "positions" cue (by-light-intensifying it) the node is positioned by the light pen; another "select node" appears, giving finally 2 nodes. "Select branch" is indicated, and an element is "picked" from the available element repertoire. "Connect" and then "Select value" appears. The value is picked from a numerical light button board displayed after the cue "Select value".

When the network is completed, the option "Analyse" is picked,resulting in the frequency response being displayed. During this time and until the output graph of "Gain (dB) vs Frequency" is generated, the cue "Running" is indicated. If the response is satisfactory a hard-copy is generated; other-wise the circuit is modified.

A flow diagram of the system is shown in Fig. 9, together with examples of typical input and resultant displays.

1.3.3.3. Advantages and Disadvantages of GINA

For the user, the following advantages are apparent :

(a) No programming experience is needed.

(b) The available options at any one time are strictly limited by the cues being given, by the intensifying of the next action to be performed according to the resident program. This prevents unallowed "moves" to be made or accidental moves, such as are possible when keyboards are present. The user is thus guided through a problem.

The disadvantages are more obvious:

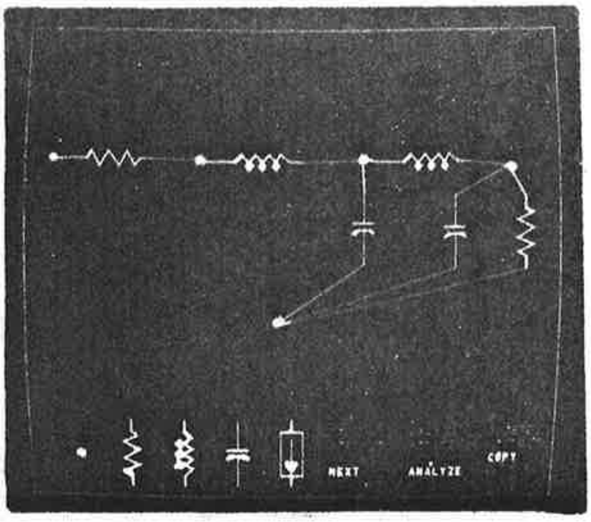
(a) There is a strict limitation in type of elements available (4 only) and in size of circuit to be analysed (20 - 30 elements). The "reason" for the latter is sometimes put forward (42) that display area size is limited and cluttering of the display will occur if circuits are too large, leading to user confusion. One suspects the real reason is the amount of required storage at the various levels, required for programs larger than the above with a larger repertoire of circuit elements.

(b) For the same reasons, only frequency analysis is catered for. To provide all of the available analysis options such as DC,AC, Transient etc with existing programs would be economically impractical to implement.

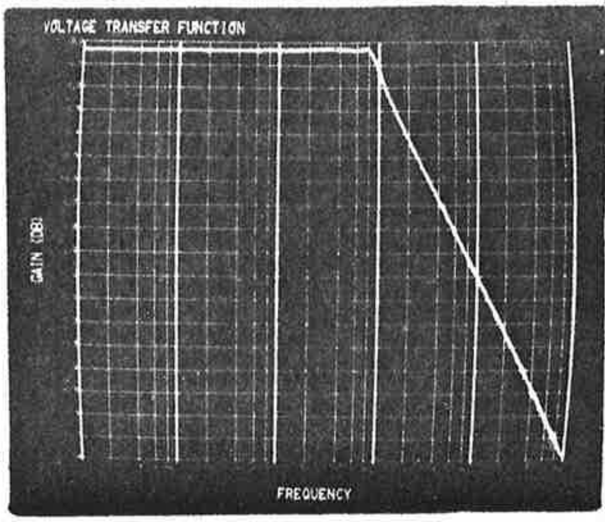
1.3.3.4. Other MCGCD Programs

Some existing MCGCD systems are shown in Appendix 2'; the tabulation is by no means comprehensive - new programs are appearing all the time.

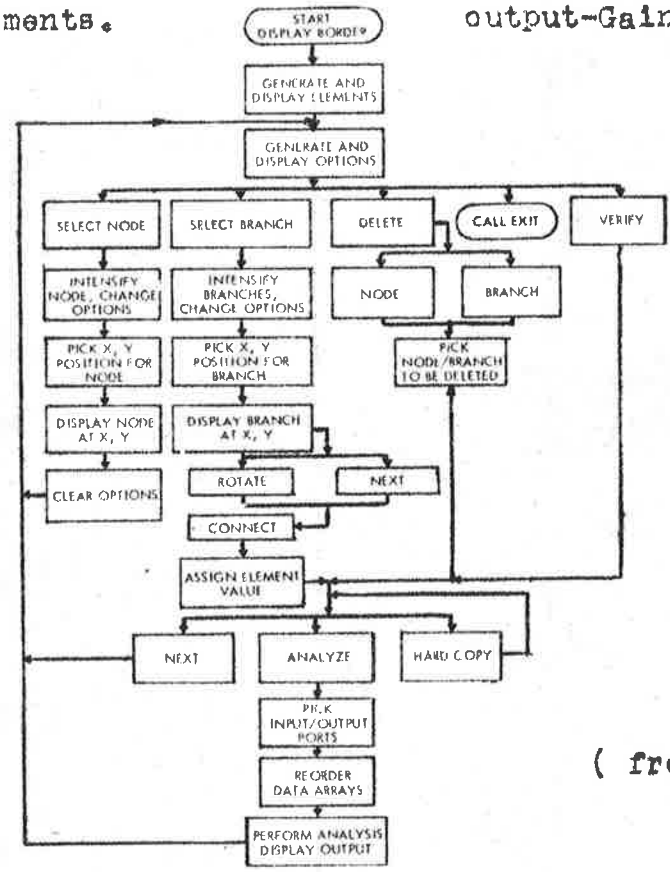
Mention must be made of what has commonly been classed as "on-line circuit design", namely the lay-out of the wiring in printed circuit cards and Integrated-Circuit lay-outs. Although an important aspect of circuit design, these strictly are problems in topology and 2-dimensional geometry and precision graphics rather than circuit analysis and design.



(a) Display of circuit and available elements.



(b) Display of required output-Gain vs Frequency.



(from [47])

(c) Simplified flow-chart with emphasis on user-interactions.

Figure 9. GINA (Graphical Input for Network Analysis)

1.3.4. Graphical Solution of Circuits

1.3.4.1. Problems of MCGCD

Compared with design problems solved using the Sketchpad approach, MCGCD has been relatively "unspectacular". As pointed out in Appendix 2, most of the causes of the limitations are due to the existing methods of circuit analysis being used, which are still, in the main, numerical. Matrix manipulations, whether in numerical or symbolic form, are still the stuff of circuit analysis. Graphics is used only to define the type of circuit elements in a symbolic form and to specify the circuit topology; the output results may also be presented graphically such as the transient response or gain vs frequency plot as in GINA. The graphic design elements are not functional but only symbolic. The element properties defining the functional relationship between parameters such as voltage-current relations, gain-frequences etc are specified by subroutines and generated and processed numerically.

In contrast, in Sketchpad the graphic elements were not only descriptive but also functional - their processing was done graphically in the sense that certain geometric constraints had to be obeyed (even though the actual computations were performed numerically).

This is the crux of the problem - for the efficient use of graphics in Man-Computer Design problems, graphics must not only be used for communication but also the geometric properties of the graphic elements must be the basis for computation.

1.3.4.2. Device Modelling

MCGCD is further complicated in that active elements (as well as some passive elements) are inherently non-linear in their characteristics; problems of accurately representing or modelling such non-linearities arise (53, 54,55), not to say of the difficulties of their manipulation during computation. Feasibility of solution may even need to be established (48).

Most network analysis programs usually model transistors by generating the Ebers-Moll equations with the appropriate coefficiants for each particular transistor inserted. Diodes are similarly modelled. For example, NET-1 (189) uses 36 parameters to insert into the Eber-Moll equations to describe a transistor; ECAP (193) however, does not accept exponential functions, resulting in transistors being represented to various degrees of complexity by parallel branches with dependent or independent voltage or current sources. AEDNET (48), specifically for the simulation of non-linear networks, can accept either piecewise linear approximations specified by a table of values (particularly for non-anlytic describable functions) or, for analytically-describable functions, by subroutine function generators. Many programs use piecewise-linear approximations for transistors or diodes to varying degrees of complexity; thus, in the simplest case, a saturated transistor may be represented by a small resistance. The same may be said for describing the exciting signals - ramps, sinusoids, exponentials and "square" wave-inputs are specified analytically, but most of the larger programs accept piecewise-linear inputs described by tables.

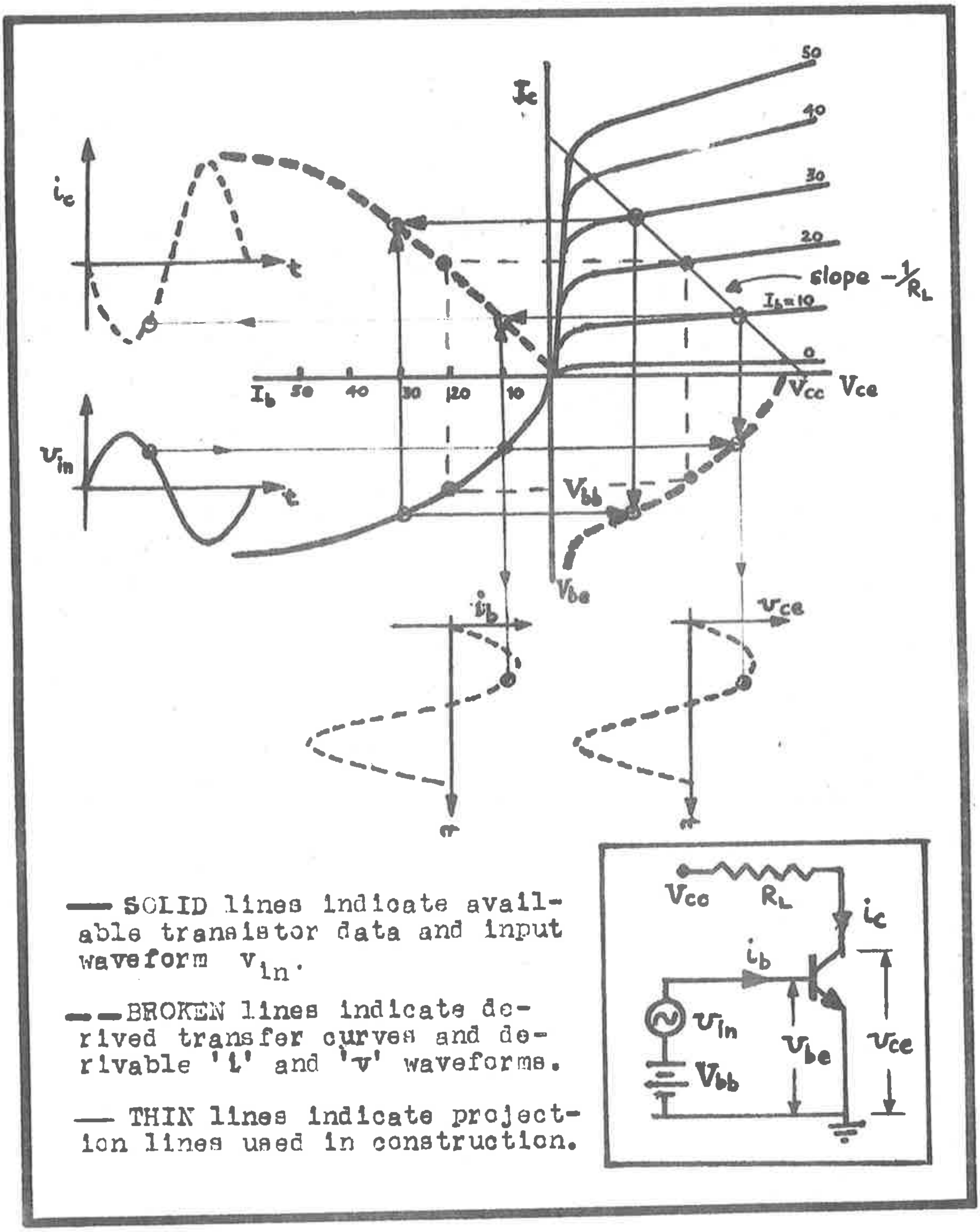


Figure 10. Graphic Solution of Hypothetical Nonlinear Circuit.

1.3.4.3. Graphic Solutions of Circuits

A quite common way of dealing with non-linear components (and sometimes the only way) is by graphical techniques. (Linear components, being trivial cases of non-linear components, quite clearly can also, if required, be solved by graphic techniques). Analogous to an algebraic transfer function of some circuit from which the output is obtained by specifying the input, so in principle, the graphic representation of such a transfer function (the transfer curve) can be obtained, by geometrically manipulating the individual transfer curves of each element. The output then is merely the point by point geometric projection of an exciting input (no matter how non-linear) against such a transfer curve; solutions then are dependent on certain geometrical constraints being obeyed, which is reminiscent of the Sketchpad method.
A simple example is shown in Fig.10.

Examples of describing functional relationships of elements graphically for a resistor, say, is a line of variable slope specifying the magnitude of resistance, passing through the origin of a set of axis representing voltage and current. More complex elements or subcircuits whith known complex transfer curves can be treated by transfer curves entered manually via a light pen or other graphic input devices (see Appendix 5); complex input waveforms may also be entered manually. The specification of elements and waveforms by such graphical forms thus no longer restrict them to man-machine or graphical communication, but also being functional relationships, make them directly amen able for computation.

Such a geometrical-based scheme of computation may seem very simplistic and without its problems but it is an avenue which warrants further exploration; it may provide new leads to overcome the many current problems such as storage problems, matrix inversion, and lack of user-computer interaction during computation (as distinct from the I/O interaction) amongst others.

Graphic computation is by nature less accurate than numerical computation; input device and display non-linearities further increase the inacuracies (the equivalent of using a curved "straight-edge" or non-linear scale in manual graphical computations!). However, during the design stage, where component values are varied or components deleted and wholly different designs with new sub-circuits are used, great accuracy is not strictly necessary - 5% accuracy may be acceptable. As soon as a design looks promising or satisfactory (the user being experienced enough to decide this),or even feasible (as in the case of several non-linear components being present), the resulting network may be numerically analysed with the existing techniques; sensitivity analyses etc may be made.

A division of techniques, labour and results, is necessary. Graphical and not highly accurate, but highly interactive user-computer methods can be utilized during the design stage, followed by numerical but precise and relatively off-line methods during the final stages of the design process. With less stringent accuracy demanded during the interactive user-computer stage, where time-sharing must be used for reasons of economy, the problems of dynamic memory allocation in various levels of storage are greatly reduced. Note that memory and memory allocation are one of the prime reasons for current MCGCD limitations.

1.4. CONCLUSIONS AND COMMENTS

The above approach of graphical computation and graphical specification of elements was one which interested the author when he began his studies. A suitable graphic console with highly interactive user-computer I/O capabilities was required. A cursory study of available interactive graphics consoles and their economics (or lack of economics!) precluded the purchase of any such consoles - starting prices were in excess of $30,000! Without the hardware, the graphical computations study would have been at best tentative, and at worst, inconclusive. The decision was taken to come up with an economic, low-cost,versatile interactive graphics console, capable not only of accepting and displaying graphics, but also of performing graphical computations. Of necessity new concepts had to be used, otherwise the result would be similar to existing graphic consoles, with similar costs, and nothing new would have been gained. The thesis explores these new concepts.

The emphasis thus far has been primarily on Man-Computer-Graphic-Circuit Design. This is only natural as of the design processes current, circuit design is perhaps most familiar to electrical engineers. However Man-Computer-Graphics has broader applications than MCGCD or even the other applications mentioned. In Appendix 1 is a table listing applications where these interactive graphic techniques have been or may be used; the classification into types of application is personal and makes no pretence of being comprehensive. However, the scope of applications indicates the growth and increasing acceptance of Man-Computer-Graphics.

Examining the table in Appendix 1, one can only agree with the conclusion that MCG -

"...opens the horizon to the solution of a large class of problems where solutions should be directed or guided by human judgement" (4).

......

SUBSCRIPT

Very recently, a paper ("The case for a generalized graphic problem solver," E.H. Sibley, R.W. Taylor, W.L. Ash, Proc SJCC, 1970 pp11-17) appeared, commenting on the aparent failure of efforts in developing further the very promising graphics computing systems such as SKETCHPAD and DAC-1. The reason, according to the authors, is the failure to realize (and thus to implement!), that a "generalized graphic problem solver" is required, which is defined as a

"software computer system which aids the engineer, scientist and mathematician in the formulation and analysis of a range of problems using the heuristics of the man to formulate a reasonable model, and the computer to aid in and augment the analysis".

The "generalized graphic problem solver" consists of two interlinked parallel loops; the first loop, common to even the most primitive man-graphics system, consisting of the user, an input device and input routines, some form of data representation and an output device and output routines; and a second loop, between the input routines and the output routine, where problem constraints (graphical constraints) routines, problem rules, analysis etc are carried out.

The presence of this second loop is the core of the "generalized" graphics problem solver, and the inability to realize it means the inability to realize graphics problem-solving along the lines forecast by SKETCHPAD etc.

A consequence of that system structure, the authors state, is that the man-machine interaction is limited only to the I/O problem phase, and is non-existant during the problem solution, as, in the second loop where the problem is essentially being solved, the problem representation may be in a form "unintelligable or untranslatable to the man".

The full value then of an interactive graphics solving system appears to be in question, as one of the prime advantages is always stated to be the ability of man-machine interaction during the problem solving stage. In contradistinction to this, our original approach was to make the user, being directly responsible for the problem, the "backbone" and the prime element of this second loop, where the problem is being solved, rather than leaving it all to some difficult-to-realise machine-implemented "general graphics problem-solving" routines.

The user thus appears in both loops, and can interact with the problem not only during the I/O stages of the problem, but during the problem solution stages. This requires a very fast, highly interactive "graphics working pad", perhaps analogous to a desk-calculator in numerical problems, where certain graphics problem constraints and solution feasabilities could be carried out, without necessarily high accuracy. One or two solution approaches could thus be "zeroed" on, which could then be solved more conventianally with off-line methods, with the necessary accuracy. This of-course necessitated the feasability study of VIDIOGRAPHIC, the proposed economical, fast highly interactive I/O graphics console.

......

CHAPTER 2

THE INTERACTIVE GRAPHICS CONSOLE (IGC) REQUIREMENTS AND EXISTING IMPLEMENTATION

2.1. INTRODUCTION

2.1.1. General

The user or designer who solves programs interactively with a computer sits in front of a desk or console which has various units such as a display screen, teletype keyboard, and various graphics input devices and function keys by which the user communicates with the computer during the interactive problem-solving phase. Such a console with the assembly of Input-Output graphics devices and communication devices is called an "Interactive Graphics Console" or IGC (See Fig.1).

Computer Graphics is still in its infancy; and a corollary to this is that, although the desired characteristics of the "all purpose" graphics console are commonly agreed on, there is considerable controversy as to how this ideal "all-purpose" console should be realized. This stems from the fact that workers from many disciplines, with different research aims, formulated whatever concepts there exist today. Secondly, as has been pointed out (56), standards in hardware and software are very difficult to implement or even formulate when the field is so open for commercial exploitation; standards have been undercut to cut costs.

Industry with a rapidly expanding market in their sights (39,57) took out patents on everything that showed promise; to circumvent existing patents, new devices, often quite exotic,were suggested and implemented (particularly in the graphics input area). Although at first sight this appears to have stimulated research, the permanent result was high cost of the resultant hardware to cover R and D expense.

This led to a second development.

High cost of hardware and low sales attenuated applications research and software developement to the point where recently a symposium was held, questioning whether "Computer Graphics was dead" (58); although the concensus of opinion was to the contrary, the main problem was pinpointed as the lack of software development and software standards (the symposium was organized by the ACM).

Hardware costs remain high. For example, graphic input tablets which enable hand-drawn information to be input to the computer, start around $2,500 (US) (60). Alphanumeric display consoles are classified as "low-cost" if the price is below $20,000 (59).

Few firms provide complete interactive graphics systems, that is, graphics display, graphics input (including hand-written input), graphics storage, user interaction facilities and associated software; most firms specialize in this or that graphics subsystem, such as displays, grahic input devices, vector or character generators and the like, i.e. on devices for which they hold patents. Only the larger and expensive systems are provided with software support (61,62,63). Software being a most elusive patentable and saleable product has been, as mentioned above, least developed. With the prospect of purchasing expensive hardware with relatively inefficient and inflexible software (witness the DAC-1 program structure (11,12)), it is not surprising that management invariably still considers MCG as too expensive to implement.

This chapter deals with the requirements of a Man-Computer-Graphics console (or "Interactive Graphics Console" (IGC)) and examines some current implementations. The emphasis is, in the main, on hardware and physical parameters; software aspects will be kept to a minimum, and then only mentioned with respect to user-console interaction and CPU-IGC interaction. The structure of programs responsible for these actions are outside the scope of this report. Each available program and language used for its implementation differs from each other.

The length of this chapter is due to two reasons. The first is the complexity of the requirements for an IGC; the major approaches to implement these requirements are indicated. The second reason is that there are few, if any, basic references one can turn to for an overview of the field (the display field is the best documented). MCG is still a new field, with few standards published or even agreed on. The IGC parameters and requirements presented here have been culled from many references.

There is no claim of this chapter being thoroughly comprehensive. For example, the section on software is very sketchy at best, not only because of the hardware emphasis being predominant, but because to do justice in describing even one MCG program in some degree of detail would double the length of this chapter. It is believed, however, that the chapter together with its references perform a useful function in this thesis.

2.1.2. Interactive Graphics Console Functions

An IGC must be capable of:

(i) accepting graphic and alphanumeric information from the user or the Central Processing Unit (CPU).

(ii) displaying graphic and alphanumeric information when requested by the user or pre-requested within a program (eg. intermediate results or trial solutions).

(iii) between initial input specification of the problem and final display of results, allowing user interaction, displaying any intermediate results as requested and hence permitting user control over problem-solving procedures.

(iv) lastly, it must be economical in initial cost outlay and CPU usage, power consumption etc.

"Economy" must be understood to be within the context of the application for which the IGC is to be used. Clearly tasks of the magnitude for which DAC-1 was designed cannot be compared with the tasks of an alphanumeric displays console for management display; neither can therefore the outlay nor running costs be compared in the two cases. Other factors to be considered here are reliability and ease of maintenance.

The above capabilities must be so implemented as to minimize what is termed "user-discomfort", and maximize graphic console and CPU efficiency ("CPU-comfort"?) (e.g. by minimizing CPU graphics processing time, graphics storage, display generation and maximizing hardware-software tradeoffs etc). As in the case of most man-machine relationships, the requirements are often contradictory. "User-comfort" for example involves such factors as size and location of display, display brightness, contrast, flicker "naturalness" of inputting of data etc. On the other hand, available memory may preclude "flicker-free" display or give "intolerable" (several seconds!) display response time to user interaction; inherent display structure (eg electro-luminescent displays) may mean low brightness and low contrast etc. Thus far, no IGC has appeared on the market satisfying the above four requirements simultaneously, particularly the requirement of low cost and economy of use.

Below, displays and display factors are first described. Graphics input, including user-interaction, follows, while overall IGC concepts such as IGC configurations, hardware-software tradeoffs, memory requirements etc. are also mentioned. As no standards have yet been specified (except to certain standards relating to alpha-numerics), the following pertinent factors involving displays, input-graphics, vector and character generators have been taken from various sources, some for proposed IGCs, other for IGCs already implemented.

2.2. THE USER - DISPLAY INTERFACE

2.2.1. Requirements

The requirements for the display interface are considered from two viewpoints:

(1) user or viewer-comfort considerations, which arise when a display is viewed by a user. Such factors as display brightness, contrast, and display size determine whether a display is "satisfactory" as far as the user is concerned. Such factors, being pschycological, vary from user to user; however certain quantititative figures apply to the majority of users.

(2) Display capabilities which concern themselves with the realizable quality of the resultant display. These include resolution, distortion, addressing resolution, repeatability etc.

In addition to satisfying the above factors, display "dynamics", which include the speed of response to display updating, and the presence or otherwise of "flicker", must be taken into account; these depend on the actual class of display being used.

2.2.2. Display Parameters

2.2.2.1 Resolution (9,64,65,66)

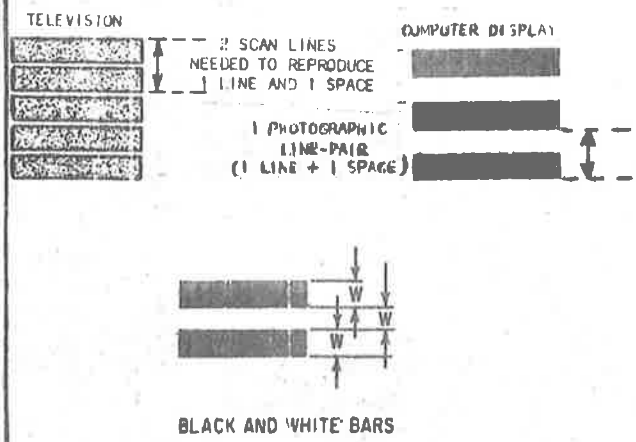
Resolution is a measure of spatial discrimination and is thus concerned with the amount of detail perceived by the viewer, and the detail generated on the display.

(a) Viewer resolution

Lack of standards precludes a quantitative definition particularly since resolution varies from viewer to viewer and also with contrast, brightness, and duration of perceived object. Particularly, it depends on the shape of the object; thus for example :

(i) two parallel black lines on a white background can be visually separated if the angle of view is about 0.5-1 minute of arc. At viewing distance of 24" this separation is of the order 3-6 mils (Fig.11(a)).

(ii) a white square on a black background can be detected if it subtends an angle of about 10 seconds of arc at the eye. This is about 1 mil when viewed from 24". A black square on a white

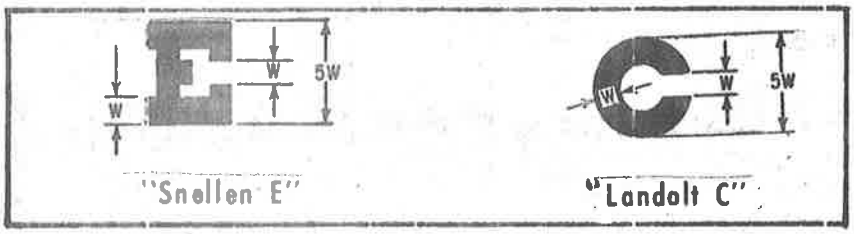


Resolution R_s in line-pairs

$$R_s = \frac{n_b + n_w}{H}$$

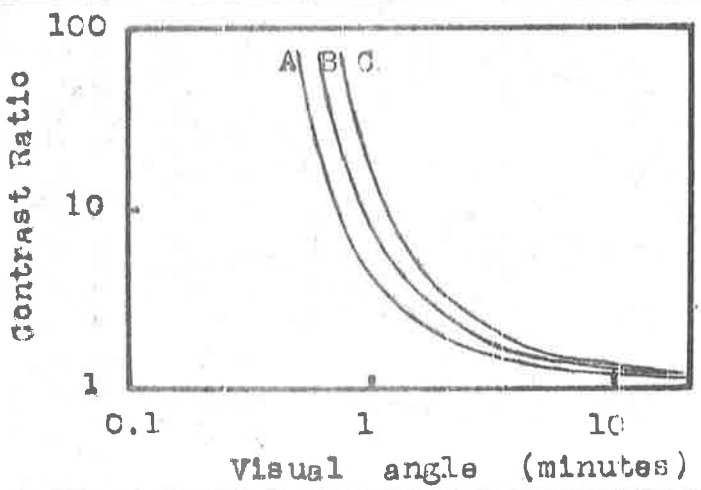
where n_b = no. of black lines
n_w = no. of white lines
H = height of total lines

(a) Visual Resolution Targets - Line-pairs and TV lines



(b) Visual Resolution Targets- Alphanumerics

Figure 11. Resolution Concepts



Display Brightness

A = 100 fL

B = 50 fL

C = 1 fL

Figure 12. Resolution as a function of Contrast

background on the other hand requires an angle of about 30 secs!

The ability to describe the shape of a complex pattern gives yet another value for resolution. This is of interest when reading alphanumerics (Fig.11(c)). Alphanumerics can be generated by inscribing them by dots (or line segments) on a 5 x 7 dot matrix. Using the "line" pair resolution figure, a character of 20 - 40 mils should be resolved. In practice a character height of at least 0.1" (100 mils) is necessary. This of-course is partially due to the spot size or line thickness available on displays which in practice is about 5 - 30 mils in width.

Thus, in practice, hardware limitations, rather than the above factors, govern the actual resolution of the display.

(b) Display Resolution

Ideally the detail capable of being resolved, and generated by the display image-forming mechanism, should be of the same order as that resolved by the user. Thus the display elements should be able to generate "line pairs" separated by 3-6 mils. In ultra-precision CRT's, spot (and hence line) diameters may be 2 mils (67); more usually spot diameters are 5-30 mils in diameters. Defocussing or increase in spot size due to different spot locations on the display area (see section 9.3) can further increase spot size by a factor of 2-3 unless corrective measures are taken. Resolution figures quoted are usually at "best focus", i.e. at centre of the display. In other experimental displays, such as electroluminescent panels (sec.2.3.7.3.), resolution is much lower than the above, being of the order of 40 - 60 mils (68).

Display Resolution may be artificially increased by software and CPU capability. Thus if within the CPU coordinates are given to 16-bit precision, whereas for display positioning they typically are given 10-bit precision, then the remaining 6 bits may be used for magnifying, "scaling" etc; and if display resolution is inadequate to show the small "fine" detail, scaling may then be used (e.g. shift each coordinate by the number of bits corresponding to the scaling required).

2.2.2.2. Brightness and Contrast (9,64,65,66)

(a) User Factors

A user perceives an object (not to be confused with "resolves") if there is a difference between the object brightness and background brightness. Brightness difference leads to the concept of "Contrast Ratio" which is given by

$$\text{Contrast Ratio} = \frac{\text{Information brightness}}{\text{Background brightness}}$$

Thus brightness and contrast are intimately linked. Brightness itself depends on ambient light, contrast resolution, duration of "flashes" from which the display is generated etc. It is measured in foot-lamberts (ft-L) (see Appendix 8.1). Contrast is considered "adequate" if it is greater than about 4:1 for white symbols on a black background, and needs be at least 25:1 for line drawings or text on a white background. Figure 12 shows the typical dependence of resolution on brightness and contrast ratio.

Brightness of a printed sheet in sunlight is about 30ft-L. A TV-raster display is typically 15-20ft-L for acceptable use; in a darkened room it can be lowered to about 5 ft-L with no viewer discomfort.

(b) Display Factors

In Cathode Ray Tube (CRT) displays, brightness depends on the phosphor used, the electron beam current, the structure of the faceplate and whether any polarizing "shield" is used to reduce glare and reflections (69).

Too high contrast and brightness leads to "dazzle" and loss of resolution. Ambient light and beam current are controllable factors; however too high a beam current greatly shortens CRT tube life. E-L panel displays have, for a typical displayed image, a brightness of about 20-30 ft-L; this may seem adequate but contrast is very low-in normal ambient light conditions, users have reported inadequate contrast. Neon-Plasma panels (section 2.3.7.4) have been reported to have brightness of up to 100 ft-L (70). Storage CRTs (section 2.5.4) have brightness and contrast within the accepted range; these however degrade with time, until after several minutes (mesh-type CRT) or several hours (DVST), the contrast ratio approaches unity i.e. the display merges into an evenly illuminated background.

2.2.2.3 Flicker and Display Refresh (9,64,65,74)

Displayed images, to be assimilated and comprehended by the user, need be maintained and displayed up to several minutes, say; this means that energy expenditure occurs to maintain the display. This energy must be supplied selectively only at the point where information is displayed. Storage of this "selective energy expenditure information" is required. In other words the display information must be stored.

The supplying of the energy to the required display location is called "refreshing". Thus an external memory need be provided and depending on what form this takes and how often the display is refreshed, this determines whether "flicker" is present.

Flicker is the awareness of the user to the refreshing of the display and makes its presence felt by a pulsating increasing and decreasing of the display brightness at the refresh rate. This not only causes viewer annoyance and irritability, but also loss of concentration (and even physical headaches!).

In TV-type displays whole images are refreshed (2 interlaced frames) at 25 Hz (here in Australia); the resultant is a "steady" image as far as brightness goes. It is due to the integrating effect of the eye on the continually refreshed images; even though their individual brightnesses are varying continually, the variations occur too fast for the eye to detect. Contrary to expectations (see Fig. 18), longer persistance phosphors still require relatively high refresh rates (15-20Hz).

(a) User Factors

It has been found that refresh rates below 15-20 Hz can cause extreme discomfort. Acceptable refresh rates depend on brightness, contrast, and in CRT's, on the type of phosphor used, and even on the size of the display. Charts are available (64) which give acceptable lowest refresh rates vs brightness for varying phosphors (Fig.18). For most common phosphors at a brightness of 10 ft-L, for flicker-free effect,refresh rates need be above 35Hz. If such high refresh rates are impractical (data-link speed limitations for example), a phosphor such as P19 may be used; refresh rate at about 10 Hz may be tolerated.

(b) Display Factors

The presence of flicker depends on the way the display refreshing is carried out. It is present in CRT displays because the phosphor persistance after excitation lasts only for several milliseconds at most, and the refreshing of each displayed point is in a predetermined order. Each point is under excitation for several milliseconds (at most), for say, 30 times per second. In displays which are generated by random-scan methods, the more complex or "dense" the displayed information, the longer it takes to read out the display file and consequently, within a given time, the less often is reading out and hence refreshing carried out. Consequently the more complex displayed images tend to have flicker more noticable.

Pseudo-storage capability of such devices as the Direct-View-Storage-Tube (DVST), Neon Plasma displays, and even E-L panels occurs beacuse each display point is refreshed continuously in a "parallel" fashion and not sequentially. In a DVST, this is done by a "flood beam" illuminating the whole display area, while in the Plasma and E-L panels, each element has its own memory provision which is initiated indedendently of any other cell. Because each display point is under excitation continuously no flicker is observed.

2.2.2.4. Display Response Time (9,71,74)

(a) User Factors

A user interacts with the display by either inputting new data (via a keyboard or by pointing to some feature on the display or directly by drawing with the light pen) and observing consequent results on the display. The delay between inquiry or command and response should not be "intolerably" long. It may be 1 to 10 seconds long before being considered "intolerable", but it depends on the user and the type of request (the user may not be prepared to wait 5 sedonds for the display of a keyed-in alpahnumeric, but may consider this delay reasonable for rotating a displayed image by some arbitrary angle).

User Response Time may be the partial or total sum of the following:-

(i) Update Response Time which is the delay between the initiation of the inputting of graphical data manually and the display of the result (or readiness for displaying the result). It depends on the CPU processing rate, the priorities (particularly in a time-shared system) and display response time.

(ii) Request Response Time, which is analogous to the above, is the delay between display of request (such as stored graphical data etc) and the request. It depends on speed of memory including access time delays.

(iii) Display Generation Response Time, which is the delay between the initiation of computer output representing the displayed data, and the time of completion of display generation. This is what is of interest here.

(b) Display Factors

As mentioned above the response time is a function of the task to be performed and hence the operations required to perform this task with the system configuration used. Any command and its execution involves the graphics console, the CPU (unless processed within the graphics console), intermediate data storage and the interconnecting data link and hence the total sum of the transmission delay, processing times and I-O response times equal to the user-request response time.

Concentrating only on the Display Response Time, as this is of concern here, the dependent factors are whether raster or random-scan is used (section 2.3.2). With raster-scan the response time is the reciprocal of the refresh rate. For random-scan displays, the response time can be resolved into:

(i) Positioning time (PT)
(ii) Character drawing time (CDT)
(iii) Vector drawing time (VDT)

Position Time (PT) refers to the time required to position the beam to any given location, either to draw a character or specify a vector initial point. It depends on the type of deflection yokes in CRTs, or matrix selection times in E-L and Plasma panel displays. It ranges from 3 - 100μs in presently available equipment (74).

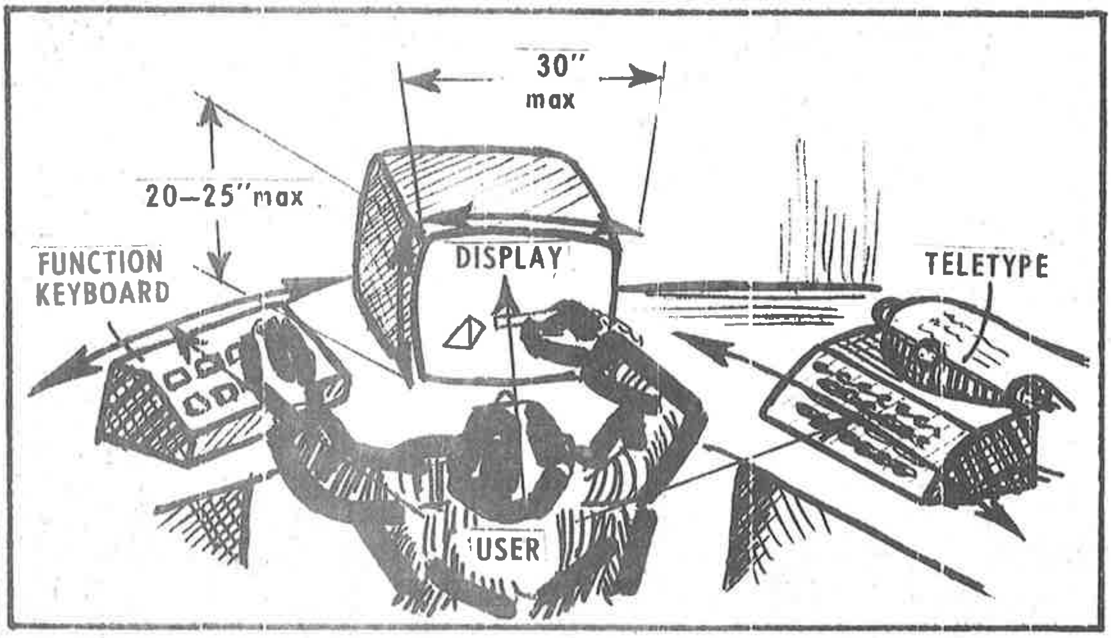
Character Drawing Time (CDT) is self-evident. Depending on the type of character generator (section 2.3.4) this time is from 2-100μs (74). In "stencil" or shaped-beam character generators, CDT is effectively zero (see section 2.3.3.5).

Vector Drawing Time (VDT) again is self evident. Vector generation can be implemented by hardware, by point plotting or by incremental plotting (section 2.3.5). Typical speed can be from 0.5 μs - 150μs per inch of vector in "proportional-time" vector generation and from 3μs to 150μs to draw maximum length vectors (length of display area diagonal) in "fixed-time" vector generation (section 2.3.5.3).

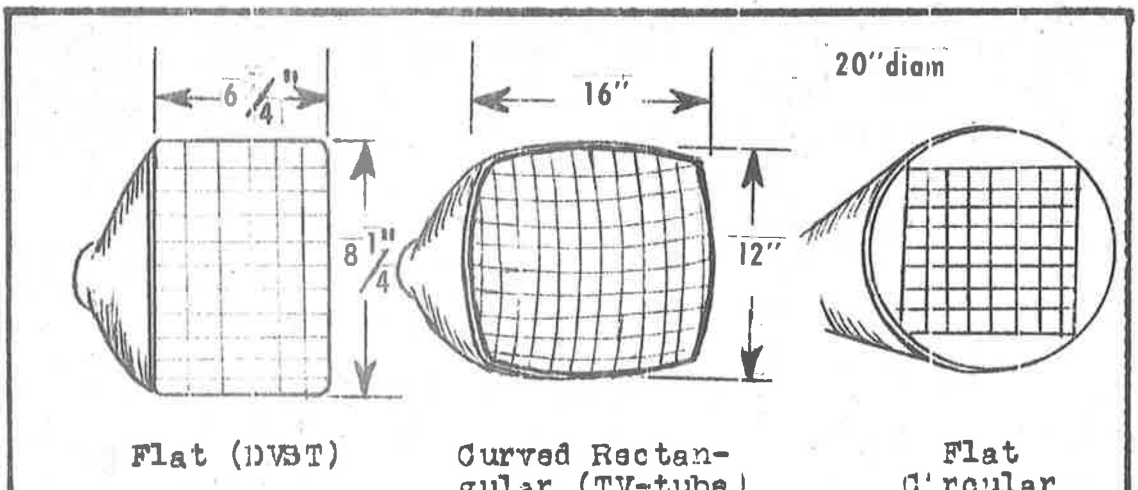
2.2.2.5 Display Area Size and Placement in Relation to User (9,64,74).

(a) User Factors

A user interacting with the display must be able



(a) Layout and size of Display and I.G.C.



(b) Typical Screen shapes and Sizes

Figure 13. I.G.C. Layout and Display Screen Sizes

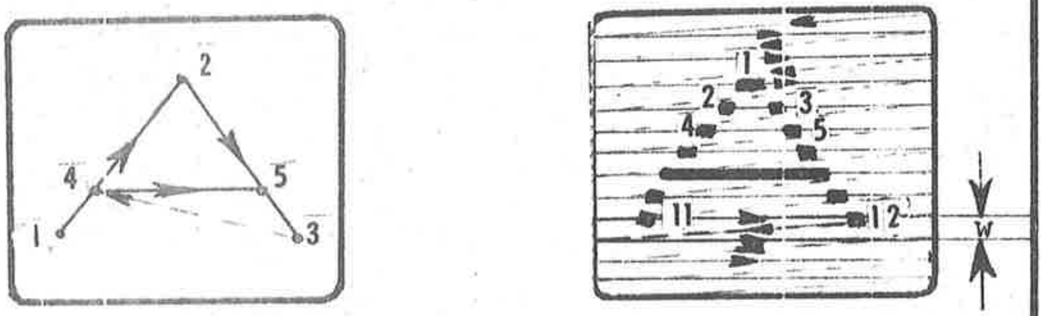
to point to any point of the display with a pointing device with no undue effort (unless a separate input and positioning device is used). With the user sitting down, this limits display area sizes to about 20" x 20" when placed about 24" from the user. The display can be vertical or tilted up to 30^{o} to the vertical.

Larger displays are either "group-displays" with little degree of interaction, or else used in precision drafting.

Placement of keyboards, graphic input pads etc is also important, particularly when visual feedback is used to help locate the placement of the input data. Figs. 1(c), 2(a), and 13(a), indicate the placement and arrangement of the IGC subsystems in relation to the user.

(b) <u>Display Factors</u>

The display area size is determined by economics and by current technology of (in the main) CRT tubes, which restricts the choice to tubes used in TV, radar or oscilloscope applications. Using TV tubes, this means curved-screen displays of 8" x 11" up to 14" x 18"; using flat-screen radar tubes, display-face-diameter tubes up to 24" are possible; oscilloscope tubes are usually too small for display use. The large flat-screen displays are used for precision graphics displays, with the area for display being a square of up to 16" x16" within the circular flat-screen. With spot sizes of 10 mils diameter, and a 2048 x 2048 addressable grid (see below) some spot overlap occurs, which enable lines with sharply defined edges to be drawn (see Fig.15). It is fortunate that the sizes fit very well with user factors such as maximum comfortable hand-movements (for pointing to displayed

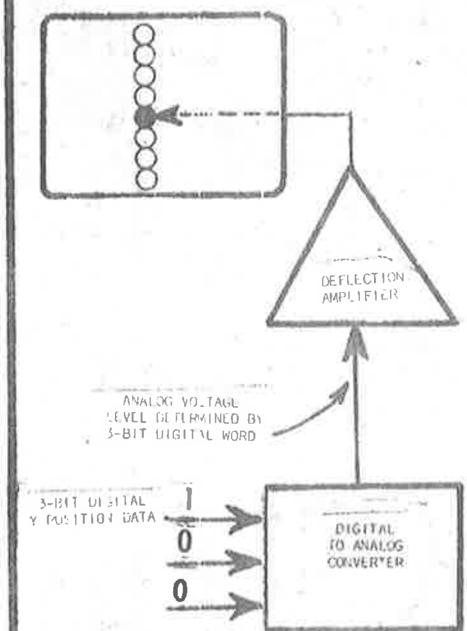


(a) Random Scan of letter 'A'

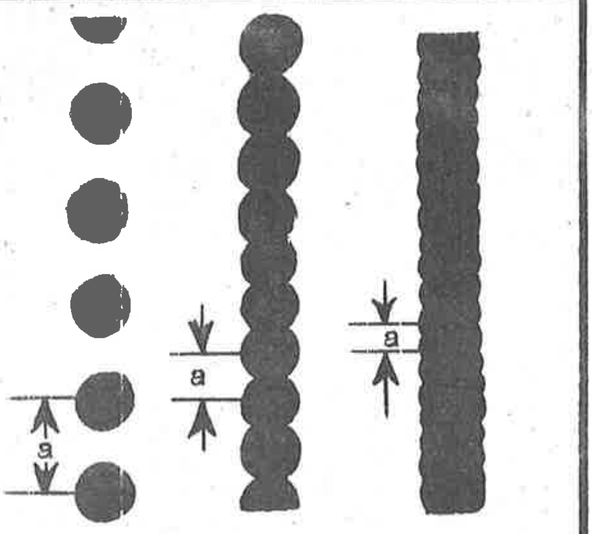
(b) Raster Scan "w" is scanning line width

Numbers indicate order in which corresponding points are accessed.

Figure 14. Random and Raster Scan Modes



Deflection System Addressability



512x512 locations 1024x1024 locations 2048x2048 locations

Line structure on same sized display with constant spot size (and hence constant display resolution) but with different addressability'a'('electronic' resolution)

Figure 15. Addressability and Resolution

information) and for eye movements to observe the display. For large or group-displays, (say 6'x8') direct-viewing CRT tubes are not feasible; projection systems need be used.

2.2.2.6 Addressability and Repeatability (74,75,76)

Interactive displays imply that the user should be able to specify, or access by pointing to, any point or set of points on the display. This means, if it was not made clear previously, that any display information consists of, or is constructed from, discrete elements (c.f. the beam spot in CRTs), each element having 2 coordinates, say the x and y coordinate. (or occasionally r and φ in polar coordinates), each specified by a certain number of binary bits.

(a) In Random-scan displays (see section 2.3.2.1), any location can be accessed randomly (allowing for time delays) by specifying the two x- and y-coordinates from which the appropriate deflection currents or voltages are obtained in CRTs by Digital to Analog conversion. In E-L displays, the x- and y- coordinates directly select two mutually orthogonal driving "strips", the intersection of which specifies the desired location.

(b) In Raster-scan Displays (see section 2.3.2.2.) each location is accessed sequentially, the complete display area requiring to be accessed before the same location can be accessed again. The display area is "scanned" by a succession of horizontal "sweeps" starting from the top and finishing at the bottom of the display, each sucesssive horizontal sweep being one line width (or when n:1 "interlacing" is used, n-line widths lower) than the preceeding. Because of the predictability of the scanning, the specification of the coordinates can be relaxed (compared with random-scan requirements) by using incremental addressing for vertical coordinates and for locations within the same horizontal line.

<u>Addressability</u> (see Fig.15) is the measure of the number of locations capable of being specifically located on the display area by specifying the digital address; it may be considered to be a measure of the electronic circuitry resolution. Typical displays are quoted as having 512 x 512, 1024 x 1024 and even in one case 16384 x 16384 addressable points. Thus, from two 9-bit coordinates to two 14-bit coordinates are used. For the 512 x 512 point display, a discrete matrix of points can be visually resolved; on the other hand, in the "finer" point display, points overlap, with lines with straight, sharply defined boundaries being capable of being drawn (for precision graphics applications).

In raster-scan displays, addressibility is limited to the number of scanning lines present, which usually is based on broadcast-TV standards. In Australia the 625-line system is used, giving about 570 -580 lines for display purposes, (the remainder being for vertical trace return to initiate each successive frame scan); a 512 x 512 element display is thus possible. Non-TV standard-equipment can approach the random-scan addressability.

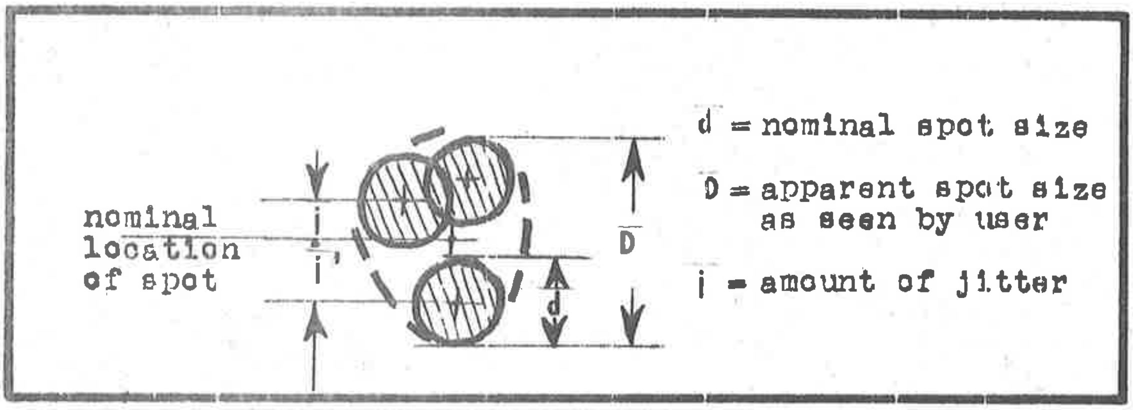


Figure 16. Apparent Loss of Resolution due to Jitter

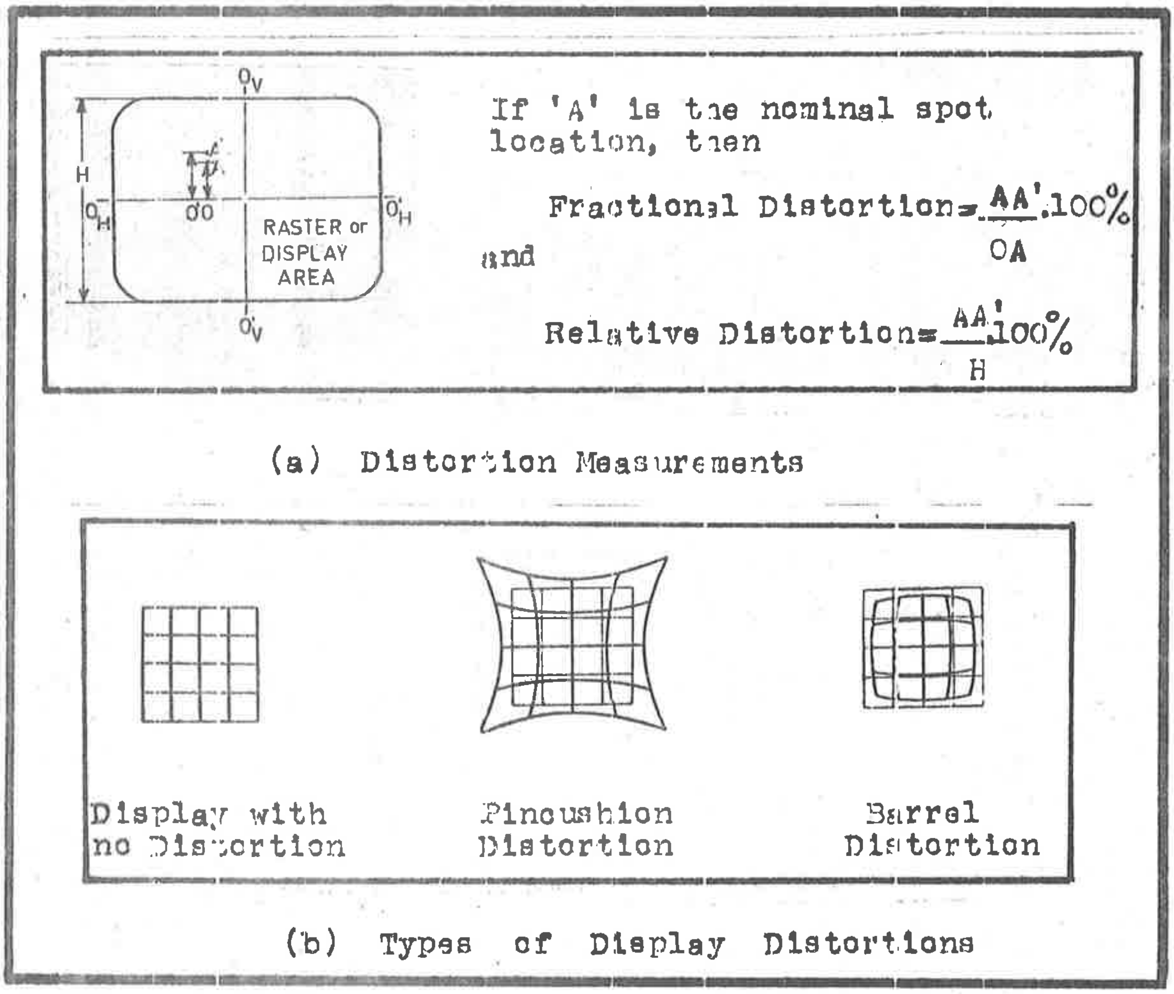


Figure 17. Geometrical Distortion Concepts

Repeatability (see Fig.16) refers to the degree with which an addressed location is consistently located at the same spot on the display area during successive addressing operations; "precision" is also sometimes used to describe this. Ideally a nominated addressed location should be located at the identical physical coordinates on the display area, at least in the short term (ie. the period during which a particular display is observed - say several minutes). In practice, in CRT's which are refreshed continually, power-supply ripples, pickup of stray fields, temperature effects etc, cause successive, nominally "identical" locations, to be shifted by about half to one spot diameter. This is called "jitter". The slow response of the eye, integrates individual spots to "form" a larger defocussed spot, causing an apparent loss of resolution (Fig.16). Repetitive jitter at below 15 Hz (supply line ripples) may cause annoying flicker effects.

Long term "jitter", called drift, due to ageing, usage, (or "abusage"), and even, over a period of several hours, by the change of ambient temperature, causes a particular spot location (and hence the whole display) to move by up to 1% of display area height.

In E-L displays, addressability is specified directly by the number of locations present, being equal to the number of x and y-line intersections; the problem of repeatability does not occur.

2.2.2.7 Geometrical Fidelity of Displays: Distortion (74,76,77)

Addressable points may have excellent repeatability over the whole display area yet there may be different and appreciable spatial separation between adjacent addressed locations in different sections of the display area.

Geometrical Distortion (see Fig.17(a)) is a measure of the spatial displacement of an addressed location from its ideal spatial location. Thus if a display area with dimensions H" x H" contains n x n locations, then centres of adjacent locations should be $\frac{H"}{n}$ apart; if at some locations this is not so, then geometrical distortion, or just "distortion", is said to occur.

Distortion may be expressed in two ways (77):

(1) Fractional distortion, D_F

(2) Relative distortion, D_R (relative to display dimension, H).

Referring to Fig.17(a), if an element whose ideal location is "A" is displayed at "A'", then taking the centre of the display as the origin (i.e. nominating that at the central axis of the display, no distortion occurs), the following are defined:

$$D_F = \frac{AA'}{OA} \text{, the Fractional Distortion}$$

and

$$D_R = \frac{AA'}{H} \text{, the Relative Distortion.}$$

A similar distortion term occurs for the horizontal distortion.

'D_R' is the more meaningful term, as any total shift in the whole of the display area (eg. long term drift) would give unrealistic figures for "D_F" near the reference axis, yet relative displacements within the display may be unchanged.

In CRTs, the distortion is mainly "pin-cushion distortion" (see section 9.4), shown in Fig.17(b) and less often, "barrel distortion". These distortions are due to the inherent screen geometries, being most pronounced when display screens are flat; in addition scanning waveforms may themselves be distorted. In commercial TV-sets, D_R may be of the order of 10 - 15%. In "good quality" CRTs, with corrections etc applied, it still is 1-3%. In certain application, orders of 0.01% linearity have been achieved (78).

Before this degree of geometrical fidelity and linearity can be reached, accurate measurements of distortion must be made. A part of the work reported in this thesis has been towards finding precise means of measuring display distortion, using Moire patterns (Appendix 11).

Display geometrical linearity is not only desirable from a user viewpoint in creating general user confidence in Man-Computer-Graphics, but is very necessary in High Accuracy Graphics, commonly called "Precision Graphics", such as I-C mask layouts, phototypesetting, and computer scanning of bubble-chamber pictures. It is also one of the main requirements for the realization of the particular Interactive Graphics Console, "VIDIOGRAPHIC", proposed and described in this thesis.

2.2.2.3 Readability (9,66,71,72)

This is a quantity used specifically in respect to alphanumeric characters and is a measure of the user's ability to detect and interpret correctly display information, and the speed with which he does it. It is measured in characters per second understood with, say 95% confidence level. It depends on character size, font, spacing, brightness contrast and, of course, on the user ("has he taken a 60 - day fast reading course?").

It has been found that alphanumerics of dimensional ratio 3:2 to 4:3 (height to width ratio), and line separation to letter height ratio of 1:3 to 1:4, are "most" legible.

As mentioned earlier, alphanumerics most commonly are formed by dots or strokes on a 5 x 7 dot matrix (see Fig.20); each letter however is on an 8 x 6 matrix to allow for character and line spacing. On an 8 x 11" display area, with 0.2" x 0.16" character size, about 45 lines of 30-odd characters each, or about 1200 - 1500 characters can be accomodated. In practice this is often less (see Appendix 4) due to spot size, display writing time and other hardware limiting factors. In the largest and acceptably fast precise displays, over 4000 characters can be displayed at the one time.

Cluttering up the display with too much "fine print" confuses rather than helps the user. What should be displayed should be relevent, easily read information, available serially on demand; a display filled completely with text is not necessarily fully relevent at any instant. Displaying the same amount of text in two or three, large, easily readable "pages" is clearly preferrable.

<u>2.2.2.9 Halftones and Colour (9,30,64,73,210)</u>

The use of colour and halftones (but mainly colour) has been often suggested and several experimental set-ups have been implemented, mainly with colour TV-receivers and colour projection tubes. Rather than use full colour, only three main colours, such as yellow-orange, red and blue-green, seem preferred,to minimize the information required (2 information bits are required to specify 3 colours).

Colour is useful in "attention-getting" displays. Thus in one case (73) alphanumeric meter readings are displayed - in green, if the readings fall within the safe limits; yellow, if marginal readings; and red, if danger levels are reached.

Another use is to indicate intensity levels, such as pressure levels in turbulent fluid flow (30).

A third and obvious area of application is where the third or height dimension is required, particularly when this height is discretely quantized. Thus in I-C layouts, several distinct layers or "planes" are present; assigning a colour to each plane, will make such a resultant multi-layer diagram clearer to follow. Present monochromatic I C layouts (236,237) appear to be so cluttered up as to be incomprehensible to any but the designer himself. The analogy between colour coding and the physical case where the depth of a plane in the I-C is indicated by the resultant colour due to selective interference between projected and reflected light is readily apparent and could readily be implemented.

Only one or two manufacturers supply colour tubes for display purposes (57).

2.2.2.10 Hard Copy Capability (8,13,80,81)

During MCG problem-solving sessions which may last up to half or to a full hour, records of required solutions or intermediate solutions are required. This is a self-evident requirement for final solution results, but trial solutions (e.g. tentative automotive designs in the DAC-1 System) may also be required; also partial results may be required.

This hard-copy capability may take two forms:

(i) "Paper hard-copy", for external files or for later casual study by the user. These may be obtained from x-y electromechanical plotters etc, "dry-printing" processes, such as Xerox, directly (through some optical arrangement) from the user-display, or from a separate small precision CRT used for this purpose, run in parallel with the user display device.

(ii) "Working hard-copy", such as obtained with a microfilm recorder or the IBM 7960 unit in DAC-1 system (section 1.2.2) which generates microfilm or 35mm film, which may be used for graphics input at later stages; this form is thus usually used for intermediate or partial solutions. Precision CRTs, separate from the user display, are used. To increase accuracy a small portion only of the CRT display area is used e.g. a 3" x 3" central area on a 5" x 5" screen; fast fine grain films are used. The film frames are indexed and stored in a film store inside the recorder from which they can be accessed on user or computer request. They can also be used to generate paper hard copy.

2.2.2.11. Optional Display Device Features (76,82)

In addition to the above parameters there are several features which are very useful from the user viewpoint but are not strictly necessary.

(a) Blink.

Certain features on a display may require immediate user attention such as, say, notification to the user that a wrong user action was performed, or for example, to indicate that the display device can accept graphics input via the light pen. The most obvious feature requiring blink notification is the cursor (see section 2.4.4.4.). Blinking the "error indicator" or blinking the "tracking cross" for the light pen is an excellent attention-getting device. This is achieved by blanking out the relevent display data say every 3 out of 4 refresh cycles.

Several simultaneous blink rates are possible; the greater the frequency of blink, the more immediate is user action required. Too many "blinking" features can however become very distracting.

(b) Variable Brightness.

For picture clarity, several levels of brightness (say 2 or 4 maximum) are desirable. Required graphic features may thus be emphasized. Increasing or decreasing the electron beam current in CRT displays for the required vectors, either by user or central program control, accomplishes this.

(c) Vector Line Forms

Again for picture clarity and user comfort (for the same reasons as in normal office drawing proceedures), dotted or broken line vectors are desirable options. They are implemented by having a parameter specified within a vector generating command word (section 2.8.2.3), which can be set by the user or the central program, which enables the electron beam on or blanks it during vector generation.

(d) <u>Character Size.</u>

For graphic images containing alphanumerics such as circuit element identifiers, or feature labels or dimensions, which need be scaled up by user request, the characters should also be scaled up appropriately. Two or four character sizes are usually provided in the more sophisticated IGCs. A character size bit(s) in the appropriate command word specifies the required character size.

(e) <u>Optical Superposition</u> (83)

In specific applications certain graphic features need be displayed repeatedly; these may be co-ordinate axes, topographical features in tactical situation displays, or some other complex features. Often too "complex" to generate or requiring, for example, colour for clarity, they may be "generated" by projecting the required features onto the display surface from films, slides etc, often through special rear-window or "ported" CRT tubes.

The tube has two display generating systems: one, the usual electro-magnetically deflected electron beam display, the other, optically projected from a film projector, and beamed through the "port-hole" in the back of the CRT tube.

2.3. DISPLAY GENERATION

2.3.1. Introduction

The fundamental task of a display device is to generate and display alphanumerics and vector-form graphics on CPU or user command. Three tasks need be performed within the display device for this:

(i) Positioning the display data
(ii) Generating the display data
(iii) Maintaining this display once generated.

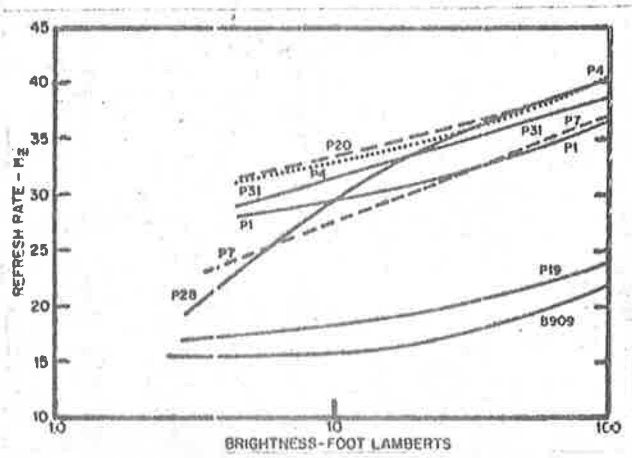
2.3.2 Data Positioning (9,74,76,84,85)

There are three basic methods of positioning data on a display

(i) random positioning
(ii) raster-scan positioning
(iii) "typewriter" or textual scan, a compromise of the above two methods.

2.3.2.1 Random Positioning (Fig.19)

In "Random Positioning" there are no restrictions in the direction in which successive vector elements or characters are positioned. The position coordinates generated within the CPU (or specified by the user via the graphics input device) are D-A converted and the result fed to the deflection coils or plates in CRT displays, or else are decoded to energize the appropriate x- and y- lines in E-L and Plasma panel displays. Finite bandwidth of the deflection amplifiers limits the speed of response and the D-A conversion rate. The "worst" case" limits the random positioning speed; this is the case where two successive points are on the diagonal extremities of the display. In E-L and Plasma panel displays, decoder switching times and excitation times of the elements limit speed of response.



Phosphor	Persistence (seconds). Time for brightness to fall to 10% of initial value
P4	60×10^{-6}
P20	50×10^{-6} to 1.8×10^{-4}
P3	38×10^{-6}
P28	600×10^{-3}
P7	25×10^{-6} to 75×10^{-6}
	400×10^{-3}
P1	24.5×10^{-6}
P19	200×10^{-3}
B909	10

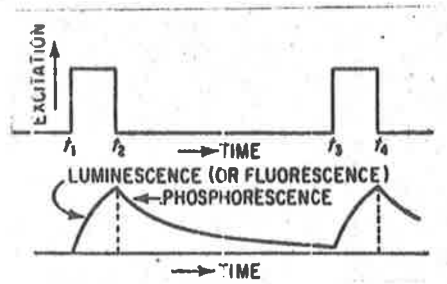


Figure 18. Lowest Refresh Rate giving Flicker-free Display (from [64])

Numbers indicate order in which letters are positioned

Figure 19. Modes of Positioning of Display Data

In CRT's with a 10" x 10" display, delays between spot positioning can range from about 2us to 100us (74). With a flicker-free refresh rate of 30Hz, this gives a flicker free display consisting of a maximum of 300 to 15000 random points. In practice, true random positioning for all display points is very rare; in most display situations, data can be "ordered" or "formated", that is, preorganized. The randomness of successive points to be displayed is reduced with a consequent reduction in bandwidth requirement in the deflection or positioning circuitry. The number of points to be displayed is thus increased if display data formatting is used in raster-scanned displays (see below).

Electrostatically-deflected CRT's usually are considered to have higher bandwidth than magnetically-deflected CRT's; however with recently introduced low inductance deflection coils this "advantage" has become marginal.

2.3.2.2. Raster-scan Positioning (Fig.19)

In "Raster-scan Display Positioning" extensive display data preorganization or formating is carried out. The formating is essentially a space-time transformation of a two-dimensional image into a line-sequential format.

In CRT displays, the electron beam is deflected in a predetermined linear fashion across and down the screen; the beam intensity is modulated with the display information and is intensified at the point where graphic information is to be located. The predetermined beam positioning means lower deflection amplifier bandwidth (and hence reduced hardware costs); this is at the cost of display flexibility. User-specified points (via say a light-pen) are detected with an average delay of half the time required to scan the complete display; the worst-case occurs when a line is being manually input in the direction opposite to scanning direction. The direction of scanning is usually as shown in Fig.14, the same as for TV-displays (and hence TV monitors or receivers are often used for these displays); spiral-scans (the origin of scanning being the centre of the display) or other scanning patterns can also be used, but are used mainly for radar displays. A normal TV-type display, with 4:3 aspect ratio (width to height ratio) using about 570 of its lines for display (625-line TV-system) with the same resolution along the horizontal line (about 750 points), gives about 4.3×10^5 resolvable points; with a 25Hz refresh rate the beam modulation amplifiers and circuits need a bandwidth of about 15Mhz.

In E-L and Neon-Plasma panel displays, raster scanning is obtained by sequentially activating, say, the vertical driver lines first; during the interval when one such line is on, the horizontal lines are sequentially activated.

Raster scanning is most often used in economic multi-console configurations (86,87,88); simple sequential memories for display refresh may be used (section 2.5.3). At the cost of reduced resolution (cf. above lines resolution), slower user input data-rates, and appreciable graphic data reordering in the CPU, many low-cost graphic displays can be serviced by one central computer (section 2.6.2.2) Alternatively hardware implementation of formating may be used; this is either via scan-conversion tubes (93), where the display is initially random-scan displayed as usual, and then is converted to a raster-scan within such a tube. Imaging a random scan display on a Vidicon TV camera performs the same scan conversion function also.

2.3.2.3. Typewriter or Textual-scan positioning (Fig.19)

This method of data positioning is a compromise of the above two methods and is used in situations where long strings of alphanumerics or text need be "randomly" positioned. The analogy to "typewriter" format is apparent. In CRT's, the beam may be randomly positioned to indicate the beginning of a new line of text or a string of alphanumerics horizontaly to the right of the preceeding character; by specifying certain typewriter-like commands such as "space", "carriage return", "end message" etc, "patterns" of text may be positioned. Between such commands, the characters are generated by usual means, such as random-dot generation, or raster-scan generators.

2.3.3. DISPLAY GENERATION (9,13,24,25,64,74,76,84,85,89)

2.3.3.1 General

The generation of graphic or alphanumeric data, once its position has been selected, can be accomplished by four basic modes:

(1) Point-plotting mode
(2) Stroke-writing mode
(3) Ramp drawing mode
(4) "Stencil" or "shaped-beam" writing mode.

2.3.3.2 Point-plotting Mode (Fig. 20(a)

Point-plotting mode is the most "primitive" method of display generation and is identical to random positioning of points, (where each point is the smallest display element (in CRT's of electron beam diameter size)); the sum of these points make up the complex graphic shape

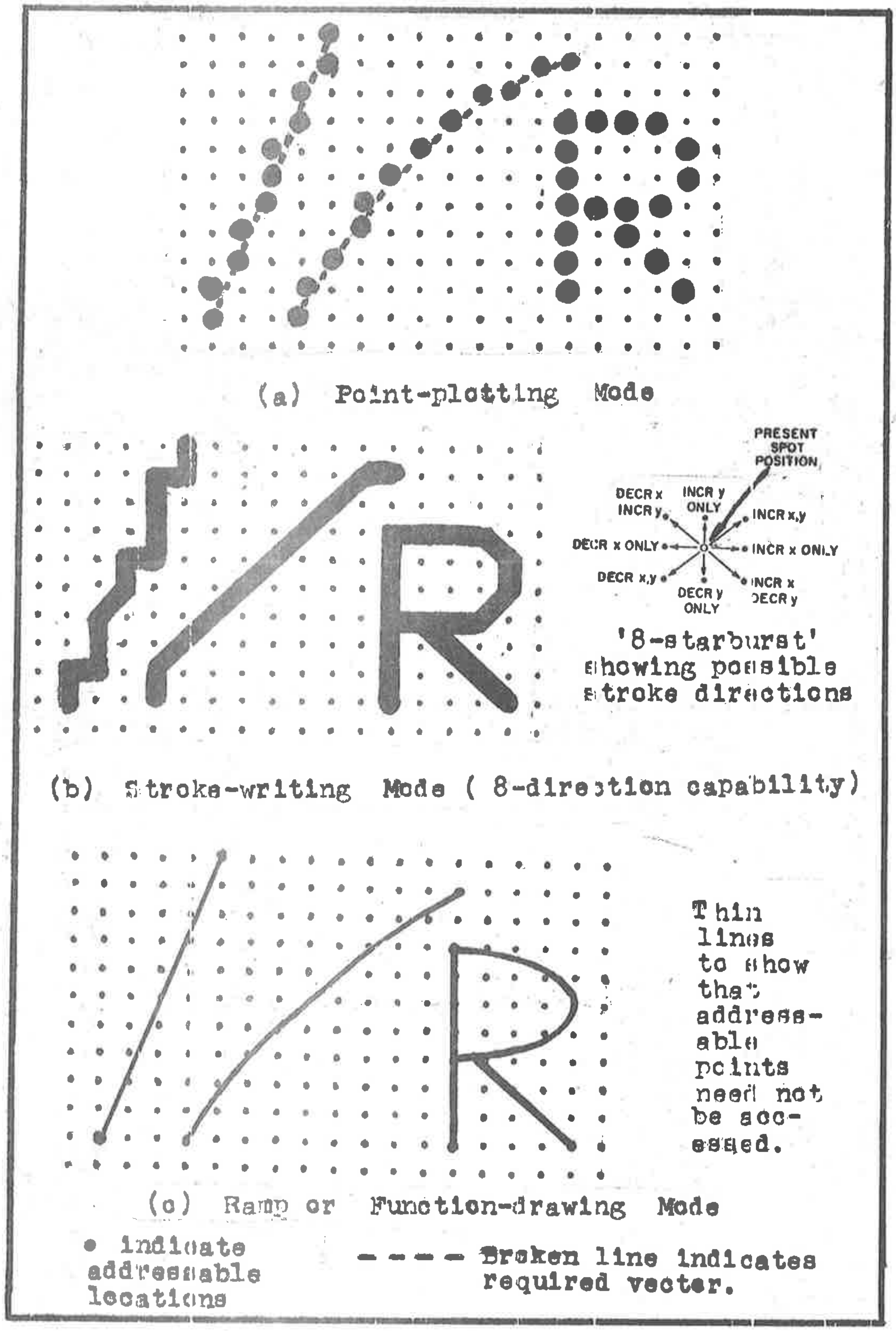


Figure 20. Vector and Alphanumeric Generation Modes

to be drawn. The points are directly positioned point by point under direct program control.

For a 1024 x 1024 addressable display, a line diagram or alphanumeric character is expressed as a string of two 10-bit coordinates (see section 2.8.3.3). Alternatively, rather than use "absolute" coordinates, relative coordinates may be used. Strings of x and y position increments, starting from some given location, specify a character or line drawing. Coordinates are D-A converted and thence fed to the appropriate deflecting circuit inputs.

Reduction of raster addressability to say a 512 x 512 matrix leads to a "beady" appearance of lines and characters.
E-L and Neon-Plasma panels essentially use this mode of display generation.

2.3.3.3. Stroke-writing Mode (Fig. 20(b)).

From any specified coordinate location, the beam can be continuously swept, in 4 (up, down, left, right), or 8 (as before and 4 diagonal positions) or even more, directions to the next addressable location. The possible beam directions form a "star-burst" pattern about their centre of origin (see Fig. 20(b)). Thus instead of the "beady" appearance of point plotting, continuous line vectors or characters are generated; however, if the addressability is still coarse, lines still have step or "quantized" appearances. Lines are continuous but not necessarily straight; if the x and y deflections are not exactly equal in sweep rate, the resultant stroke has a "bow" in it, although the end points are correctly specified. Again both alphanumerics and lines may be drawn with this method.

2.3.3.4 Ramp Drawing Mode (Fig. 20(c)).

For straight lines, all that is strictly necessary to specify a line are the end points or an initial point and the slope and direction; the length may be computed. Specifying end-points and performing simple logic and arithmetic operations, control signals are generated for a programmable ramp generator whose output are the appropriate analog signals for the deflection coils or plates.

In one case (incremental vector generation) (24,25) Incremental Integrators (I.I.'s) controlling the deflection circuits are incremented by the number of steps specified by the "Δx" and "Δy" increments from the end point of the previous vector. The "Δx" and "Δy" increments are signed to indicate direction.

For slope and length vector specification, a scheme similar to the following is used. Say 12 bits specify some (short) vector: 4 bits specify slope (22.5° increments) and 8 bits specify the length. An "up-down" counter controls the deflection circuitry from the end point of the last vector. The slope bits control the D-A converter scaling and whether the slope is positive or negative. The 8 length-bits control the number of counts. Alternatively the 4 slope-bits control pulse widths which are fed into the I.I.'s for a period proportional to line length (see Fig. 29).

An interesting method, purely analog, using the orthogonal summation of sinusoids on the display screen (Lissajous Figures) has been implemented in the ARTRIX System (90). If into the x-deflection plates of the display CRT, a waveform of the general form

$$x = R_1 \sin(\omega_1 t + \varphi_1)$$

is fed in while into the y-deflection plates a waveform of the general form

$$y = R_2 \sin(\omega_2 + \varphi_2)$$

is fed in, then both lines and circles can be generated.

For $\omega_1 = \omega_2$,

(a) if $\varphi_1 = \varphi_2 = 0$ then,

$y = \frac{R_2}{R_1} x$, a straight line.

<u>Note</u>. This scheme can be extended to all forms of conics:

(1) An ellipse: - if $\varphi_1 = 0$, $\varphi_2 = \frac{\pi}{2}$, $R_1 \neq R_2$, then

$$\frac{x^2}{R_1^2} + \frac{y^2}{R_2^2} = 1$$

(2) a parabola: - if $\varphi_1 = \varphi_2 = \frac{\pi}{2}$, $\omega_2 = 2\omega$, then

$$y = \frac{2R}{R_1^2} \cdot \frac{(x^2 - R_1^2)}{2}$$

(3) a hyperbola:- if $\varphi_1 = \varphi_2 = \frac{\pi}{2}$, $\omega_2 = \sqrt{2}\,\omega$, then

$$\frac{2x^2}{R_1^2} - \frac{y^2}{R_2^2} = 1.$$

(b) if $\varphi_1 = 0$, $\varphi_2 = \frac{3\pi}{2}$ then,

$y = -\frac{R_2}{R_1} x$, a straight line of negative slope.

(c) if $\varphi_1 = 0$, $\varphi_2 = \frac{\pi}{2}$, $R_1 = R_2$ then,

$x^2 + y^2 = R^2$, a circle of radius R.

With very high "ω", vectors can be generated in a few microseconds (see Fig.30).

Circles and other conic sections or more complex curves may also be generated with I.I.'s as above, with gated feedback links between the separate I.I.'s controlling x and y deflections (25). Complex curves may be generated by sections of exponential curves (91,92); in this case, the resultant vector increments can be shown to be specified, on the average, by less than 2 bits per increment.

Processing time is much shorter than before as no longer does a list of points or strokes, constituting the line or curve, need be produced; storage for the display file is also reduced. The tradeoffs between software and hardware are clearly evident; display generation speed advantages occur (e.g. a full length vector (across a display diagonal) can be generated in under 20µs), at some expense of accuracy and increasing hardware costs. Another advantage is lessened programming and CPU requirements.

2.3.3.5 "Stencil" or "Shaped-Beam" Writing Mode (Fig22)

For alphanumeric characters, or commonly used symbols such as in military applications (planes, ships, triangles etc), templates or stencils containing those symbols can be inserted within the CRT near the electron-gun section; the electron beam passing through the stencil is "extruded" or shaped by the particular symbol at the location of the beam, and after further deflection for positioning, is displayed as that symbol. When specifying a symbol or character to be displayed, deflection signals are generated for a local set of deflecting plates or coils which position the beam to pass through the appropriate stencilled character; the shaped beam is re-centred with the tube axis and then normally deflected for positioning as before.

The character font is fixed by the template; new tubes are required for different fonts. The "Charactron " tube (Fig. 22) is the "classic" tube of this type (93). Up to 20,000 characters per second may be generated.

Arbitrary line drawings could be drawn using such a tube (having a "point" or even "strokes" on the stencil) using dot or stroke drawing methods. However, the method is restricted usually for symbol displays.

2.3.4. <u>Alphanumeric Generation (9,85,93,94,95,96,97)</u>

2.3.4.1 <u>General</u>

Alphanumeric generators convert codes, which specify the desired alphanumeric, into control signals causing the specified character to be displayed. They are usually hardware implemented.

Many techniques have been implemented, probably for patent reasons (alphanumeric display consoles are much more common than those with vector generation capabilities), and because the problem is basically much simpler than line generation of complex curves, as

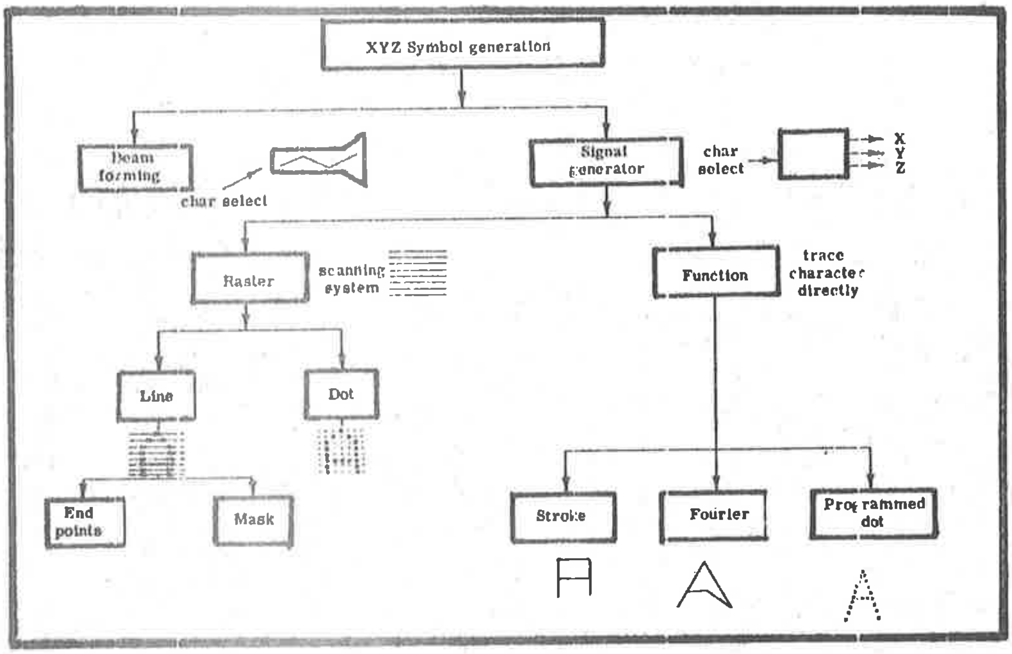


Figure 21. Classification of Alphanumeric Generation

(from [94])

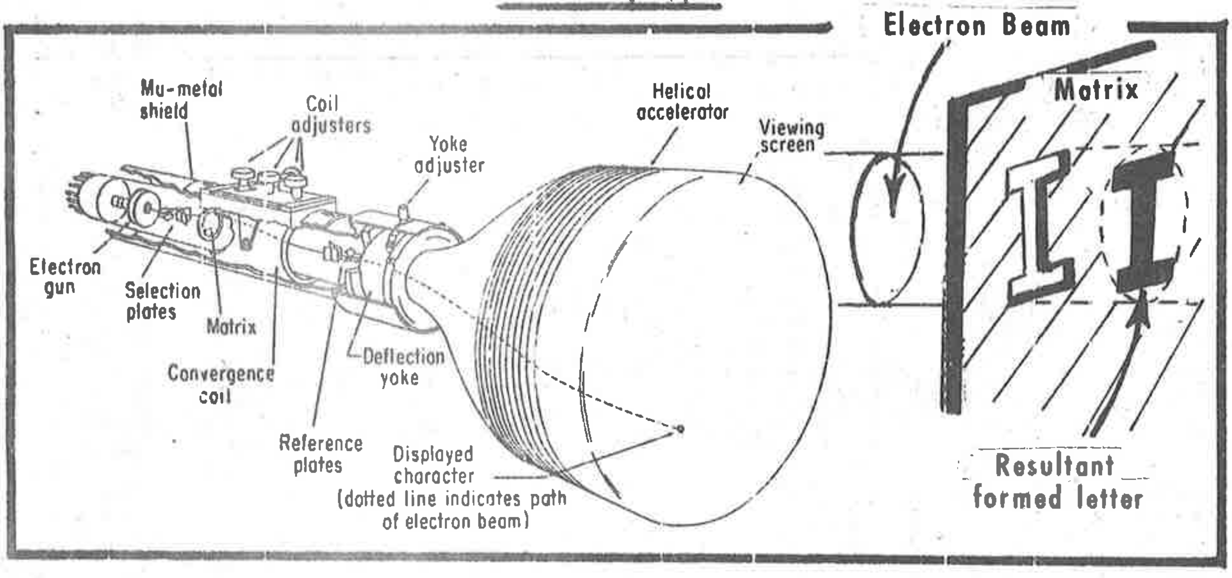
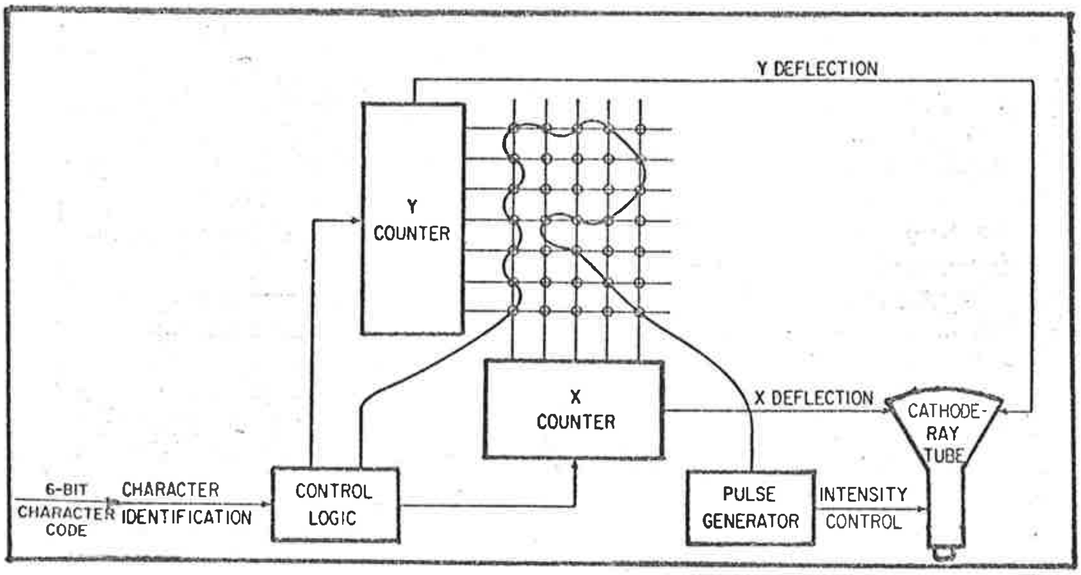
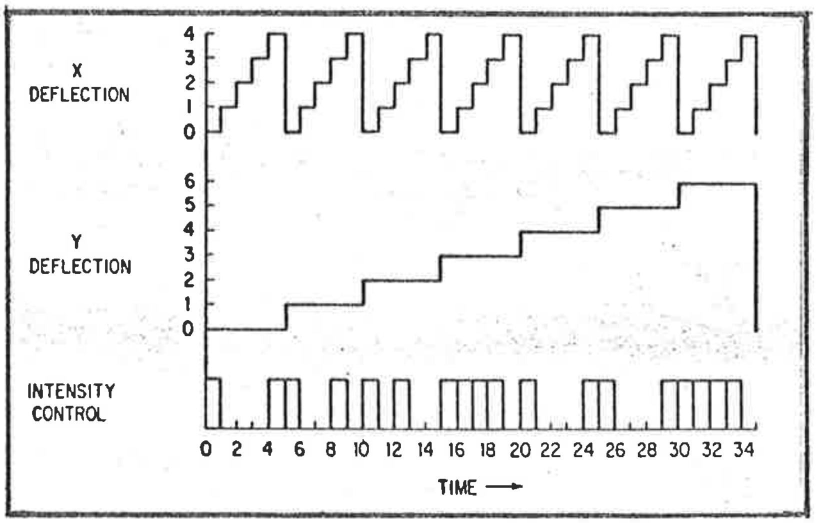


Figure 22. Beam-formed Generation ('Charactron')

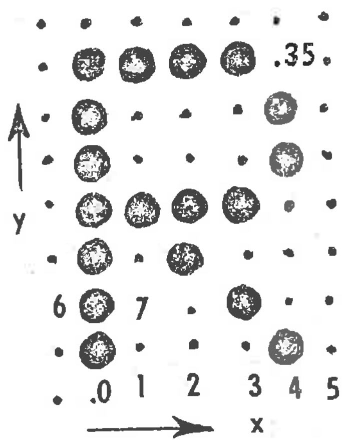
(from [93])



(a) Block diagram of Generator



(b) Deflection and Intensity waveforms



(c) Resultant Character

In this and the following figures the time units on the axis correspond to the numbered sequence indicating the sequence of character generation.

Figure 23. Fixed Format Dot Generation (using core store)

(based on [35])

character shape is fixed and the only variable is positioning. Distortion in character shape can be more tolerated than in line graphics; some of these distorted characters are even touted as "different" fonts.

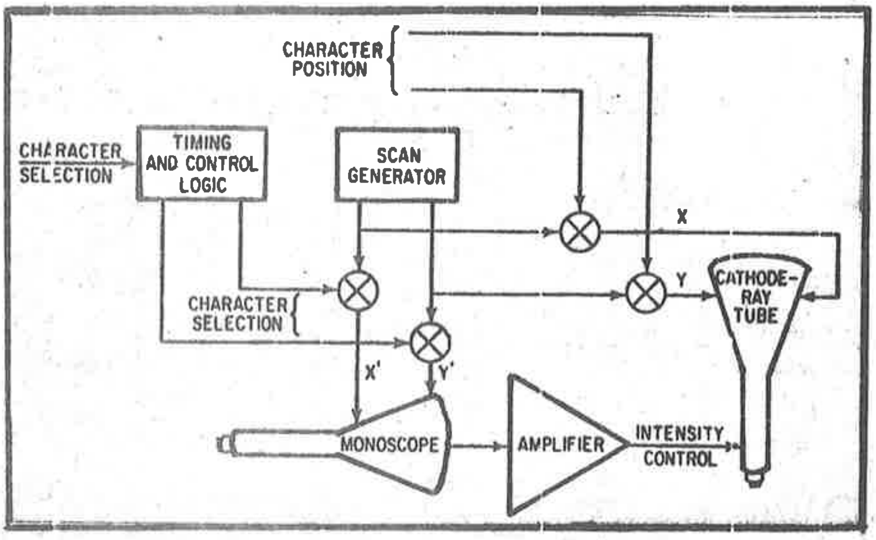
The principal methods of generation are (Fig.21)

(1) Beam shaping (described in section 2.3.3.5)
(2) Various raster-scanning methods
(3) "Function" generation

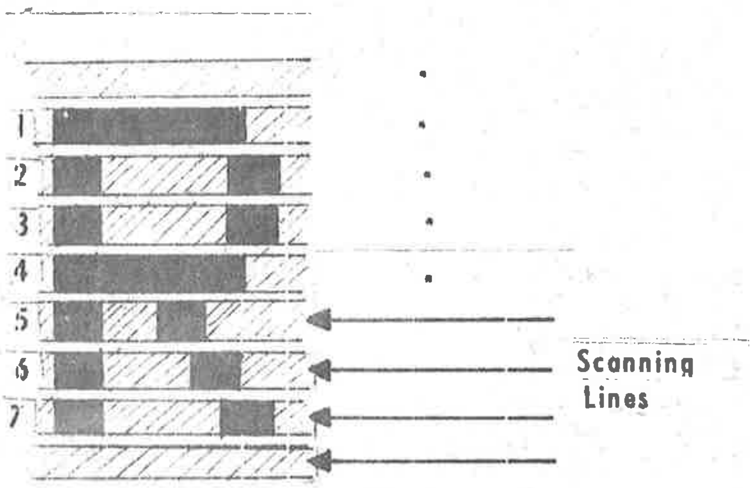
2.3.4.2 Raster Scanning Character Generations (Fig.23)

In a CRT, the beam is grossly positioned at the location where the character is to be generated, and a small, local raster-scan is made corresponding to the size of the character. Usually characters are formed on a 5 x 7 position matrix; in some schemes, a 5 x 7 array of memory locations (e.g. in a magnetic core store) is associated with each alphanumeric (of which there may be typically 64, 96, or 128 different characters and symbols); in a core memory, a read wire is strung through cores corresponding to the character shape. The character is generated by strobing each core sequentially; a pulse is generated in a strung core which then modulates the beam intensity. Dots or horizontal strokes may be generated. Other forms of Read-Only-Memories (ROM's)are also used, including a Flying-spot scanner with a character mask in front of it; the beam intensity signal is the output from the detecting photo-tube. Character generation times of about 1us can be achieved with the fastest ROMs.

The "monoscope" is another read-only store character generation using a CRT, the "monoscope". The character "mask" is drawn on the target screen with



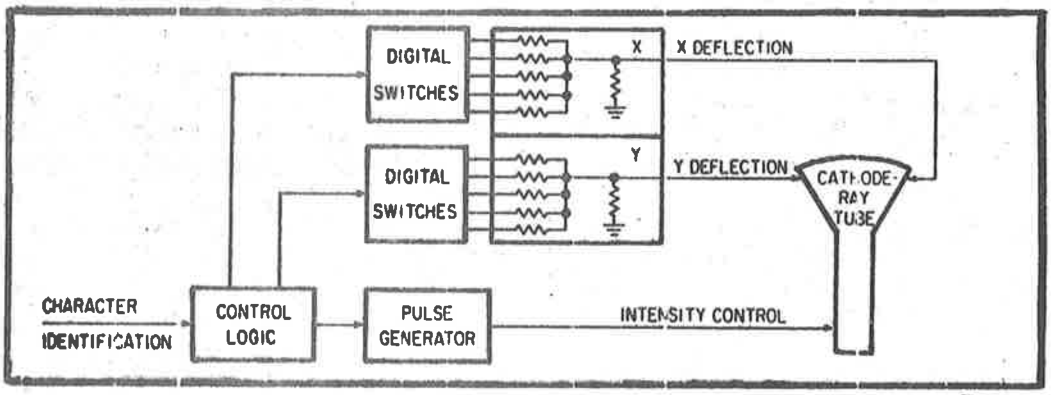
(a) Block diagram of Generator.



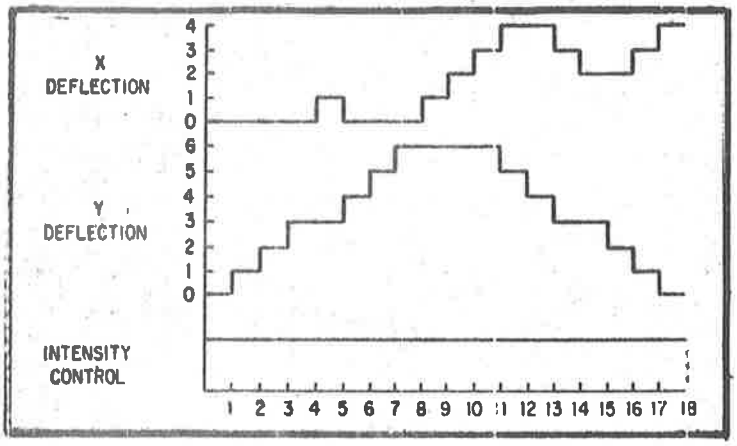
(b) Resultant Character.

Figure 24. Fixed Format Scan Generation (using Monoscope)

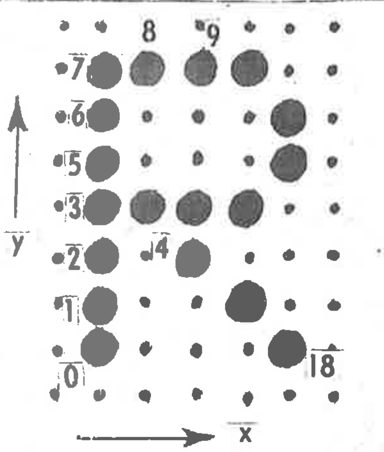
(based on [95])



(a) Block diagram of Generator.



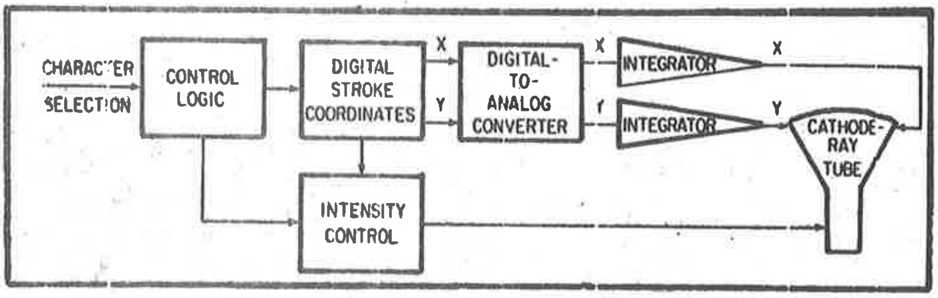
(b) Deflection and Intensity Waveforms.



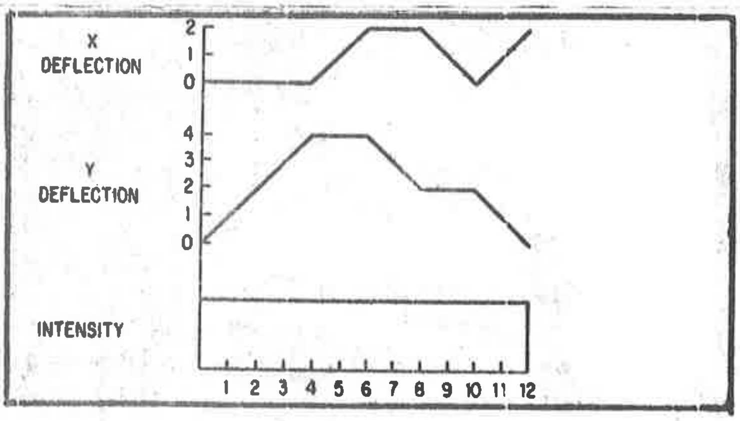
(c) Resultant Character

Figure 25. Programmed Dot Generation (resistor storage)

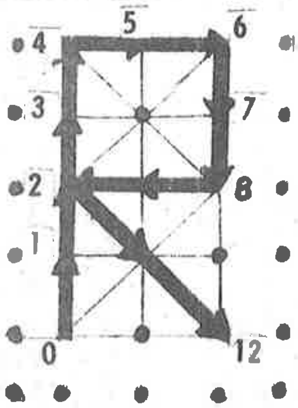
(based on [95])



(a) Block diagram of Generater.

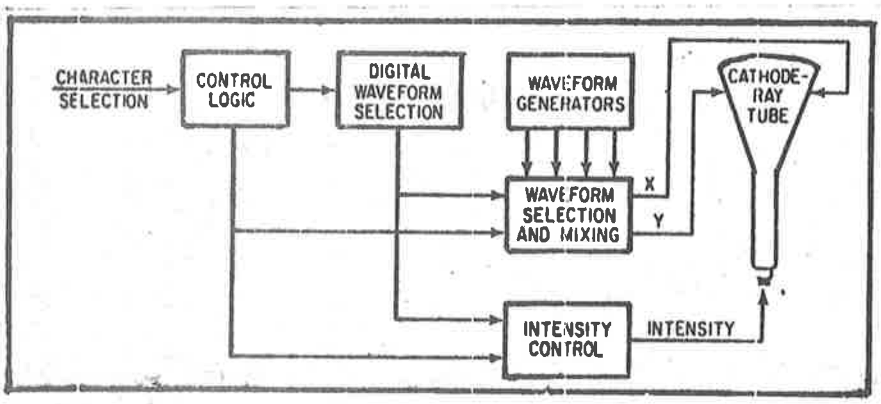


(b) Deflection and Intensity Waveforms.

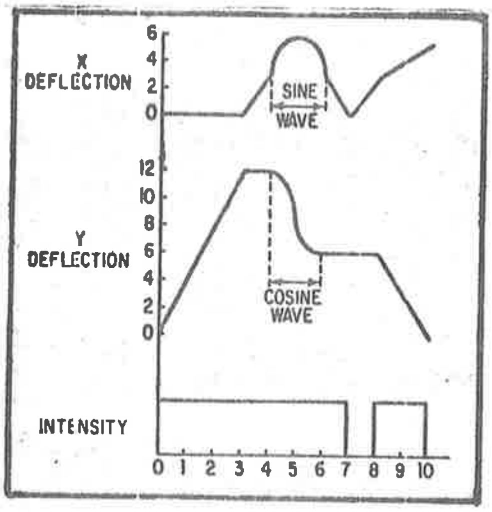


(c) Resultant Character.

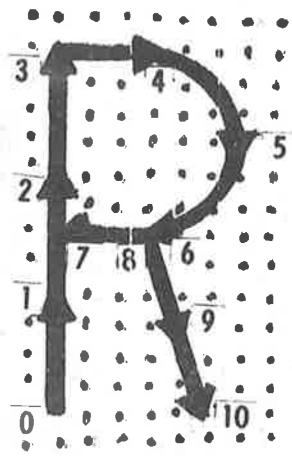
Figure 26. Stroke-writing Generation (based on [95])



(a) Block Diagram of Generator.



(b) Deflection and Intensity Waveforms.



(c) Resultant Character.

Figure 27. Lissajous Figure Generation (based on [95])

secondary emissive material. When the appropriate character is scanned, the secondary emitted electrons are collected by an anode, amplified and used to intensity-modulate the electron beam of the CRT display on which the character is to be displayed (see Fig.24).

A disadvantage of Raster Scanning Generation is that lower brightness than in other methods is available, as a large proportion of the character generation time does not have the intensified beam on (see Fig. 20(a),(b)).

2.3.4.3. Function-generated Alphanumeric Generation

The deflection signals causing the character to be displayed may be by the "Programmed" dot or "Point-plotting" mode (Fig.25), the "Stroke - writing" mode (Fig.26), or by hardware generators using segments of Lissajous Figures or Fourier Series Sequences (Fig.27). The appropriate sequences of coordinates or control signals are ROM-stored. Character drawing times of a minimum of 3µs are possible. In the Fourier Method, combinations of harmonics of various magnitude and phase are combined orthogonally to generate waveforms which then are fed to the x and y deflection circuits as shown in Fig.27.

2.3.4.4. Comparisons (8,74,94,95)

From a user viewpoint, dot-matrix characters (whether raster scanned or point-plotted) are considered to have poorest quality, while shaped-beam characters have best quality, particularly if complex characters are required. The disadvantage of the latter is inflexibility (new characters, new tube!). Another factor to be considered is that character generators may be shared between several consoles; a beam-shaping tube however is good only for its own character generation.

Time factors may be important; programmed dot characters require about 100µs for generation; beam

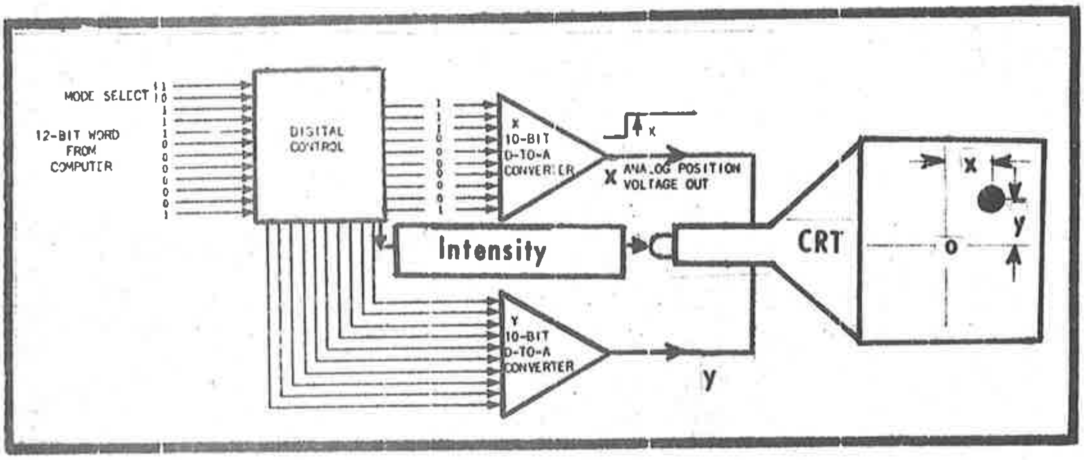
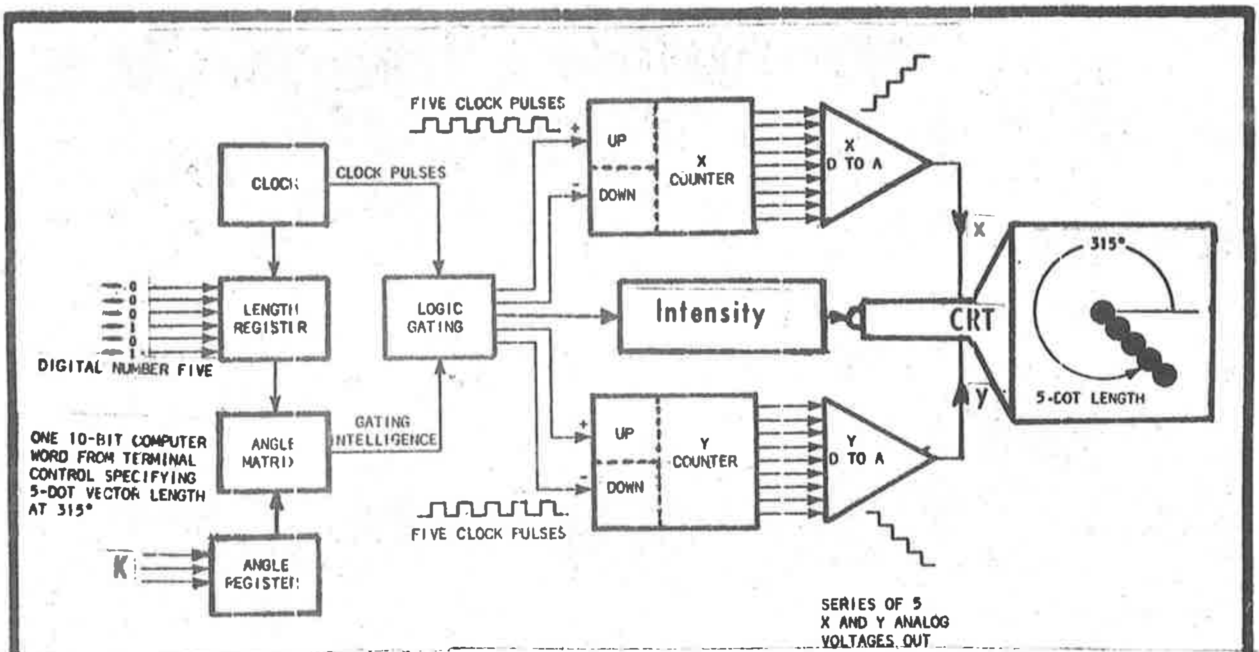


Figure 28. Simple Point Plot Vector Generator

Two 12-bit words are required for each point plotted.



Three 'angle' bits K specify one of the 8-starburst directions.

Figure 29. Stroke Writing Vector Generator

(based on [84])

shaping requires only the character selection-deflection time, which using electro-static deflection, may be only 1-2μs. To the above character generation times must be added positioning times. The number of characters displayed in random positions at adequate refresh rates is in the range of 150 to 5000, while with "Typewriter formating" it is about 250 to 8000 (74).

2.3.5 Vector or Line Generation (9,25,64,74,84,85,96,99,100)

2.3.5.1 General

A vector is a positioned line having a specified origin, a length, and orientation. Vector generators convert or translate line definitions in alphanumeric or binary form into controls signals, causing the defined line to be displayed.

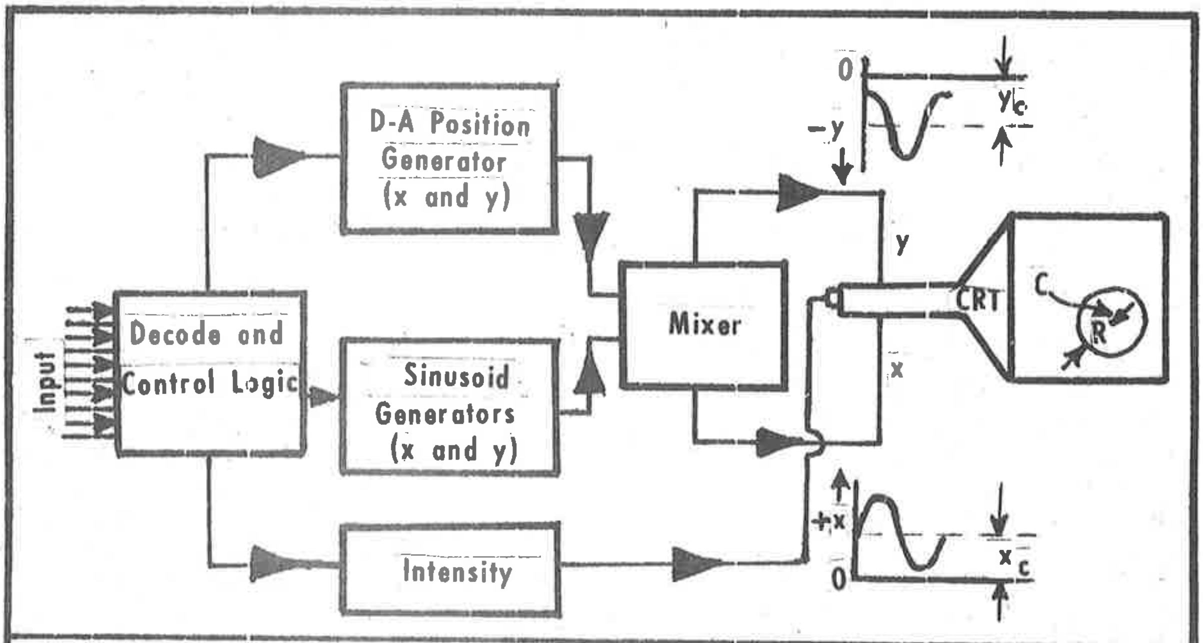
The methods available are :

(1) The Point-plotting mode (sec. 2.3.3.2 and Fig.28).
(2) Stroke Writing mode
(3) Random Drawing mode
(4) Raster-scan mode
(5) Storage Properties of Displays.

2.3.5.2 Stroke Vector Generation (Fig.29).

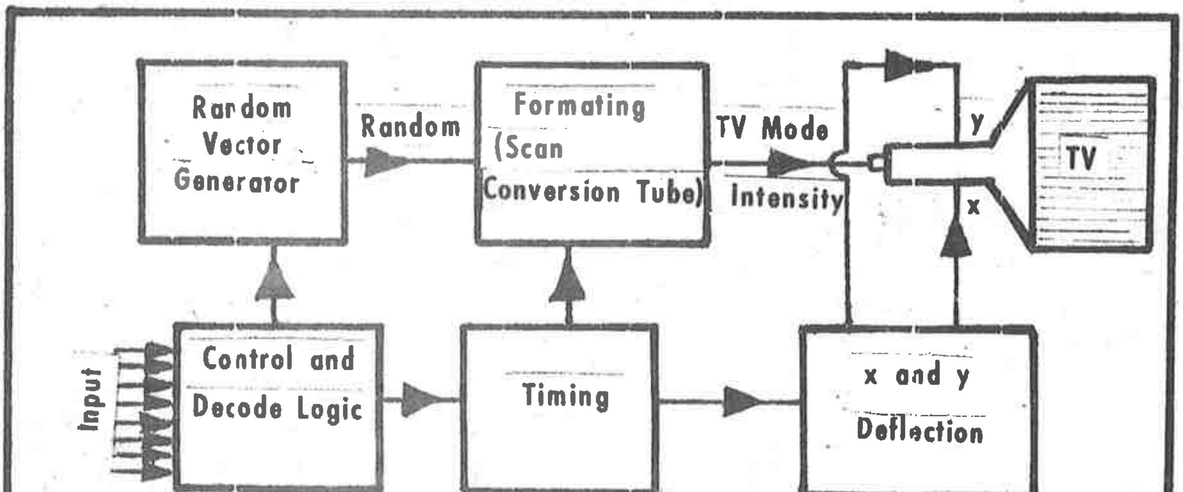
In stroke vector generation (see section 2.3.3.3), the stroke lengths are of limited length and of limited angular orientation, being approximately of half an alphanumeric character in length (of the order of 0.1"). Such strokes are concatenated to produce vectors, which on closer observation appear "stepped",or at least not smooth, as in Fig.20(b).

Vector specification may be by specifying successive addressable locations, with the continuous strokes being swept out and generated between them (this requiring excessive storage), or better still, by specifying the initial point or beginning of the vector, and successively specifying one of the 8 possible (or more if available) "starburst" directions.



Input vector commands specify position 'C' of centre and 'R' the radius of the circle.

Figure 30. Lissajous Figure Vector Generation



Formatting for scan-conversion is hardware implemented.

Figure 31. Raster Scan Vector Generation

2.3.5.3. Ramp-mode Vector Generation

This is the method essentially as described in section 2.3.3.4. However display brightness variation due to line drawing speed may result. This depends on which of the forms of writing mode are used:

(a) Constant time, variable length vector generation (using simpler hardware than its alternative), implies a constant amount of beam energy (beam current x time of beam sweep) being transferred to different amounts of phosphor (depending on vector length) and hence producing different brightness, the shortest lines being brightest, and the longest lines being dimmest. Compensating circuits to alter beam current are required. However, the fastest writing speed (for longest vectors) are achieved with this method.

(b) Variable time, constant speed vector generation generates, as implied, constant vector lengths in constant time giving uniform brightness, and hence resolution without additional compensating circuits. Long vectors take proportionally a longer time to generate.

2.3.5.4. Raster-scan Vector Generation (refer to section 2.3.2.2 and Fig.31)

When the raster-scanning circuit timing is precisely timed (usually from the beginning of the top horizontal line and the left-hand edge of the raster area), then the time during each scan can be associated with two coordinates specifying some location; the vertical location "y" is given by the time from the start of the vertical scan, while the horizontal location "x" is given by the time from the start of the horizontal scan.

Re-ordering graphic data by sorting in ascending vertical coordinate magnitude, and within a given vertical co-ordinate, rearranging, by sorting in ascending horizontal coordinates, rearranges display data into raster scan form. At "matches" between the horizontal and vertical sweep timing and the corresponding ordered coordinates, the electron beam is intensified, thus displaying a point at the required coordinates. If the speed of data re-ordering is fast enough, a complete new display can be generated within a scan time, i.e. of the order of 20 - 30ms. Any loss of timing synchronism, leads to curved "straight" lines and other image defects of a geometrical nature. Section 2.3.2.2 gives alternate methods of data formatting for raster scanning.

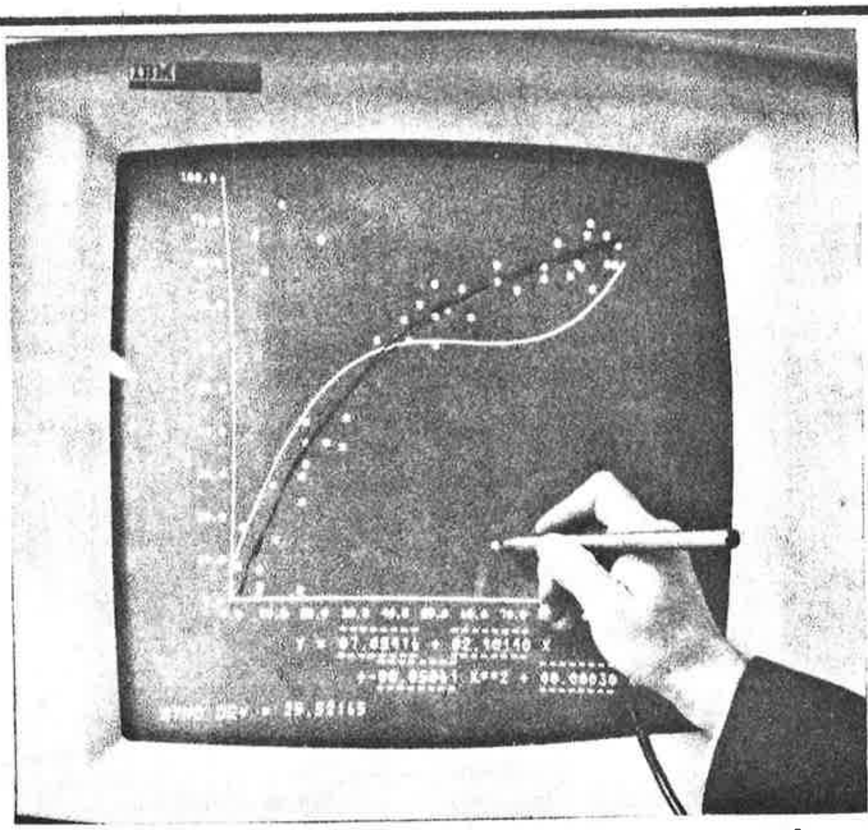
2.3.5.5. Storage Properties of Displays (96)

In certain applications where moving targets are displayed (e.g. traffic control, radar etc), CRT screen phosphors may be selected with long persistance; thus when a moving target is tracked, a continuous trace can remain visible for several seconds or more due to the persistance of the phosphor. This method is not used in computer-driven displays.

2.3.6 Other Graphics Operations

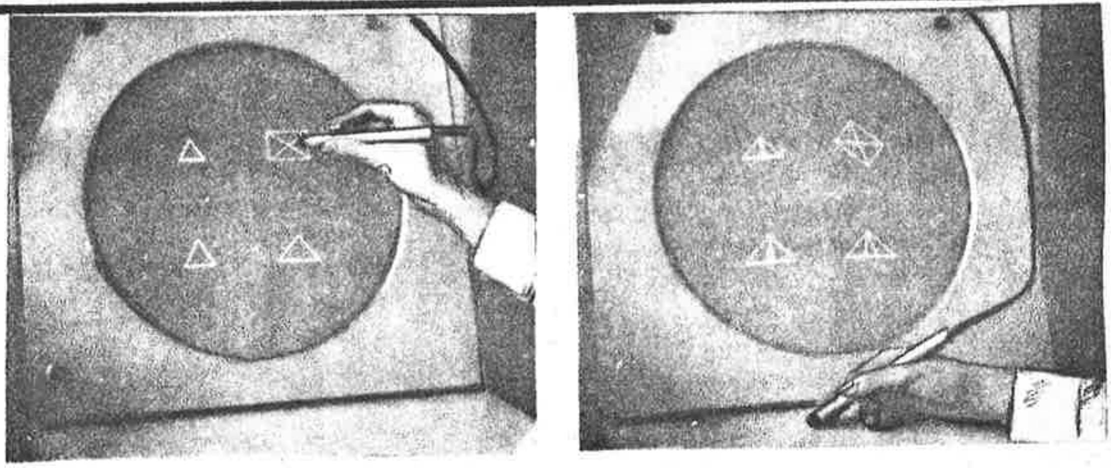
2.3.6.1 General

It has been the experience that in computer graphics and particularly in Interactive Graphics such as in DAC-1 and SKETCHPAD, simple alphanumeric and vector generation is inadequate; certain functions to be performed on user-request recur quite frequently. Other than selecting a point or feature, and deleting a line or



With the Light Pen,points and objects may be position-ed or,if already displayed,may be selected for further operations (as above).

Figure 32. Graphic Operations - 'Pointing' (from [62])



The four views of a pyramid are rotated clockwise by 30°

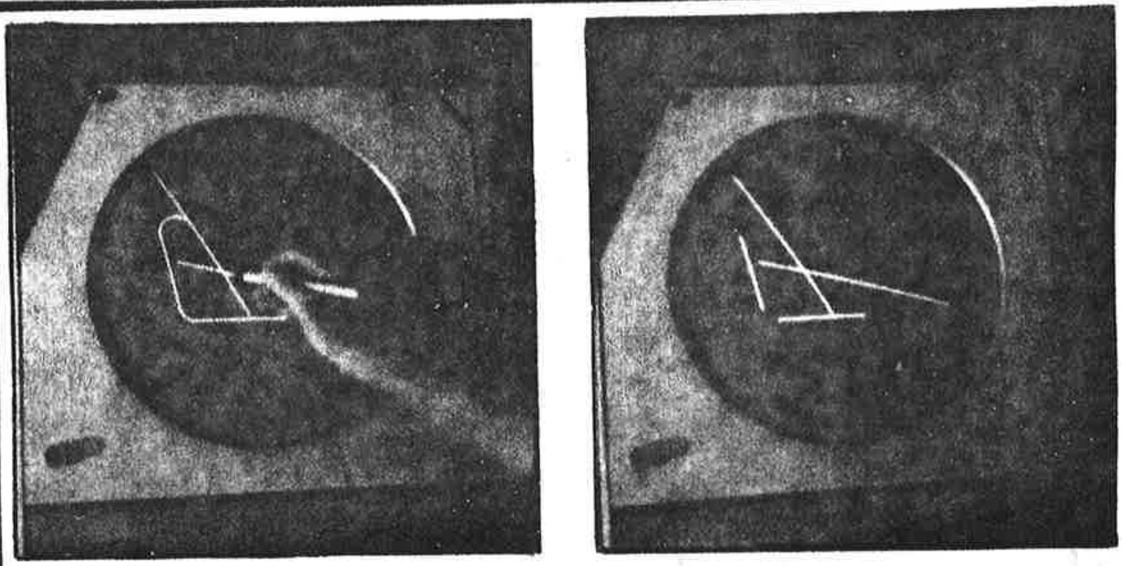
Figure 33. Graphic Operations - 'Rotation' (from [4])

feature, there are included also requests for rotating a given view by a certain specified angle, scaling some section of a display by a given magnification, translating some graphic element or section of a display to a new position and "replicating", whereby any given number of complex identical elements may be generated by specifying or pointing to one such element already displayed.

All of these functions can be performed by subroutines and can thus be considered as CPU operations rather than IGC functions. However the frequency of occurence of these functions in MCG, and, as in SKETCHPAD, for reasons of speed, efficiency and cost, make hardware implementation of these functions attractive; locked at in this aspect, these functions belong to IGC display tasks. A recently released commercial unit, a descendent of the graphics console on which SKETCHPAD was implemented, incorporates the above functions in special purpose hardware units (63).

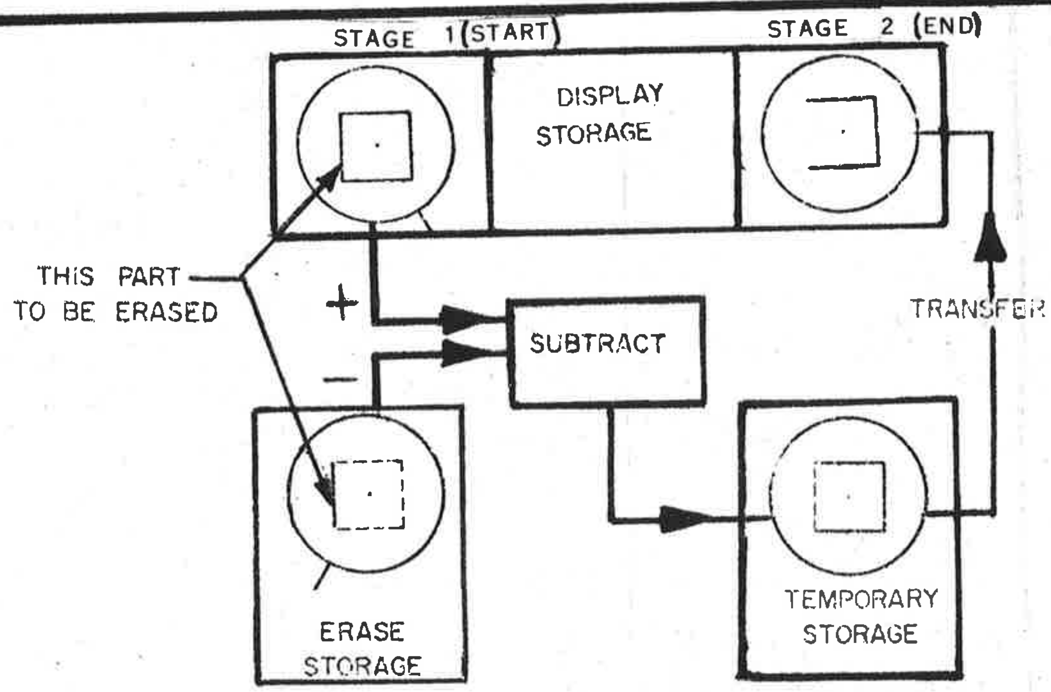
2.3.6.2. Pointing or "Picking" (101) (Fig.32)

Pointing by a user may be to specify a point (or set of points to indicate end-points of vectors, or boundaries) or to an item of a "menu" (which is a set of options presented to the user and from which he may select one to indicate his course of action - see section 2.8.4.1). This may be accomplished by a "light-pen" (Appendix 5) which is a light sensitive device which picks up the phosphor luminescance of the CRT within the field of view of its "pen tip", or else via a graphic input pad or tablet, points on which have a 1:1 positional relationship with the display screen and which may be accessed via its own pen or stylus.



The three rounded corners are 'pointed' to and the 'erase' function button depressed.

Figure 34. Graphic Operations - 'Erasure' or 'Deletion'



Each 'storage' is a 'Memotron' Storage Tube

Figure 35. Hardware Implemented Erasure (from [90])

The pen in both cases generates a signal which identifies the location of the pen. This location may be associated either with a function (an "item" of the "menu") or else, with user activated keyboard functions, may initiate a function at that point - e.g. "erase" the point (which clears the display file at that location) or it may specify the initial point of a user-specified vector. Pointing to a sequence of display locations in turn and activating certain function keys, may generate vectors between these specified points, circles etc.

2.3.6.3 Display Deletion or Erasure (Fig.34).

User-display interaction not only consists of drawing subelements such as points, lines or other sub-elements but may also necessitate modifying already drawn details by erasure or deletion. Erasure may be achieved as mentioned above: specifying the end points of a vector, which needs erasing say, and depressing the "erase" function button which results in clearing that vector from the display file. Similarly a graphics tablet or pad may be used in specifying the end points.

An interesting solution to erasure was implemented in the ARTRIX System, which is a hybrid analog-digital system (90). Three storage tubes of the Memotron type (section 2.4.5.2) are used. One tube is the "Display Storage" which corresponds to the refresh display memory holding the display file. The section selected for erasure is carefully traced out, vector by vector, with a light pen; this generates a video signal (the display is raster scanned) corresponding to the erased section and stored in a second Memotron Tube, the "Erase Storage". This erase information is inverted and analog-added to the display storage when both are read out via raster scanning; the result is stored in the "temporary storage", again a Memotron tube. The inverted information cancels out that in the display tube giving the net image with the erased elements omitted. The net result is re-transferred to the display storage tube. All 3 storage tubes need have accurate 1:1 position relationships and run off the same timing or synchronising system. This is illustrated in Fig. 35.

2.3.64 Rotation (25,26,63)(Fig. 33)

The new coordinates of an object which is rotated (but not translated) by some arbitrary angle are given by the matrix multiplication of the old coordinates with a 3 x 3 "rotation matrix", the elements of which are dependent on the 3 rotation angles (a rotation around each of the 3 coordinate axis) making up the rotation angle. Rotation of a graphic figure (or a "3-dimensional" projected figure) can be achieved by specifying the 3 rotation angles (digitally) via the keyboard; special purpose hardware does the rest. Alternatively a hand-activated rotatable "Crystal Ball" (Appendix 5) which controls 3 mutually orthogonal shaft-encoders or the like, generates 3 analog quantities which specify the necessary rotation matrix coefficiants (28).

Essentially, matrix multiplication is by Binary Rate Multipliers (BRM's); the resulting 3 components in each coordinate direction are summed in the appropriate x and y accumulating registers whose outputs control the deflection circuitry.

As in SKETCHPAD, complex conic functions (ellipses, parabolas, hyperbolas) may be generated by using a circle, parabola, or hyperola generator in conjunction with a rotation matrix.

2.3.6.5 Scaling (25,26,63,102) (Fig. 36)

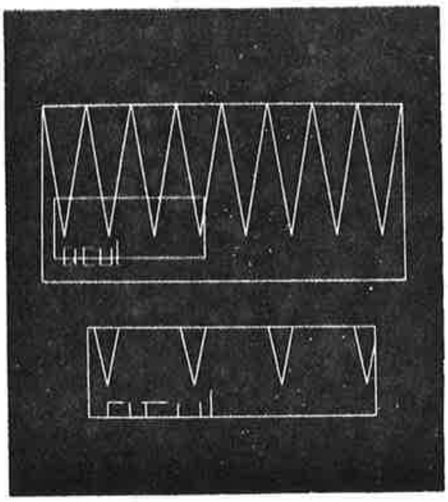
In engineering graphics or perspective graphics certain segements of the displayed image may require "closer examination". Such a segment requires boundary coordinate specification, usually via a light-pen or other display-pointing device, and the degree of magnification; the specification of the segment to be scaled is called "windowing". Again this is a form of coordinate transformation on certain specified display coordinates and hence the scaling can be expressed by a matrix. Scaling factors are usually "x2", "x4", etc and in one case (63) are "x4096".

Hardware implementation has advantages particularly in speed; in one case (details of which are not given) lines can be scaled in 2-9µs as compared with about 1ms per line using SKETCHPAD-type software (63).

2.3.6.6. Translation (26,102) (Fig.36)

In graphics applications such as circuit specification, I-C layout, or applications where several elements of the same shape are required, elements often require relocation to some newly specified position (e.g. via a light-pen); occasionally simple rotation by $\pm 90^{\circ}$ or 180° may be required.

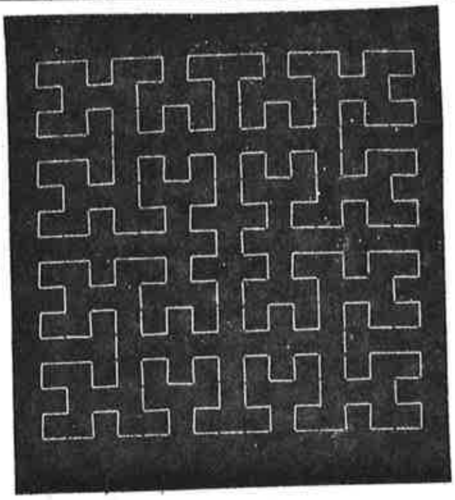
In the simple case of translation, by say "Δx" and "Δy" units, these coordinate increments are added to all the element coordinates.



The original image is shown in the upper half; the section to be scaled up is delineated by forming a 'window'. The scale of magnification is specified, the location of the scaled-up window is indicated, and the result is displayed in the bottom half of the display area.

Figure 36. Graphic Operations - 'Windowing' and 'Scaling'

(actual photo from [63])



Complex patterns can be constructed by a short sequence of very simple commands. The original 'master' element appears to be "  "

By rotating (through 90°, 180°, 270°) this element and indicating the desired locations and short segments, the resulting pattern is displayed.

Figure 37. Graphic Operations - 'Replication'

(actual photo from [63])

2.3.6.7 Replication (25,26,63)(Fig.37)

Replication is the ability to generate and display any number of figures at given specified locations by pointing to the "master" figure required (which has already been constructed) and pointing to the desired position locations. This may be considered to be analogous to alphanumeric character generation, where the "characters" are the arbitrary geometrical figures to be replicated. In I-C layouts, where symbols such as squares and rectangles need be positioned at various locations, this is a very useful feature. More complex figures can be drawn from much simpler figures.

2.3.6.8 Comments

The above operation namely, Rotation, Translation and Scaling, being coordinate translations can be accomplished by a common 4 x 4 matrix (3 rotations, 1 translation). A hardware realization of this has been implemented in the LDS-1 (Line Drawing System-1), a commercial graphics console, with a "matrix multiplier" unit (63).

The other two operations, Pointing and Deletion, are fundamental to user interaction in MCG; both depend on the ability to detect points or locations on the display area. The rest can be done by software. Light pens or graphic pads or tablets and pens have been developed for this.

2.3.7 Form and Types Of Available Displays (70,103,104,105,106)

2.3.7.1 General

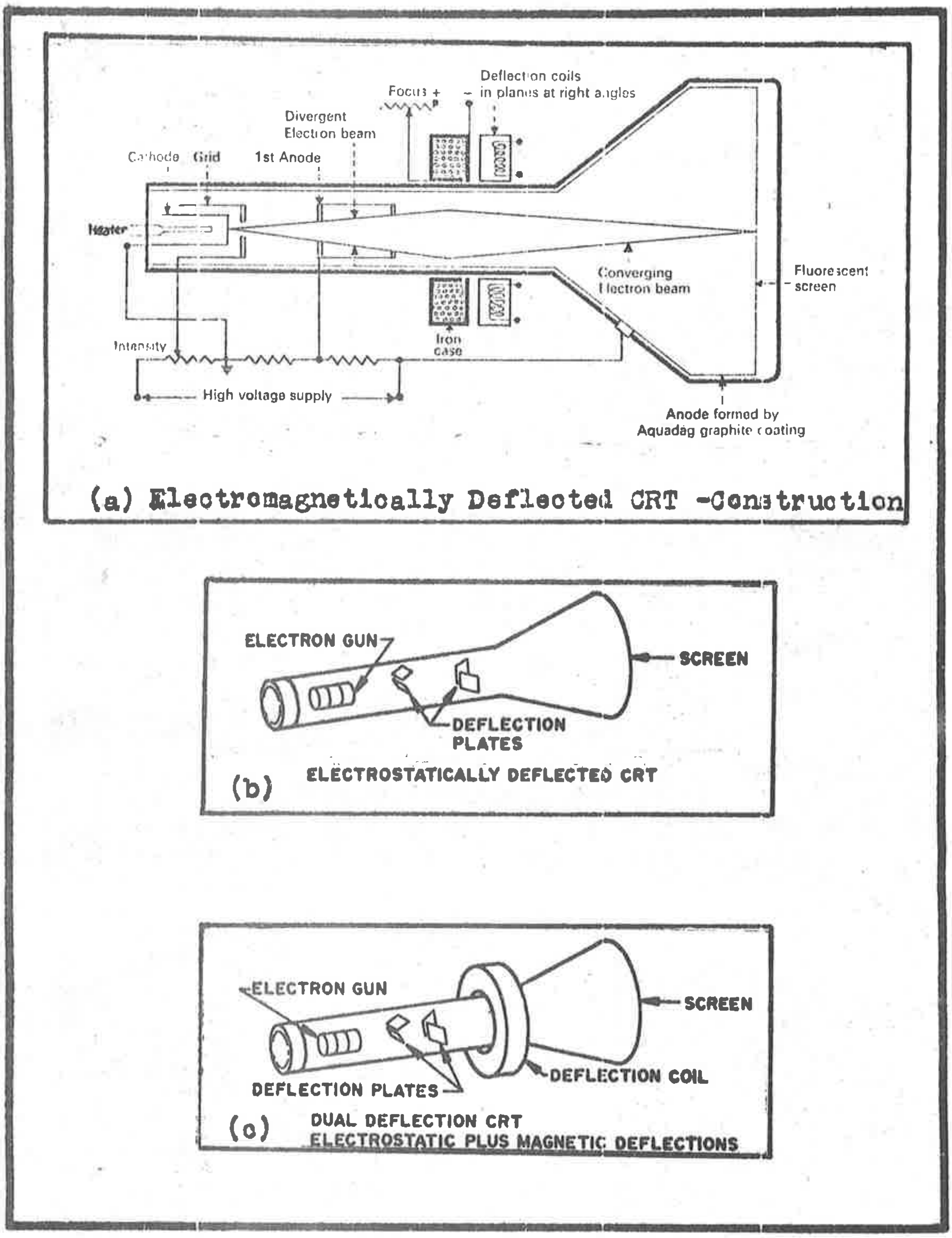
A display provides a visible representation of alphanumeric or graphic data, which may be either computer or user-generated, for the benefit of, and evaluation by,

the user. "Visible representation" implies the presence of light energy in specified locations determined by display information while "computer-generation" and "user utilization" implies fast-positioning and fast-generation of light energy. These requirements greatly limit the choice of electro-luminescent effects which can be used for displays. For these reasons the form that displays take is very limited, being almost exclusively CRT's displays using the effect of "cathode-luminescence". For completeness's sake, other experimental displays or effects suitable for displays are mentioned also. Direct-View-Storage-Tube displays are mentioned in section 2.5.4. Also, "single function" or "individual" display, such as illuminated control panels, indicating lights, alphanumeric readouts etc are excluded.
Appendix 3 compares displays using different luminescent effects.

2.3.7.2. Cathodoluminescent Displays - CRT's (95,103)

Cathodoluminescence (CL) is the process whereby electron beam energy ("cathode rays"), incident on a phosphor screen, are converted into visible light energy. Cathodoluminescence is the basis for CRT's. Phosphors are materials which emit visible light under the influence of electric fields or incident kinetic energy. Various phosphors have different rise-times and persistence of luminescence - these can range from nano-seconds to several seconds (see Fig.18). The resultant brightness is more than adequate for viewer comfort.

Electric accelerating fields are used to create the incident kinetic electron energy; in addition transverse electric or magnetic fields cause beam deflection and hence positioning of display information. The use of electro-magnetic fields for display generations and posit-



(a) Electromagnetically Deflected CRT -Construction

(b)

(c)

Figure 38. Cathodoluminescent Displays - Cathode Ray Tubes

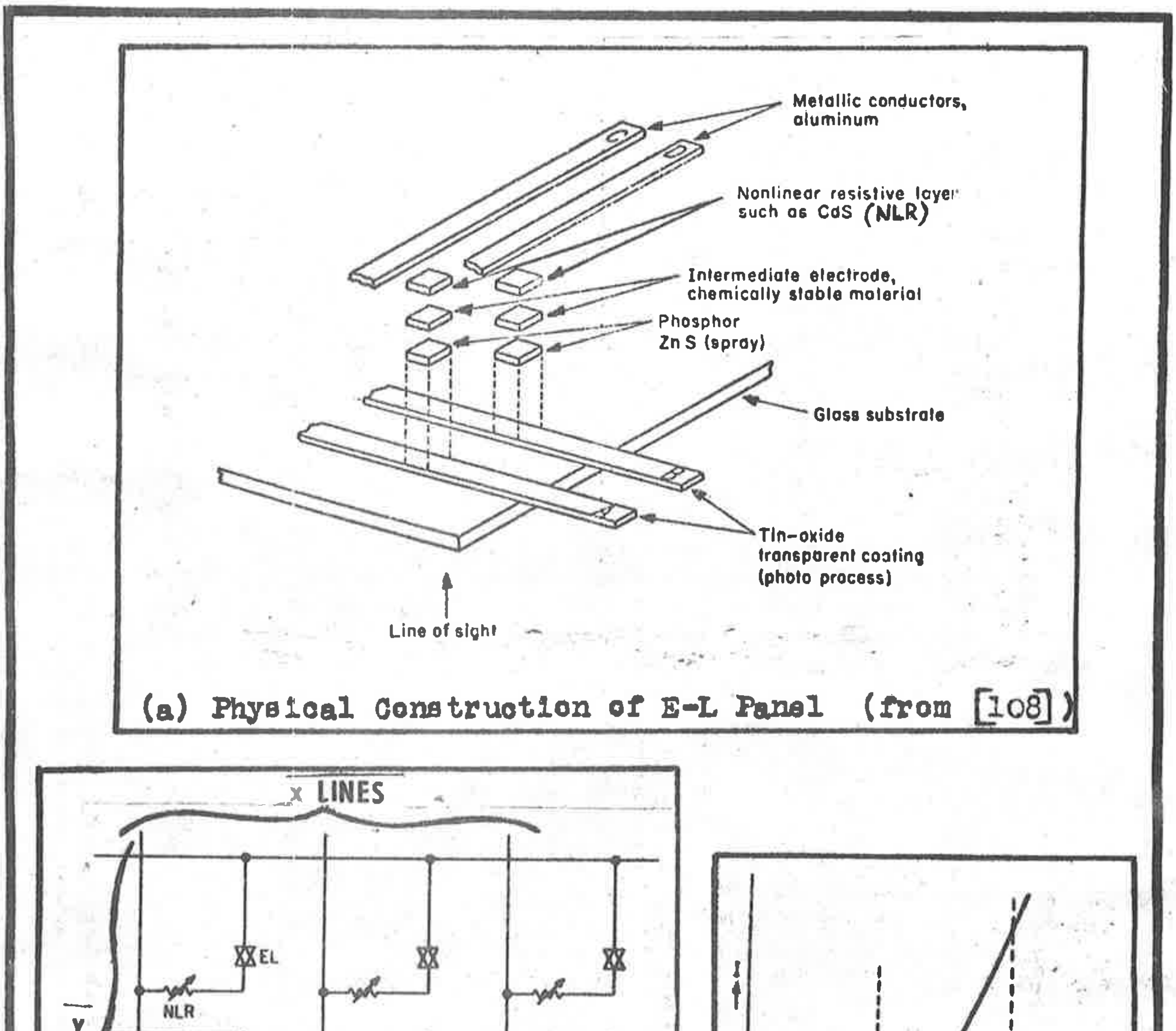
ioning,with the speed advantages of these methods, has ensured the preeminance of CRT's in the display field. The construction and the two main types of CRT's are shown in Fig.38.

The operating details of a typical CRT will be described in more detail in later chapters (4 and 9). Appendix 4 lists several commercial CRT displays, the listing being based on display versatility.

2.3.7.3. Electroluminescent Displays - E-L Panel Displays (68,70,107,108,109,110)

Electroluminescence is the process whereby electrical field energy within a phosphor layer is converted into visible light energy. Since electric fields within a region are caused by two conducting planes at different voltages, then, for a phosphor within such a region to have its emitted light visible to the user, one of the conducting planes needs be transparent; tin oxide is one such transparent conductor used. To obtain discrete resolvable locations on which to generate the display, the electric-field-specifying planes are split into sets of equi-width, equi-spaced parrallel line conductors, one set in an x-direction say, and the other set orthogonal, in the y-direction. The intersection of these orthogonal sets of lines specify the display locations; by selectively energizing "x" and "y" lines, an adequate electric field is generated within the intersection location, thus making that location visible (Fig. 39(a)).

Experimental displays (68,109) have been made, but have been found to be inadequate for computer-driven or interactive displays for the following reasons:



(a) Physical Construction of E-L Panel (from [108])

(b) Section of E-L Panel

(c) Nonlinear Resistor Voltage-current Characteristic.

Figure 39. Electroluminescent Display - E-L Panels

(i) the displays have low brightness and low contrast.

(ii) the relationship between electric field strength and resultant luminous energy generation, although extremely non-linear, does not have the desired switching characteristic i.e. there is no thresh-hold electric field intensity above which the light energy generated becomes visible, and below which none is generated. The result is that all elements in a selected row or column may be visible: contrast suffers as a result. However, non-linear resistors have been used to make this threshhold sharper resulting in "luminance discrimination" ratios of about 10^4 (109). (See Fig. 39(c)).

For these same reasons, more complex display information requires that more complex switching, selection circuitry and memory are required to ensure that unwarranted locations are not switched in (110).

(iii) slow response time (typically milliseconds).

(iv) complex memory is required. Effectively this means a storage element need be associated with each location. This frees the CPU from regenerating the display; it only generates it once. However when even a single new point need be generated, the whole display is erased and the new updated one is generated.

(v) Problems also exist in switching high AC drive voltages.

2.3.7.4. Gas Electroluminescence - Ionization Plasma Display Panels (107,111,112,113)

When certain gases (e.g. neon) have electric fields impressed across them, ionization of the gas occurs with resultant generation of visible light, the brightness being proportional to the ion current flow.

The major advantage is that the voltage level V_0 at which ionization occurs is higher than V_m, the voltage at which ionization is maintained; the extinction voltage, V_E, is much lower still. Thus if some bias voltage, V_S, such that $V_E < V_S < V_0$ is present, then locations or gas "cells", once "ignited", remain on.

A display unit consists of a glass sheet with an array of holes (locations) drilled, within which a gas such as neon is trapped by it being sandwiched between 2 other thin glass sheets, and on which thin glass sheets, orthogonal x and y sets of transparent conductive lines as in the E-L panel displays,are deposited. Their intersections are located at the gass-cell locations. About 40 cells per linear inch have been realized, giving comparable resolution with available CRT displays (Fig.40). Plasma panel displays have not progressed far beyond the experimental stage as, thus far, the following problems still occur:

(i) slow switching ("ignition") and extinction times (approximately 20µs)

(ii) unwanted triggering of adjacent cells occur by ignited cells.

(iii) lack of uniformity and stability in cell composition and structure (e.g. uneven gas pressure) leading to non-constant igniting voltages; depending on how long a cell has previously been ignited, the ignition voltage may vary by up to 30% of its nominal value.

With these problems solved, plasma displays would hold great promise, particularly as cells can be ignited with "light-emitting" pens (as distinct from the usual light pens which are light receiving), leading to direct user-inputting of data onto a display (see Appendix 5).

2.3.7.5 Injection Luminescence - Light - emitting Diode Arrays (LED Arrays) (70,114,115)

When certain semiconductors (e.g. Gallium Arsenide, Gallium Phosphide) have electrons injected into their P-region, then, during the resultant recombination of holes with these electrons, visible light is emitted.

LED's can be made as small as 0.02" x 0.02", which gives about 50LED's to the linear inch. Other advantages are that switching voltages are Integrated-Circuit Voltage compatable, high brightness, and fast switching times (about 1-5ns). At around a price of $1.00(US) apiece, a 1024 X 1024 array would cost around $$10^6$(!) for the LEDs alone; a cost projection of about $0.10 per element by 1980 (70) still make the cost prohibitive. Its main attractiveness is reliability, flat-panel display construction, low-voltage drive circuitry, in addition to the advantages already mentioned. A proposed LED display panel is shown in Fig.41, with a 4-layer diode providing the memory element for each display location.

2.3.76 Other Forms of Displays

(1) Arrays of small bulbs have been occassionally suggested but size limitations (15 bulbs to the linear inch at the best), low switching speeds (milliseconds) and relatively high cost $0.50(US) each have limited their use.

(2) Magneto-optical effects (106), liquid crystals (103), photochromic effects (106), oil-film systems ("Light Valve projectors") (106), Laser Displays (106), 3-D displays (116), have been suggested but they all suffer from one or more of the following disadvantages: slow speeds, lack of flexibility, size limitations and cost.

2.4. THE USER - INPUT INTERFACE

2.4.1 General (9,60,119)

The function of a graphics console is not only to display final or intermediate results but also to accept data in alphanumeric or graphic form from the user, either at the beginning of the problem or else during the interactive phase of the problem run.

Many devices, some very ingenious, have been proposed, not only having been developed for certain applications (e.g. for pointing ability or for drawing non-linear vectors etc) but as mentioned before, to circumvent patent or other proprietary rights.

2.4.2 Requirements of User-Input-Interface

A user input-interface should have the following capabilities:

(i) alphanumeric input capability to specify commands or data; electromechanical keyboards are normally used for this, although "light button" keyboards have been proposed and implemented (4,120).

(ii) 2- or 3- Dimensional graphic input capability either by "free-hand" specification or else by tracing from hard-copy "masters". Such graphic information may be plots of one variable against another, transfer function curves, map contours and the like. This form of graphics information, which otherwise would be expressed numerically, need be accurately drawn in. Various forms of "graphic input pads" or "tablets" are used for this. Light-pens are inefficient for this purpose as extensive software support is required.

(iii) 2- or 3- Dimensional topological input capability, such as for specifying circuit topology, block-diagram interconnections etc. Positional accuracy is not required, merely topological accuracy.

(iv) Pointing ability to some specified point or feature on the display, to either specify some command (e.g. "light button" commands), or to perform some function on the specified graphic feature in conjunction with a teletype keyboard or function keyboard (e.g. pointing to the boundaries of a "window" for scaling). Light pens, "cursors" and the like are used for this.

It should be emphasized here again that "pointing ability" is the basis of Man-Computer-Graphics. Through it, the user can, in theory, be linked to any part of the resident program during its solution run and be able to alter that part and hence any subsequent sections of the program by (subsequently) pointing to "light-button" commands, or by the use of keyboards or function buttons; that is the essence of Man-Computer Interaction. A display device to make computer memory contents and other associated files available to the user for accessing, is necessary for this.

User graphic-input capabality, whether for specifying precision graphics or topological information, is very useful from a user viewpoint and "elegance" of input, but is not strictly necessary; the same data can be easily input in the form of numerical tabular data on tape or cards, although this may perhaps be cumbersome. A graphics input device is thus not strictly necessary for "graphic" data input!

In addition, when assessing or selecting graphic input devices, the following points should be kept in mind:

(1) Naturalness or ease of use - a pencil or pen-like device with a "fine-tip" seems most natural for user comfort.

(2) Compatability with the display - particularly in size of input area to the size of display area.

(3) Ease of interfacing with the CPU and display device. Input devices based on exotic effects and requiring "fancy" transducers should be excluded. Speed of response should be fast enough so as not to tie up excessively the Input-Output Control System.

(4) Software support to interpret the input signals must be simple; CPU and memory requirements should not be excessive.

(5) Geometric accuracy in positioning or generating the graphic input is highly desirable.

(6) Economy in capital outlay and in use is desirable.

2.4.3. Input Interface Parameters

2.4.3.1 General

It would be easy to say that performance factors of a graphic input device, such as resolution, geometrical distortion (or lack of it!), speed of response etc, should be of the same order as those for display devices. In fact, however, these requirements can be relaxed.

Display device parameters should ideally be matched with the corresponding parameters of the observer's eye. Graphic input devices on the other hand are limited by the capabilites of the user's hand. For example, the eye can discern whether a line is straight or not, or can resolve 2 parallel lines of say 0.005" separation; but it is a rare person who can draw a truly "straight" line, and a rarer one still, who can draw it straight in any specified direction, at some specified location. Even tracing from a hard-copy master rarely gives a good match between original and copy.

Manual inputs of curves, lines, etc are rarely used in precision graphics; for such precision work electro-mechanical, precision x-y digitizers, or micro-film readers such as the IBM 7960 unit of DAC-1 (11) are used.

Pointing, however, can be performed very accurately by hand; for this reason "hand-drawn" lines or "vectors" are specified by pointing to the two end-points, with computer-generated vectors drawn in-between. Circles are "drawn" by indicating the desired centre location and numerically specifying the radius. In generating non-linear curves, small inaccuracies can be tolerated; hand drawing is thus acceptable. Topological information, such as specifying circuit or flow diagrams, requires neither high resolution nor geometric accuracy, implying that manual drawing is adequate.

2.4.3.2. Resolution

Resolution is a measure of the fineness of detail capable of being accepted by the input device, and hence of the complexity of drawings which can be manually input. Ideally, this resolution should be of the same order as the display device to which it is interfaced; thus the input area of the device should be resolvable into say 1024 x 1024 addressable locations.

Commercial devices recently available achieve this quite easily (see Appendix 5); some have even 12-bit resolution per x- and y-coordinate. "Discrete-element" input pads, such as the "Grafacon" (see Appendix 5) have locations specified by intersections of mutually orthogonal x and y lines. Their resolution is inherently limited by the number of locations, which is of the above order of magnitude. "Continuous" type input pads, in which pen locations are indicated by continuous analog quantities, can have resolution determined by the quantizing resolution and accuracy of the A-D converters; this does not necessarily mean that analog-type input pads have superior resolution.

The tip or "nib" of the pen or pointing device should not be larger than the dimensions of a resolvable point; this tip size may determine resolution. In a light-pen for example, the field of view from which a response is obtained (the pen being light detecting) is of the order of 0.2" x 0.2", a "blunt" nib indeed (!), requiring extensive software support to accurately position a point.

2.4.3.3. Geometrical Linearity and Accuracy

Hand-drawn inputs, as noted above, are not too accurate, with, at best, several % of distortion being present; consequently up to several % of geometrical distortion can be tolerated. With transparent input pads superimposed over the associated display surface, such as in DAC-1, such distortions would be apparent by the positional mismatches between input points and resultant displayed points; but the same problems of mismatches occur when an input pad, separate from the display, is used to attempt to precisely locate some point on the display, with the user's eye being the feedback control element controlling the position. Better geometric accuracy is thus desirable although, if very expensive, is not strictly warranted.

With "discrete-element" input pads, geometrical distortion depends on manufacturing tolerances and the artwork of the masks which determine the position of the orthogonal x- and y-transparent conducting drive lines; for a 1024 x 1024 addressable grid, linearity is 1 bit in 10bits in the worst case i.e. about 0.1%. With analog-type input devices, geometrical distortion is about 0.5 - 1%, becoming progressively worse nearer the boundaries of the input area (see Appendix 5). With light pens, positioning accuracy is about ± 0.05" at best in a 10" x 10" display i.e. about ±0.5%. It can be increased by soft-ware. By scaling-up the area of interest by say "x16" and then positioning the pen, and reducing again by "x16", the positioning accuracy can be increased to about +0.03%. The operations here include "windowing", "magnifying" and "reducing" (see section 2.3.6).

2.4.3.4. Repeatability or Positioning Precision

Repeatability does not enter into the picture (no pun intended!) in "discrete-element" input devices;

unless the line-selection circuitry fails, a nominated point will always be positioned at its selected lines intersection. In analog-type devices, only long term drift may be of consequence, giving rise to positional mismatches between the nominated input point and its displayed location; this may be corrected with biasing voltages or by physical realignment of the superimposed input devices over displays. Short-term "jitter" is of little consequence as a particular figure etc is input only once within a short interval.

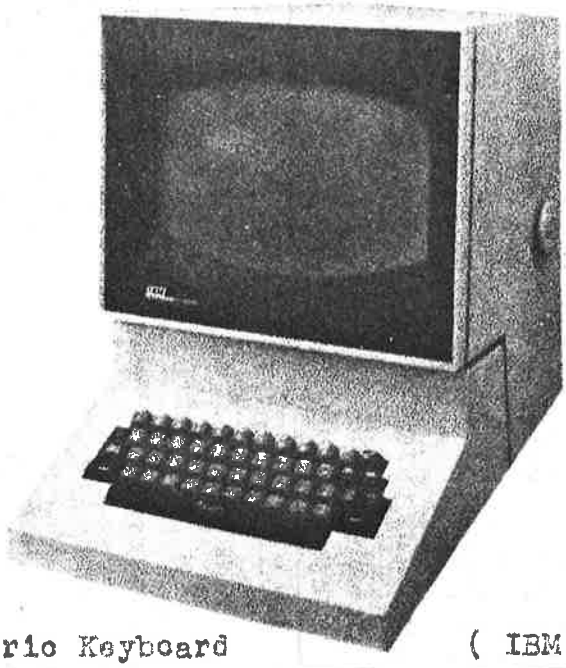
2.4.3.5. Input Area Size

The active area of the input device should be of the same order of size as the associated display area. This goes without saying in superimposed "transparent sheet" input devices positioned over display areas (e.g. the DAC-1 input device - Fig.2(a)). In a "desk-type" input device or "writing-pad", the similar sizes of the input and display areas greatly helps the visual feedback process of the user when positioning some input data at required locations, with the effect of positioning being observed on the display.

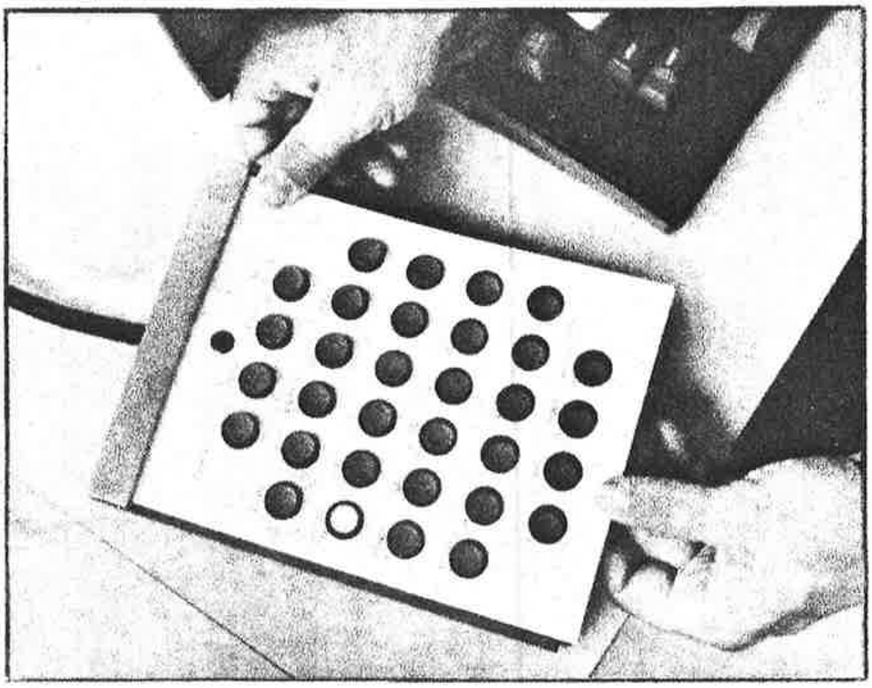
With display area sizes ranging from 8" x 11" to 20" x 20", it is not surprising that input device sizes are typically 10" x 10", with the largest (Grafacon) being 20" x 20".

2.4.3.6 Input Response Time

Hand drawing speed is probably no faster than about 1"-2" per second when tracing. With a 10" x 10" input device of 1024 x 1024 points resolution, this corresponds to about 100 - 200 points per second, or 5-10ms per point.



(a) Alphanumeric Keyboard (IBM 2260)



(b) Function Keyboard
An 'overlay' labelled for a specific application is shown being changed - it is identified by the computer by a punched code inside the frame. (on an IBM 2250)

Figure 42. IGC User - INput Devices - Keyboards

Some devices have a data conversion rate of up to 4000 points per second, corresponding to a writing speed of 40" per second; this may approach the maximum hand motion speed holding a pen, but hardly the speed at which a straight line (the simplest input vector) can be drawn.

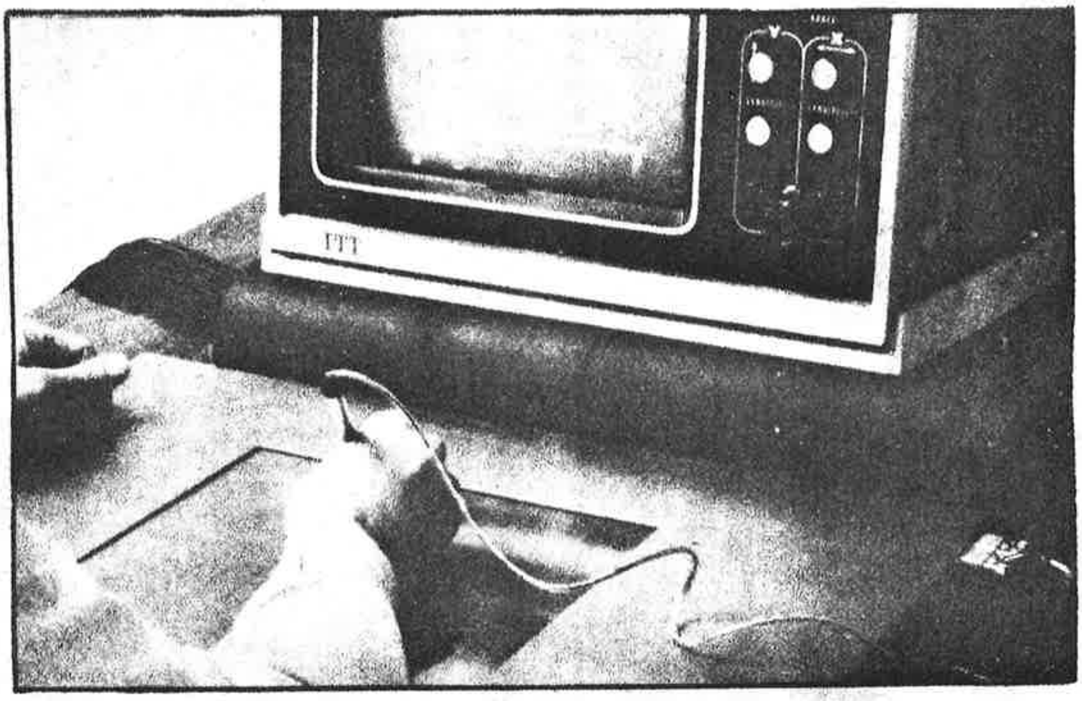
2.4.3.7 Hard-Copy Capability

When free-hand inputs are generated, the user may use a master record, or "hard-copy", from which he traces or copies, or else he may like to have a record of the input he has specified "free-hand", which again becomes a "hard-copy". A light-pen precludes tracing from a hard-copy, as this attenuates or even completely blocks out the light from the display device, from reaching the pen. For these same reasons, direct voltage-pickup pens (section 5.6, Appendix 5) do not permit a hard-copy for tracing or record purposes, to be inserted between the pen-tip and the input device surface as paper etc is an insulator. Other input devices using capacitive effects, pressure effects, or ultrasonic surface-waves, can use a ball point in the same pen housing and adjacent to the pickup "pen" device, to generate the hard copy, as a thin sheet of paper on which the master-drawing or record is on, does not greatly interfere with the device performance.

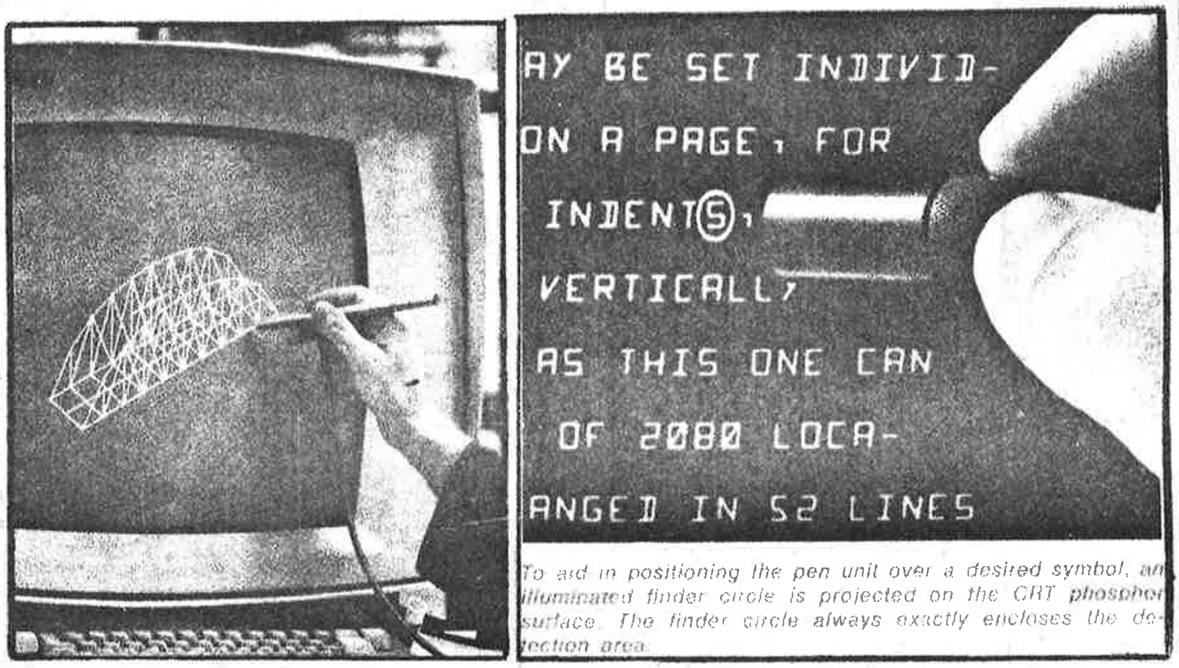
2.4.4. Forms and Types of Available Input Devices. (9,25,60,119,121,122)

2.4.4.1 Teletype Keyboards and Function Keyboards

Keyboards may be the standard teletype (Fig.42(a)) to input numerical data, labels, requests, and specify procedures or else they may be "function" keyboards or



(a) Graphic Tablet or Pad ('Rand Tablet') (from [125])



(b) Light Pens - IBM (left), Sanders (right).

Figure 43. IGC User Input Devices - Pointing Devices

buttons where specific keys or buttons are associated with specific functions or subroutines on a 1:1 basis.

For different problems (e.g. circuit analysis, graphical curve fittings etc), labels on the function keys may be changed by plastic "overlays" (9,62) as shown on Fig. 42(b). Sets of overlay panels may go hand in hand with associated programs. These overlay panels, when inserted, activate switches (different overlay, different combination of switches) which automatically call the associated program from disc program libraries.

2.4.4.2 Two-dimensional Graphic Input Pads or Tablets

These may be in one of the two configurations:

(i) the transparent input pad or surface which is superimposed over a display surface (11,123) or else has the display projected onto it (124). The advantage is the ease of positioning data relative to displayed data (see Fig.2).

(ii) the "desk-type" input pad, positioned adjacent to the display, and used as an ordinary input pad (60,125,127,128). Positioning of data relative to displayed data is much more difficult, as rapid eye movements are required from the pad to the display to help accurate positioning. One commercial form is shown in Fig.43(a).

Positioning of data and subsequent resolution may be either:

(a) due to the fact that the display area is digitized inherently by addressable locations being the intersections of discrete x- and y-selection lines such as in the Grafacon (125); in that device, locations are determined by a unique pulse train from each location, capacitively picked up by the probe tip and decoded.

(b) or the position depends on some analog quantity picked up by the pen or probe, which then is A-D converted to specify the location. For this purpose, orthogonal voltage fields (123,129) may be set up between opposite pairs of sides of the display area; or else surface waves may be alternatively launched in the x and y direction, or even sound waves may be used. (127). Electric potentials sensed by a direct-contact probe or the time delay between the launching of waves and probe pickup, determine the location of the probe. Appendix 5 lists the working principles of several available input devices with pertinent parameters.

2.4.4.3 Light Pens (25, 129,130,131)

The usual light-pen (occasionally a "light-gun") is a positioning device used in conjunction with CRTs which has within its tip a light-sensitive element which is sensitive either to the initial luminous glow of the excited screen phosphor or to a certain colour,"fast" phosphor if the screen has several phosphors. When pointed to a desired position, a pen-output pulse is generated when the electron beam passes under the pen tip. In raster-scan displays, there is a guarantee that the beam will appear at any specified location with the refresh rate period. In random-scan there being no such guarantee, the pen needs to move a point, via a "pen-tracking" program, until the desired location is reached. This is done by having a "cross" ("+") displayed somewhere in the field of view of the pen and detecting its position by detecting its generation and display by electron beam. The coordinates of the "cross" as detected by the pen, locate the pen tip position.

The average light pen has a view area of about 0.2" x 0.2" (Fig.43(b)) and hence has poor resolution unless software support is provided to "centre" the pen (see section 2.4.3.3.). A light pen is relatively inexpensive; (several $100's) but the software support required is quite extensive; for this reason it is inefficient when used to draw vectors etc.

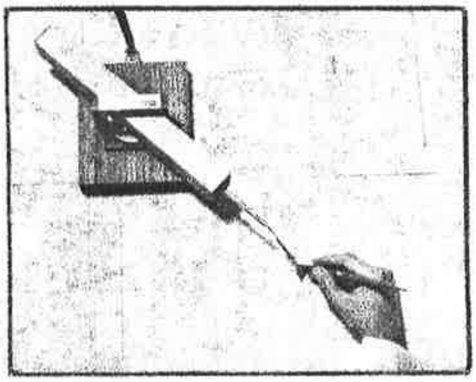
A different light pen (a light-beam generating pen as distinct from the light-sensitive pen above) has been suggested in conjunction with a Plasma Panel Display (113). By pointing the pen at required Plasma cells locations and "squirting" light at them, the cells would be ignited.

An interesting device with 3-dimensional position specifying capability is the "Lincoln Wand" (132), which can specify locations within a 4' x 4' x 6' volume. The "wand" senses its tip location 25 times per second by picking-up ultrasonic waves emitted from 4 sources located at the boundaries of the working space. An analogy between the light-pickup pen is evident; again as with the light-pen, no "hard-copy" is possible as this would interfere with the generated ultrasonic waves.

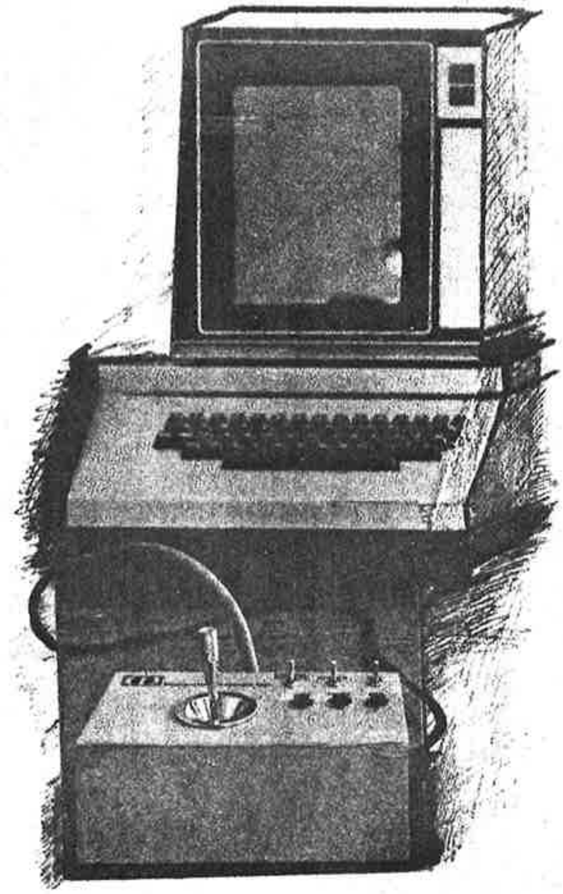
2.4.4.4. Cursors, Caret or Tracking-across Input Devices (9,119,121,129)

"Cursors" or "Carets" are symbols displayed on display screens and having shapes such as "underlines",

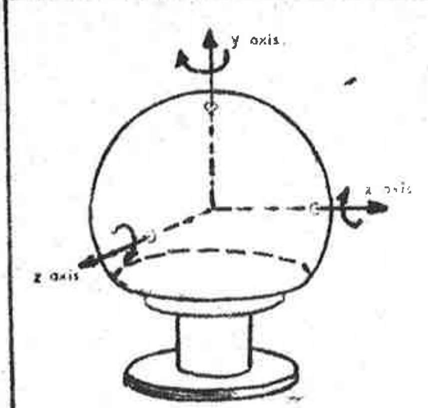
(a) X-Y Digitizer on a CRT. (Graphic Displays)



(b) Rho-Theta Digitizer (Belt Beranek and Newman)



(c) Joystick (Computer Displays)



FOR EACH AXIS	
CODE $g_3\ g_2\ g_1\ g_0$	MEANING*
1 0 0 0	Highest negative rate
1 1 0 0	Middle negative rate
0 1 0 0	Lowest negative rate
0 0 0 0	Reset Position. Zero rate
0 0 1 0	Lowest positive rate
0 0 1 1	Middle positive rate
0 0 0 1	Highest positive rate

* Considering clockwise rotation as positive

(d) 'Crystal Ball' for rotating Displayed images (from [25])

Figure 44. IGC User Input Devices - Electromechanical

squares or circles; in alphanumeric displays a cursor indicates the location where the next symbol specified will be located and displayed. The cursor position is advanced by one character space either by generating a character, or by depressing "space", "new line" and other textual positioning keys. Alternatively, "free-hand" control (that is random positioning of the cursor not made up of a sequence of keyboard "typewriter" functions) can be achieved with several available, hand-controlled analog devices, such as:

(i) "Joysticks" (119,121) shown in Fig.44(c). Hand pressure on a small vertical joystick, generates analog signals (via shaft encoders) to the deflection circuits which move the cursor or caret by a distance proportional to the deflection of the joystick within the area of freedom of movement. Alternatively a "Spherical Ball" which is hand- or finger-rotated achieves the same effect (25). Alternate forms may be spring-loaded joy-sticks; pressure on the "stick" in any direction moves the cursor in that direction on the display with a constant velocity and for the length of time that the pressure is applied.

Accurate enough to position points, both devices require extreme skill to generate graphical images with any degree of accuracy.

(2) SRI "Mouse" device (119,129). To overcome the difficulty in generating graphic data with joy-sticks this device, hand-moved, has been developed. It has a pair of orthogonal drive wheels (x- and y-positioned) driving rotary potentiometers, the movement of which generate voltage signals proportional to the total rotation of

the wheels and hence to the distance moved by the "Mouse". Originally developed to select elemets from arrays, it can input graphics by following the required trace. Its physical shape (see Fig.1(b) limits the accuracy of tracking. A block diagram is shown in Appendix 5.

2.4.4.5 Electromechanical X-Y Digitizers (119)

These are the "dual" of X-Y plotters. The hard-copy master drawing is traced by the pen-tip and data is input by depressing a switch . X-Y Digitizers are used mainly for precision graphics. The digitizers come in all shapes and sizes from an input area size of about 8" x 10" to about 6' x 10'; sometimes they are superimposed over the display area as shown in Fig.44(a). In another case, the pen is on a cantilevered arm with the output being in polar coordinates; not surprisingly it is called the "Rho-Theta" Transducer (Fig.44(b)).

The main disadvantages of these devices are the mechanical encumbrances and "stiffness" in use in guiding pens affixed to mechanical members.

2.4.4.6 Phototube Input Devices (86,90)

These are Vidicon, Image-Orthicon tubes etc used in TV-camera-like applications. When focussed on some hard-copy diagram they produce appropriate signals for display generation, particularly for raster-scan TV displays. Both raster scans (i.e. input-device scan and display-device scan) must be synchronized to the same timing source. These are strictly not user-interactive devices but are mentioned for completeness' sake.

2.4.4.7 Hard-copy Scanners (11,112,133)

In certain problems (DAC-1 automotive design for example) or certain comprehensive MCG systems, working files and libraries of graphic data in hard-copy graphic form, such as on film, exist; these may need be called on during a problem run. Although they may be called up by keyboard action or light pen action, the mechanics of the process of call-up and the devices used will be briefly mentioned.

Films are stored in a film library, and a particular frame may be located and brought for input purposes by electromechanical access devices (each frame being indexed). The input device is a "flying-spot" scanner or some such device, either raster-scanned or random scanned (i.e. "vector following"). The electron beam spot, scanning the screen surface, is very intense and of short duration (100's of nano-seconds). The film is inserted between the scanner and a detector, usually a phototube. Data on the film occurs where the film is transparent; thus at the location where the "flying" spot coincides with transparancies the phototube registers an output. The coordinates of this output are at the coordinates of the "flying" spot and hence are the coordinates of the graphics data. More sophisticated software can "track" a vector (cf. random-scanning vs raster-scanning in display devices) thus reducing the time to read-in a graphic record.

2.5. GRAPHIC INFORMATION and DISPLAY STORAGE

2.5.1 General

The fourth major requirement of an IGC, after the ability to accept manually-specified user graphic information, the positioning of graphics information on the display device, and the generation of alphanumerics and vector information on the display, is the short and long-term maintenance of the displayed information.

A user requires a finite time of several seconds to "absorb" or comprehend a newly displayed image of vectors and characters; in addition he may require the displayed information for several minutes during which he examines it in detail, evaluates it and modifies it if necessary. Even the taking of permanent "hard-copy" of displayed output results may require a sizeable fraction of a second.

CRT displays are in the main, "dynamic" or "soft-copy" displays, that is, unless they are "refreshed" periodically, they disappear. Direct-View-Storage-Tubes (DVST's), have inherent display storage (B.5); the graphic information need only be "written-in" once, and remains visible for up to several hours (although the recommended duration of storage is less than 15 minutes). The disadvantage is that any new display image updating requires the generation of the complete new image; this makes DVST's excellent as far as displays go but at the cost of user-display interaction. Other electro-optical effects used for displays such as E-L, LED's also require refreshing. Plasma-panel Displays have inherent storage.

CRT displays are generated at, or above, the "flicker" rate (see section 2.2.2.8), usually at the rate of 25Hz or more. Display regeneration implies that the displayed data need be available during the time the display need be maintained; a <u>display refresh memory</u> is required for this. This is usually distinct and separate from the memory which stores graphics data from past displays (for recall) or for future displays, or the fast storage which holds the partial graphics information which the CPU at that instant is processing.

The requirements for such a refresh memory are as follows: -

(i) capacity - should be capable of storing the maximum number of data points capable of being displayed.

(ii) readout speed should be such that the complete memory should be readout at refresh rates.

(iii) the up-dating or writing-in should be random (i.e. memory should be random-access).

(iv) economical and reliable.

These requirements are relatively modest compared with the requirements associated with the CPU memory; consequently to eliminate tying up fast storage and excessive time loading on the CPU, separate display-refresh memories are used.

The current philosophy is to consider refreshability as of prime importance, at the expense of updating flexibility. This is not only a result of economic considerations but also because the majority of display units at the present are alphanumeric displays for text or alphanumeric tabular application, without graphic capabilities. Serial memories are thus quite adequate for the purpose of refreshing. Secondly, in multi-display console configurations, for display hardware economics, TV-raster scan techniques are most advantageous (86,87,88), with modified TV equipment being used. As TV diplay generation is serial in nature, again serial memories can be used.

Other factors which determine the type of refresh memory used inclued the type of system configuration used, whether "stand-alone" IGC, or multi-console systems (see section 2.6), the type of data link between the CPU and the IGC, the form of the other levels of memories storing other graphic information, and whether serial or parallel data transfer is being used. All of these factors determine whether a "display processing" subsystem is required, which, amongst other functions, converts data formats and data rates between the CPU and the refresh memory; this is mentioned more fully in section 2.6.1.

Memories in this context will be divided into two main catagories:

(i) Random Access Memories (RAM's)

(ii) Serial Access or Cyclic Memories (SAM's)

DVSTs (Direct View Storage Tubes -Bistable) having inherent storage will also be mentioned, along with some miscellaneous image storage-effect memories. Figures 45 through to 48 show system configurations using these major forms of memories for display refresh.

2.5.2. Random Access Memories (RAM's) (134,135,136,137)

In a RAM, there is no predetermined order in which memory locations are accessed. The most common form of RAMS are magnetic core stores, although with Large Scale Integration, thin film and semiconductor memories are becoming economically feasible. With cycle times down

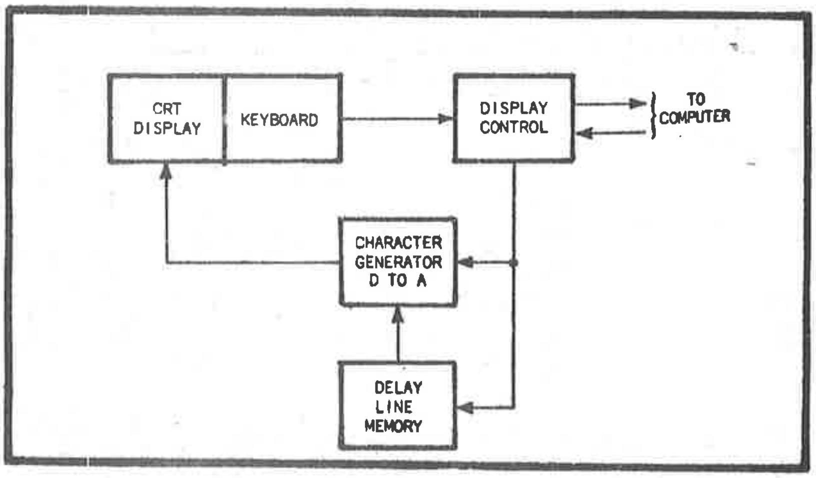


Figure 45. Simple Alphanumeric Display Using Delay Line Storage for Display Refresh

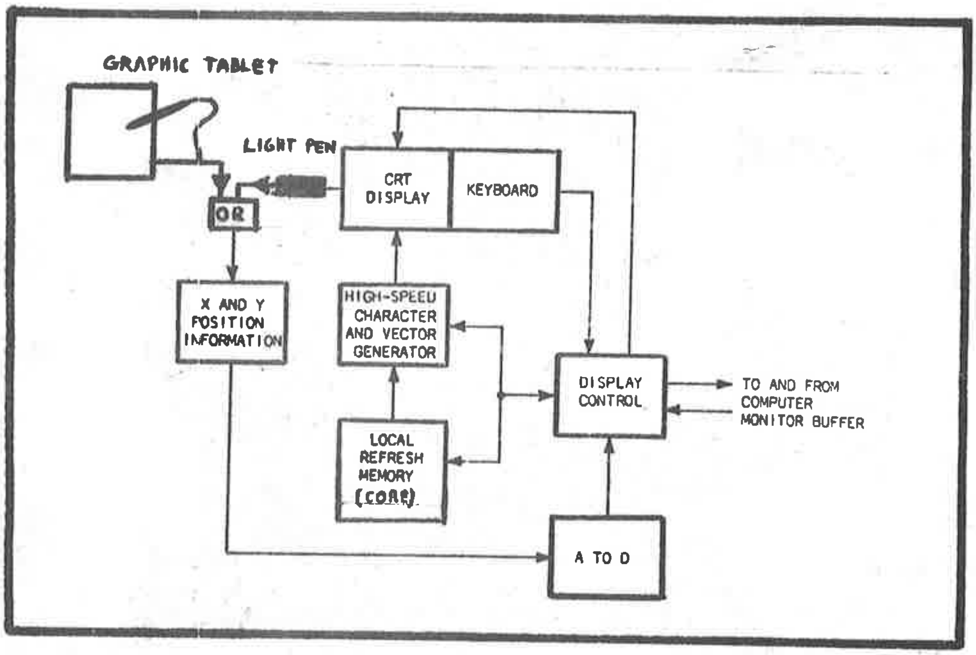


Figure 46. Full Graphic IGC Using Random Access Core Storage for Display Refresh

to 500ns memory cores are the fastest and most reliable (as far as precise addressing and location of information go i.e. repeatability). Cost at present is about 3 cents per bit, including driving circuitry (135).

Data can be read-in from the CPU (e.g. processed graphic results) or else read-in from manual graphic input devices. Up-dating display information can be inserted by selectively erasing the corresponding storage locations and writing-in the new information; otherwise for a wholly new display image, the memory is wholly cleared and the new display data is written-in.

In displayed images containing repetitive elements at various positions ("replicated elements" - see section 2.3.6.5), the "master element" or "master drawing" need be stored only once in RAMs; then, only the location of these other positions together with a reference to the location of the master element need by stored to generate the required repetitive elements. In serial or cyclic memories (see below section 2.5.3), the display information for each replicated element need be stored as many times as it is to appear on the display image.

Trade-offs between memory, hardware and speed are possible:

(i) In random-dot or stroke graphics generation, graphic data is stored by either storing the graphics coordinates or by storing the "Δx" and "Δy" increments for continuous vectors.

Say a display image consists of 5000 characters (about the maximum capability of the best displays). With a 64-character repertoire, 6 bits are required to specify a character; and for random positioning, another 20 bits are required to specify the location of each character (10 bits for each coordinate on a 1024 x 1024 location display). A total of about 1.3×10^5 bits of storage is required!

(ii) In raster scan positioning and generation, with minimal hardware required (no D-A conversion required as in random scan above; only good synchronization between raster scanning timing and memory interogation), for the same number of characters and same display size, the amount of memory required is $1024 \times 1024 \approx 10^6$ bits. The memory is sequentially interrogated at refresh rates (required memory cycle time of 30ns in this case for acceptable refresh rates!); the memory locations storing the display information give pulse outputs which intensity-modulate the display CRT electron beam.

These two hypothetical extreme cases do not take into account display response times, available memory cycle times etc, but they do give a comparison of the amount of storage required between random and ordered scanning schemes in display refreshing. In random-scan refresh modes using RAMs, refresh rates depend on the amount of information to be displayed at any instant. The less there is to display, the more often within a given time it may be refreshed. Conversely, high density information displays have low refresh rates and thus annoying display flicker.

2.5.3 Serial Access Memories (SAMs) or Cyclic Memories (135,138,139,140).

2.5.3.1 General

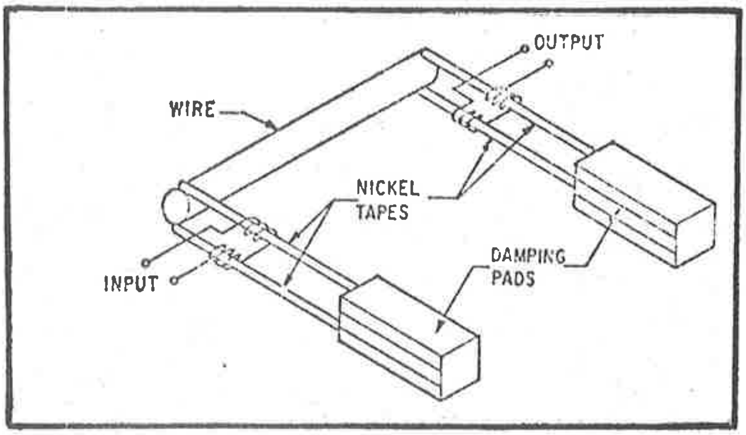
As mentioned previously, the main advantage of serial memories is cost, at the disadvantage of poorer updating flexibility. In alphanumeric displays where serial typewriter formating is used, or in raster-scan refreshing, this updating inflexibility is not a great disadvantage.

2.5.3.2 Simple SAMs

The most common form of SAMs are delay lines made of glass, quartz and magnetostrictive materials or those using ultrasonic surface wave effects (139). Mainly used in small desk-top alphanumeric graphic consoles, their capacity is 1000-2000 characters, refreshed at 25-30Hz.

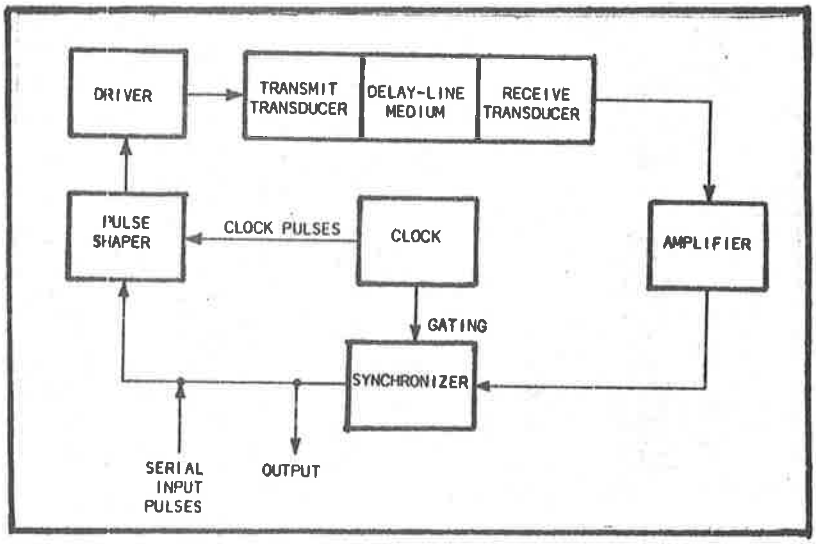
Once inserted into such a delay line SAM, data propagates down it and emerges at the output (after one refresh cycle period), only to be reinserted back into the memory (see Fig.49). Only at the input (and thus at the output) can display data be modified by erasing and rewriting with new data.

The address of a particular bit is serial and is proportional to the time between the instant of the whole display file entering a new cycle at the input of the memory, to the instant the bit in question re-enters the memory. Time markers can be used to indicate various regions in the display by initiating timing cycles which help specify the time in a cycle (from the beginning of the cycle), and hence the locations where new data is to be inserted. A simple block diagram of a delay line memory is shown in Fig.49(b).



Current pulses at the input generate magnetic fields causing magnetostriction in the nickel tape (corresponding to the 'transmit trans-ducer') which are transferred to the wire as stresses and propagated to the end where the reverse takes place (the 'receive tranducer').

(a) Schematic of Magnetostrictive Delay Line



(b) Simple block diagram of Delay Line Memory

(from [138])

Figure 49. Serial or Cyclic Access Memory - Delay Lines

Temperature and pressure changes etc. change the effective delay line length and hence the circulation time and thus the absolute position corresponding to a given time marker. Repeatability is thus poor; fortunately in alphanumeric displays, absolute positioning is not of great importance. Typical maximum data rates, at 50Hz refresh rates for 2000 characters specified by 6 bits, are 6.10^5 bauds (excluding the timing markers) which fall well within existing technology. Typical cost is about 0.4c per bit for a magnetostrictive delay line (139).

2.5.3.3. Shift-register SAM

Discrete component SAMs, with MOS elements currently being in vogue (from economic viewpoints, with a projected cost of 0.6c per bit (135)), seem very promising; they are not in use yet as array sizes with the storage capacity required for display refresh are not yet manufactured. The discrete storage elements can be arranged as **a linear shift register**; extra control or input terminals at specified shift register intervals lead to shorter update times compared with the update times of the above form of continuous delay lines. A 2000 character display with 6-bit character specification (excluding spaces etc) requires a 12000-bit shift-register; MOS switching speeds are adequate for a data rate (refresh data rate) of 6.10^5 bauds.

2.5.3.4 Semi-serial SAMs (86,87,88,135,141)

Magnetic drum stores and video disc stores come under this category. Storage consists of many serial tracks (of some ferrite material) with stored data, each track capable of being separately accessed, written-in and erased in a fixed manner. The disc or drum rotates past a set of reading-writing heads, one for each track.

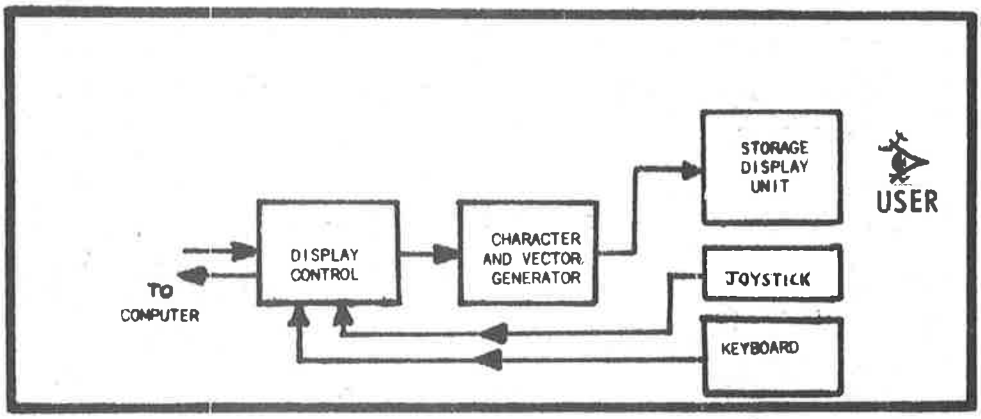


Figure 47. Graphic Display Terminal Using a DVBT

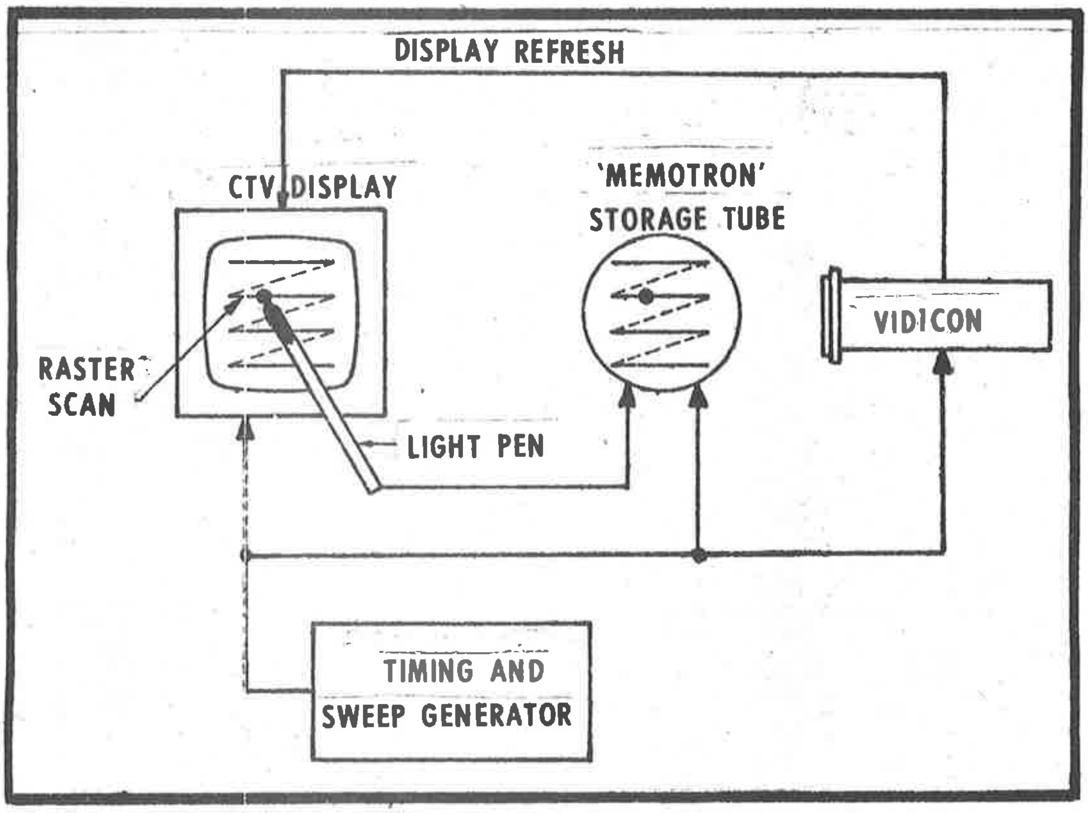


Figure 48. Graphic Display Terminal Using Storage Tube and Vidicon Camera for Display Refresh

(based on [90])

Multiple tracks enable data formating such as for typewriter format; thus alphanumerics can be generated on a 5 x 7 matrix, either by dot or stroke generation. Using 7 tracks in parallel, alphanumerics can be refreshed directly rather than storing strings of 6-bit code to activate character generators. For multi-console configurations, only one character generator is thus required rather than a character generator for each console (87).

For updating raster-scan displays at acceptable refresh rates, the finite rotation speed of drums and the finite packing density of locations necessitates "interleaving" of several adjacent tracks for each display, requiring say 4 - 6 tracks or more, if characters are to be directly stored (86,87). As discs are faster and have greater storage capacity than drums, they are used for this purpose rather than drums.

2.5.4 Direct View Storage Tube (DVST) (134,142,143,144,145,156)

2.5.4.1 Storage Tubes.

DVSTs have inherent storage and thus need no external refresh memory (see Fig.47). They are gaining wider use, their main virtue being low cost (around $2500 for an 6" x 8" display, and $8500 for a DVST including keyboard, and vector generators (144,145)).

The main disadvantage is the lack of true user interaction as with a light-pen, for example. Selective erasure in existing DVSTs is not yet possible, although a "write through" feature is possible as explained below. Whenever selective erasure or updating is required, the whole display need be erased (within 0.5 seconds) and then the new updated image displayed. Storage is of the order of several hours but for fast display erasure (0.5 sec)

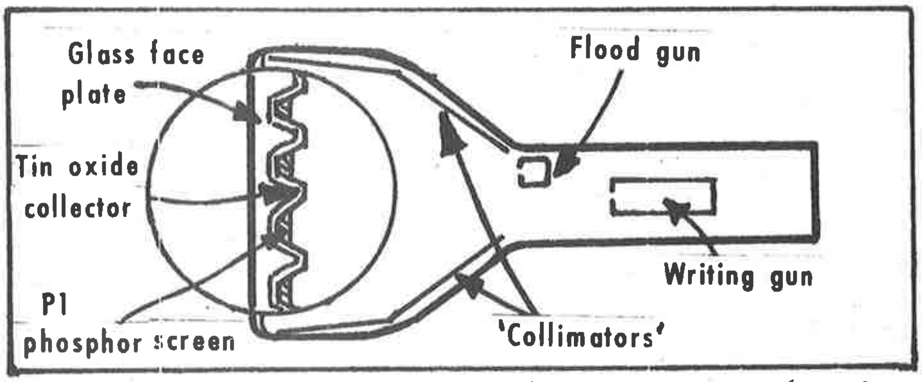
15 minutes is the recommended viewing time. Manual graphics input can be via a graphics input pad or a "Mouse" or "joystick" (see section 2.4.4.4). Resolution is such that up to 4000 legible characters can be displayed on a 6" x 8" screen.

Briefly storage is due to a bistable effect when a phsophor screen is "written on" with a "fast electron" beam (i.e. containing electrons causing secondary electron emission on striking the target), and then illuminated or "flooded" by "slow" electrons (i.e. electrons not giving rise to secondary electrons); these latter electrons supply just enough energy to keep the initially excited areas (by the fast electrons) phosphorescing.

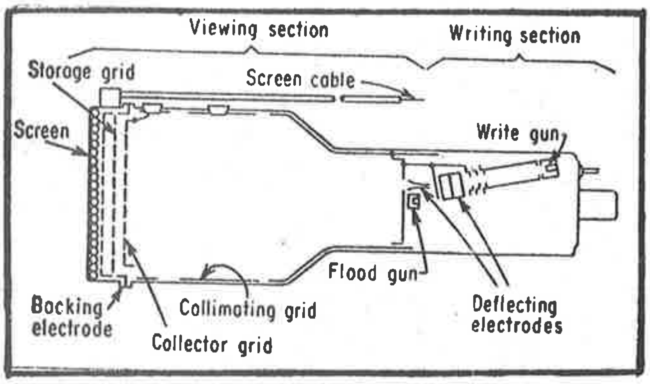
As can be seen in Fig.50(a), the actual phosphor target is quantitized, not only to improve addressability and repeatability, but in conjunction with the raised collector, to result in a more even intensity spot over the display area, by ensuring the electron beam falls perpendicularly to the screen surface.

Beam positioning traces (via a "caret" symbol) are not stored as the beam current for this is lowered and is continuously swept in a localized pattern; no storage takes place. When the beam is positioned, write-in can be accomplished normally.

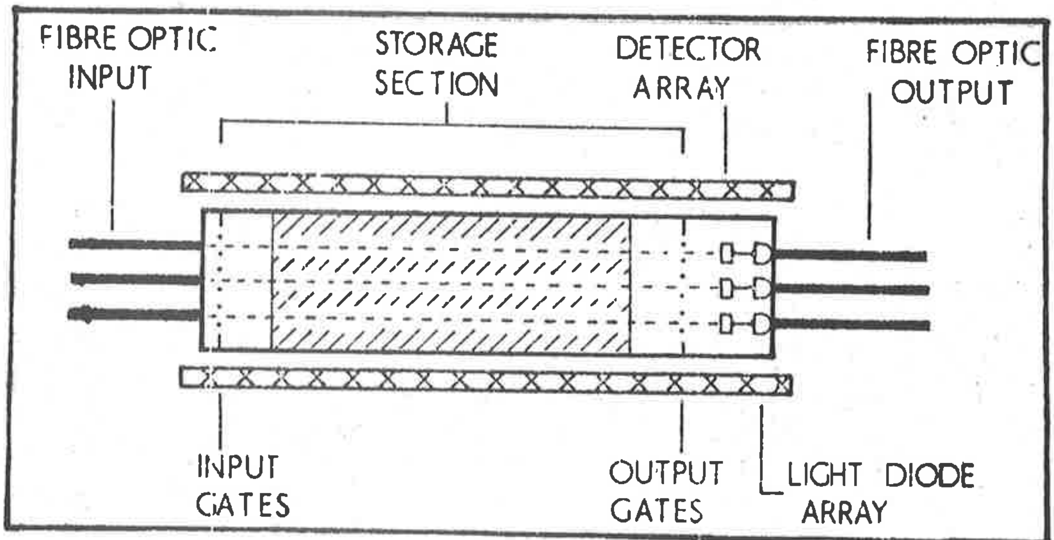
These DVST are distinct from the older forms of DVST which had separate storage meshes in front of the phosphor screens, but which had no write through facilities (Fig.50(b)).



(a) Direct View Bistable Storage Tube (as in ARDS)



(b) Direct View Storage Tube ("mesh type" - RCA 6866)



(c) Proposed Electron Image Tube (from [151])

Figure 50. Various Analog Graphic Image Stores

DVSTs are becoming very popular as display terminals, their main advantage being low cost and good quality display. User-display interaction is however rather limited, as mentioned above, restricting their use mainly to display purposes.

2.4.5.2 Other Analog Storage Tubes

Certain older-type storage tubes, namely the "Memotron (93,147,148) (although there are several other forms) (93)) have been implemented for graphics information storage, particularly in the ARTRIX system, mentioned in sections 2.3.6.3. These tubes can, when properly designed, store and display images for many hours; the contrast ratio is low, about 3:1.

As in DVSTs, two beams are used, a high-voltage accelerated beam of about 3KV, used for writing, and a "flood beam" of "slow" electrons accelerated by about 300V to keep the image displayed. The storage target, unlike in DVST, is a separate fine metal mesh, placed about $\frac{1}{8}$" from the phosphor screen. The functioning and construction of the tube are adequately described in reference (147), while an IGC application is given in (90,149,150) and shown in Fig.48.

Reading out the stored information for refreshing the display is done by imaging the visible stored image on a Vidicon-tube TV camera which is raster-scanned in synchronism with a TV-type display device. Writing speed is of the order of about 2 x 10^5 inches per second (it is a 5" diameter tube) which is compatable with the TV-raster display rate of the TV monitor display; erasure of the display takes about 0.2 seconds. A light pen is used to write in new information. Selective erasure of display information is done by using three Memotrons as described in section 2.3.6.3.

Such purely analog memories suffer from inherent geometrical distortion (2-3% distortion) and poor repeatability due to temperature and voltage supply variations. They are adequate for applications such as ARTRIX was designed for, (analog information processing and analog vector and display specification and generation with a CPU), but are inadequate for the usual IGC systems where vectors are specified and position-controlled digitally.

2.5.5 Miscellaneous Image Storage

Various "image stores" have been proposed which are based on low-velocity electron beam "segments" emitted from photocathodes on which the image, or drawing requiring storage, has been focussed. These electron beams are serially fed into long tubes having strong axial magnetic fields. These electron beams follow the lines of the axial field backwards and forwards within a certain region, being "reflected back" by energizing fine meshes ("electron gates" (151)) across the tube section at both ends of the tube regions (151,152). It has been reported that 100 x 100 dot array images can be inserted at any one time (100 x 100 dots in parallel) and several (up to 100) such arrays can be fed in sequentially. Thus within one tube, 100 segments of 100 x 100 dot arrays can be stored and read out or erased at the instants when they reach the appropriate mesh. A two dimensional optical analog to a delay-line can be seen. Fig.50(c) shows such a proposed tube. Still in the experimental stage, cost of tube and fibreoptics (to specify the 100 x 100 dot array) for readin and readout, make them uneconomic for at least several years to come.

2.6 IGC SYSTEM CONFIGURATION

2.6.1 General IGC Configuration (9,25,61,62,76,106,145, 153,154,155,156,157,158)

A typical IGC system configuration is shown in Fig.51. It includes:

(i) input devices such as:
- -pointing light pens (although a joystick or trackball may be used instead).
- -teletype keyboard and programmable or variable "overlay" function keyboard.
- -a graphical input pad.

(ii) a CRT display device.

(iii) display generators such as:
- -vector generators
- -character generators
- -circle and other special function generators
- -position control for character positioning and vector positioning.

(iv) the IGC Processing and Control Unit.

The control, specification and maintenance of the displayed image is achieved by the IGC Control and Processing Unit which for our purpose includes the display refresh memory. This unit may consist of a small modified control computer or "mini-computer" with an extended memory, either random-access or cyclic.

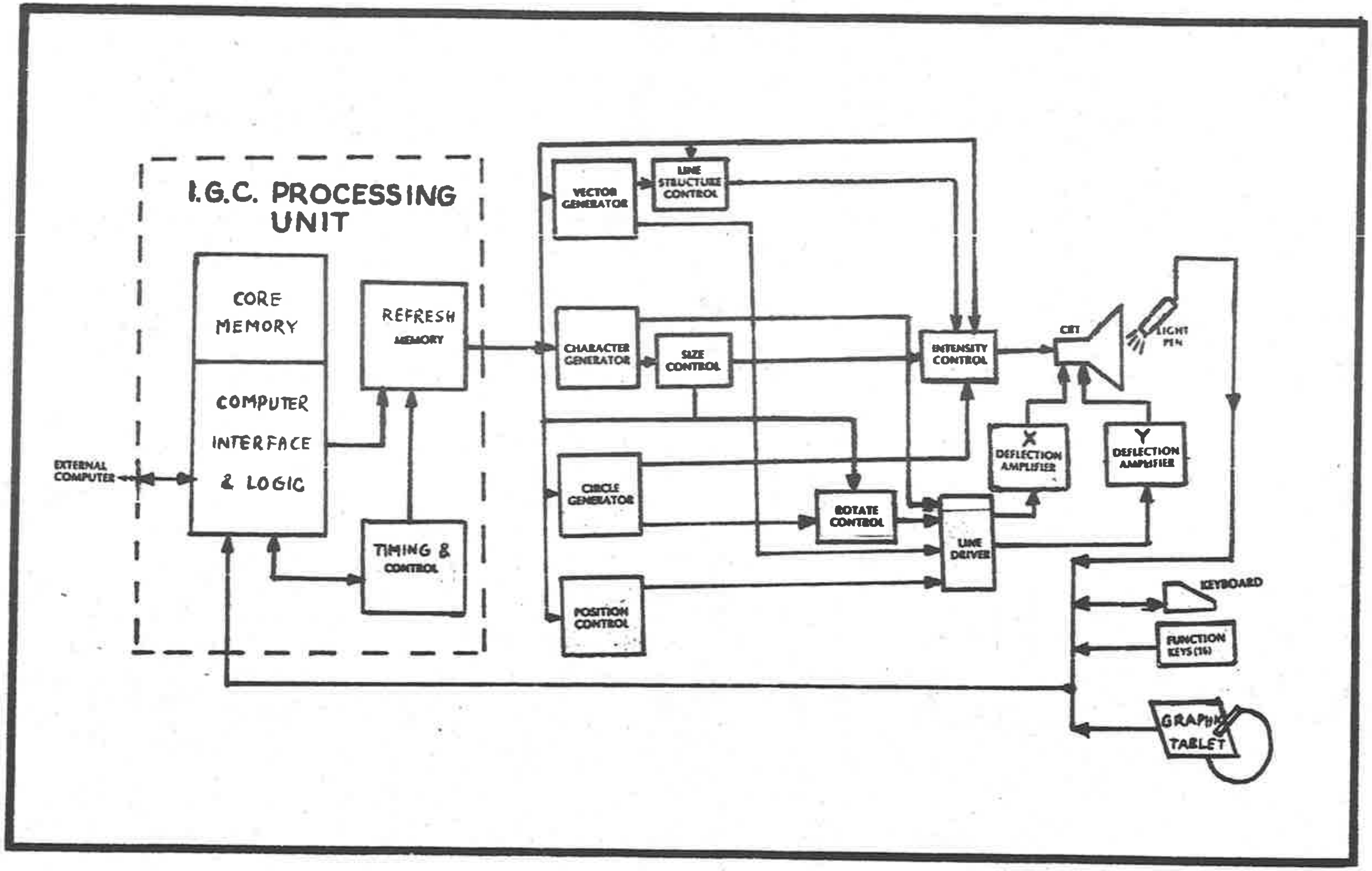


Figure 51. Typical IGC System Configuration (Digital System) (based on [61])

Facing p.106

The Computer Interface Logic in conjunction with the control and timing unit performs:

(i) data format conversion between the CPU and the IGC from the complex graphic data structure and commands of the application program, and generation of graphic routines.

(ii) data rate conversion for speed compatability between the CPU data link and the IGC.

(iii) handling of the interupt functions associated with the display and input devices and allowing input of data therefrom.

(iv) routes the display refresh data into the appropriate display generators.

(v) provides the logic for subroutining and other IGC commands (see section 2.8.3).

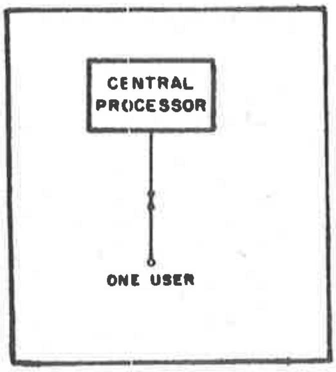
(vi) depending on the remaining core store available, simple computations may also be made within the IGC. Some typical IGC commands which are processed within the IGC processing unit to activate the various display generators are given in section 2.8.3.

The IGCs shown in Fig.1 all have the general system configuration described above. Appendix 4 lists the characteristics of several commercial IGCs.

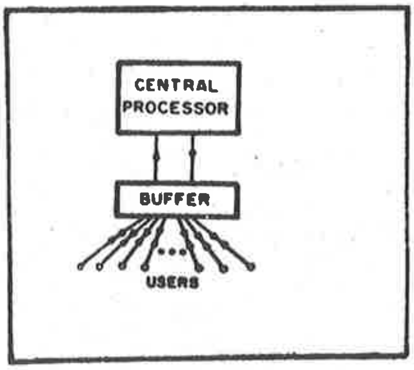
2.6.2 IGC - CPU Configuration

2.6.2.1 General (41,154)

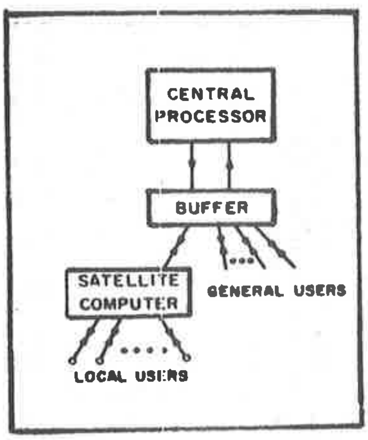
The possible CPU - IGC configurations will be briefly mentioned as the requirements for the IGC, including the type and amount of storage required and also the capability of the CPU-IGC data link are determined by the configuration (154). For example, a large CPU servicing several widely separated IGCs, may have only telephone line data links with data rates of 1200-1400 bauds;



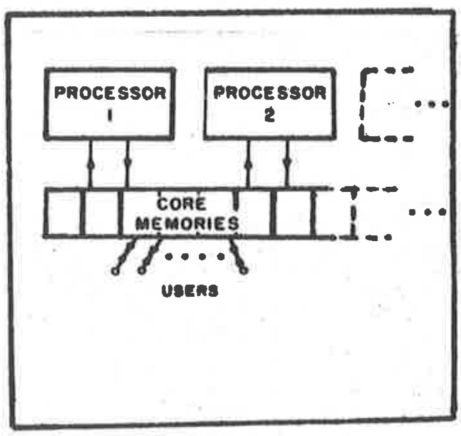
(a) Stand-alone configuration with dedicated computer.
(eg. DAC-1, SKETCHPAD)



(b) Multistation configuration
(eg. Project MAC, VICTAB)



(c) Satellite computer configuration.
(eg. IBM 2250-2480)



(d) Memory centred approach.
(eg. MULTICS System)

Figure 52. Possible CPU-IGC Configurations (from [41])

on the other hand a CPU servicing a single experimental IGC in the same building has no such data transfer rate restrictions. Consequently speed buffering, buffer storage and serial-parallel data conversion problems exist in one case but not in the other.

There are 4 major configurations possible (41) listed roughly in their chronological order of development:

(i) stand-alone configuration (CPU - single IGC).

(ii) multi-station configuration (CPU - several IGCs).

(iii) generalized multi-station configuration with associated satellite computer (CPU-satellite computer - several IGCs).

(iv) MULTICS approach (several CPUs - several IGCs).

These are shown in Fig.52.

The common features of the above configurations and the factors which perhaps necessitated their progressive developments are: -

(i) the necessity of buffer registers and buffer storage to match data rates and data formats between CPU and IGCs.

(ii) the necessity of CPU processing-time economy, display maintenance, display updating, and graphic input operations required the implementation of "satellite computers", "buffer computers" or "graphics interface computers" (88).

Thus economies of operation, particularly in time-shared, multi-console systems, are achieved when the CPU is accessed only for numerical calculations and display file generation while graphics generation display maintenence, and I/O control is either performed by satellite computers or by the IGC.

2.6.2.2. Stand-alone or Single IGC Configuration.

This configuration consists of a single graphics console directly under CPU and user-control; all display generation and display refresh, and numerical and graphical processing is done in the CPU. The earliest and least economic approach, it is warranted only for special purpose systems such as in DAC-1, where the amount and degree of interaction and fast update responses require full CPU control. The CPU serving a single IGC is sometimes called a "dedicated computer".

2.6.2.3 Multi-station Configuration

A centralized CPU may control and service several tens of terminals on a time-sharing basis; not all of the terminals need be IGCs - more often than not, they are teletype terminals. A buffer performing data format-conversion, data-rate matching and associated storage is necessary between the CPU and the console or terminals. The number of IGCs capable of being serviced depend on the CPU capability and complexity of tasks involved in the generation and maintenance of the display; thus only one or two consoles with SKETCHPAD-type capability can be accomodated on the one CPU (Project MAC), whereas several hundred alphanumerical consoles displaying current textual files with keyboard updating can be accomodated by a single large CPU. The VICTAB System (Victorian Totalizator Agency Board) is an example of the latter (159); the latest horse betting information and other relevent data is displayed (and keyboard updated) on 96 alphanumeric display consoles (raster-scanned) with a CDC 3300 as the central processor.

Remote terminals with vector display capabilities need more stringent data link requirements than alphanumeric displays. Typically in a 1500 baud telephone line data link about 120 characters per second can be transmitted (obviously a local display refresh memory is required). For modest vector drawings consisting of several thousands vector increments (specified by two "x" and "y" increments of 6 bits each, say) the diagram data transmission time would be in the order of half to one minute, unsatisfactory to even the least demanding users. Consequently in remote graphic displays, coaxial cables provide the data links.

2.6.2.4. Peripheral or Interface Computer Configurations.

ICG-CPU configurations using "peripheral", "satellite", or "interface" computers are currently accepted as the most economic solution for MCG; they can be associated with single station display terminals or multi-station situations. An extension of this is the "cluster mode", where several peripheral computers are controlled by a CPU, each peripheral computer controlling several stations. Economic, small but efficient control computers (costing around $8000- $25,000) are readily available (160); they relieve the CPU of time consuming (and hence cost consuming) I/O functions control, can perform formating of graphics data for display device compatability, generate the required data, and even may be used to perform small computations; they also perform speed buffering and data format conversion. The OLCA system for MCGCD (Appendix 2) used a PDP-5 as the pheripheral computer servicing a single IGC console; an IBM 7094 was the CPU. Recently, "INTERGRAPHIC", an interface computer specifically for IGC applications, has been proposed and built (83); up to 50 display terminals may be serviced by it.

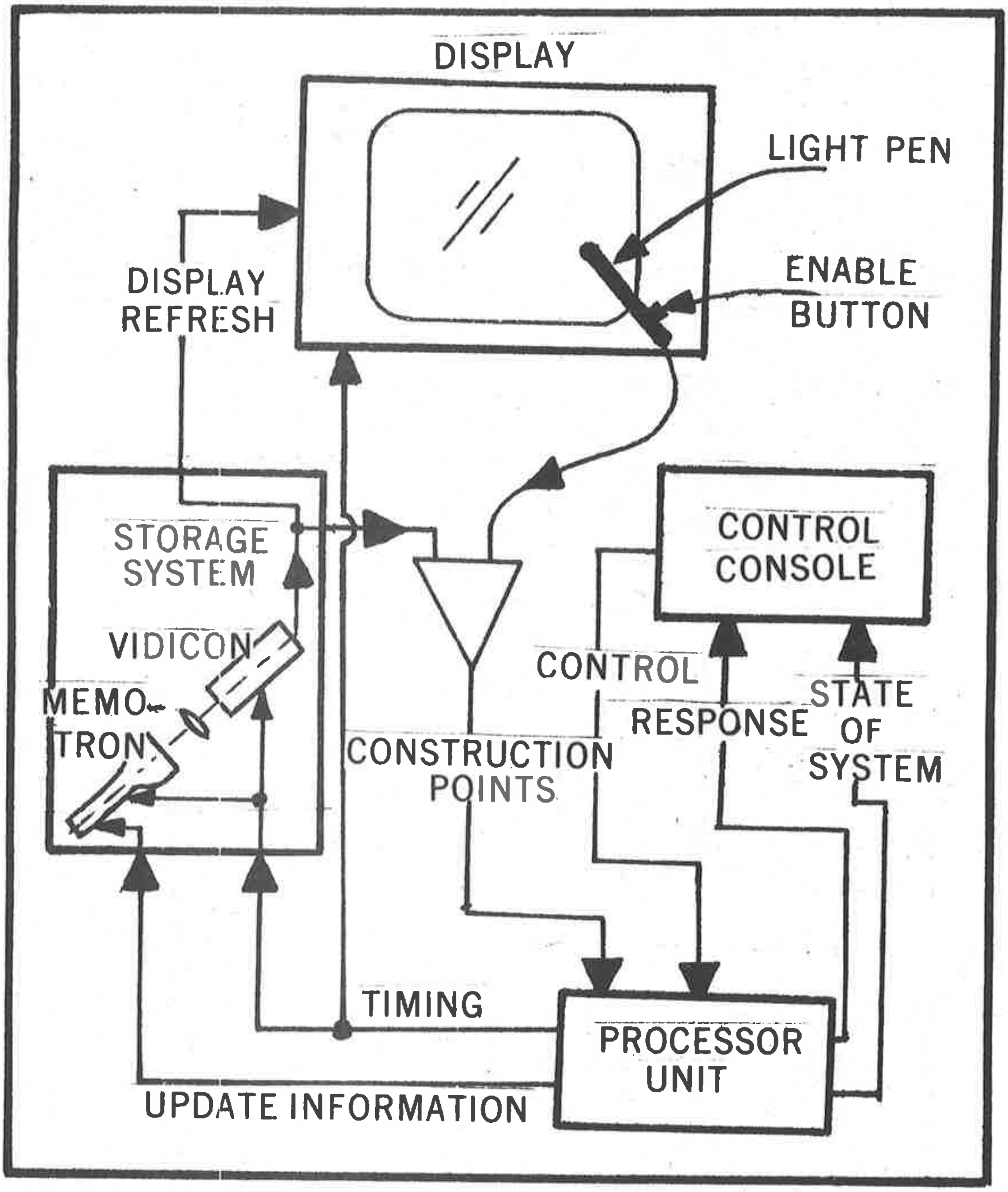


Figure 53. ARTRIX IGC System Configuration (Analog System)
(based on [150])

2.6.2.5 Multi-station - Multi-CPU Configurations

In large computing installations, terminals and the available central processors may be all interconnected via a very large multi-level memory; this ensures that at any instant no CPU is overextended. It may be considered an extension of the multi-station case above. The "MULTICS" approach (161) is an example of this; it is a "descendent" of Project MAC at MIT.

2.6.3 Analog Graphics Systems - ARTRIX (90,149,150)

All of the above systems and configurations are digital; processing and graphics specification (both user-specification of graphics and, until the final generation of the deflection waveforms, the system graphics specification), are digital. Yet CRTs, by far the most common of display devices, are analog in nature; user interaction and user graphics drawing is also inherently analog. The ARTRIX system, conceived at the U. of Illinois, attemped to make use of these facts. It was a feasibility study for a "self-contained hybrid graphics processor". User specified vectors, circles and arcs, generated by analog means (via Lissajous figures described in section 2.3.3.4) were manipulated to implement all graphics constructions requiring a "straight edge" and a circle, without a digital central processor. The output of ARTRIX however can be input to a digital computer. The aim was to minimize costs by using analog hardware for tasks which are inherently analog.

A line and circle generator were used (analogous to "straight-edge" and "compass") based on Lissajous figures. The display storage was based on Memotron storage tubes with Vidicon camera tubes as readout (section

2.4.5.2). The operator-controls and display interaction are with a light pen, used in conjunction with a commercial TV 23" monitor; about 15-20 control buttons and light indicators specifying functions such as "circle", "line" "erase" etc are also provided.

Poor resolution and drift due to temperature posed problems; but as a feasibility project it provided interesting economic solutions to fast vector and curve generation and user-graphics interaction. The system is shown in Fig. 53.

2.7. HARDWARE - SOFTWARE TRADEOFFS

2.7.1 General (82,88)

The emphasis within this chapter and in this thesis is on IGC hardware; software however cannot be ignored, not only because it provides the means for specifying and processing the problems requiring solution, but because of the possibilities for software-hardware tradeoffs. Thus on the one extreme, with even the most minimal graphics display hardware (raster-scanned TV-display devices) but with "infinite" memory available and suitable software (however inefficiant), even the most complex MGC problems can be solved. On the other extreme, an "infinite" array of hardware devices to generate any desired line, complex curve, alphanumeric etc and other special purpose processing hardware, activated by minimum software can again in theory solve any complex MGC problem which has a solution. In between these two extremes there is ample scope for tradeoffs. The existance of IGCs with inbuilt vector and character generators, compared with the early, simple display systems, indicate that the natural compromise between hardware and software

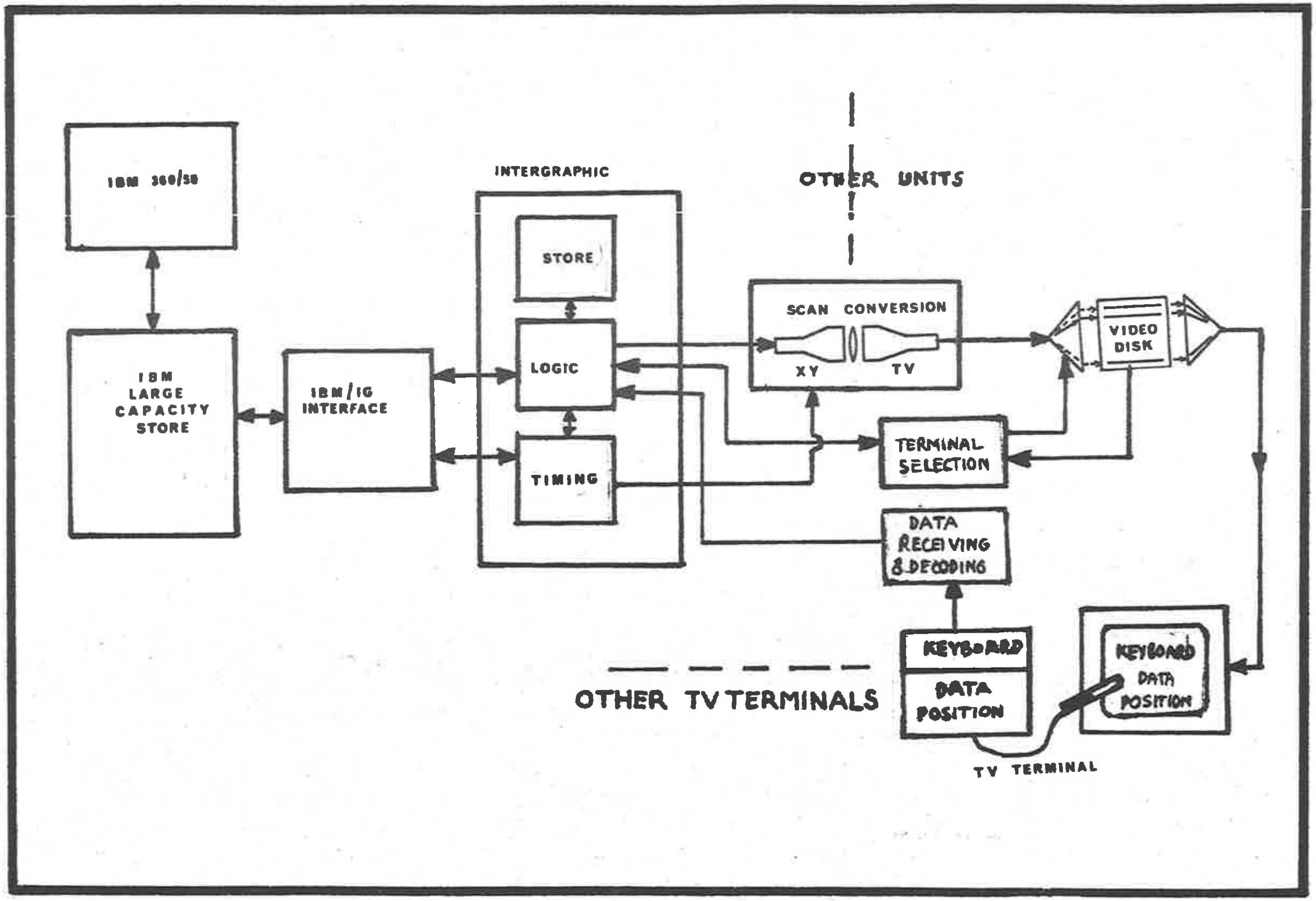


Figure 54. INTERGRAPHIC - Microprogrammed IGS Satellite Computer with System Configuration. (based on [88])

has been recognized and implemented.

2.7.2 "INTERGRAPHIC" - Efficient Tradeoffs

An interesting and elegent solution (at least for multi-station configurations) to the above problem of tradeoffs was reported by Rose (88). This was a special-purpose graphics interface computer, "INTER-GRAPHIC," linking a central computer (IBM 360-50), performing the problem processing, with up to 50 user consoles with interactive graphics capability (via a light-pen and key-board). The function of INTERGRAPHIC was to control, generate and refresh the displays; rather than use hardware vector or character generators, micro-program sequences from a fast Read-Only Memory (ROM) generate the necessary vectors and alphanumerics. These micro-program sequences can, as user experience grows or the system is expanded, be extended. Both cartesian (x and y) and polar (r,φ) mode (for conics etc) for vector plotting are available. This method of vector generation and fast ROM's (100ns cycle time) together with fast CRT displays using electro-static deflection, allows incremental point plotting rates at 10Mhz.

To overcome display refresh problems (problems in memory capacity and speed) and the cost of high speed precision CRT's, only several such electrostatically deflected CRT's are used; the displays on these are random-scan generated, and then scan-converted either via scan-conversion tubes or by imaging the random scan display onto a Vidicon camera (i.e. hardware "formating" and not software formating), to TV-rate raster-scanning. A similar principle (i.e. hardware implemented scan-conversion) was used in the "STARE" hard-copy output system at the Bell Telephone Labs (157). These resultant displays, being the outputs from the scan-conversion tubes, are stored on video-discs, which then are used as the refresh memories. This scan-conversion scheme allows the use of slightly modified, economic commercial TV-receivers as the displays. The user accesses the CPU via his console consisting of TV-display lightpen and keyboard. The instructions are writtin in PL-1, with special graphical orders inbedded. The PL-1 with the special graphical orders are executed in the CPU whereas the graphical orders are partly interpreted in INTERGRAPHIC and partly in the CPU.

Although several existing IGC systems use small, off-the-shelf computers to service IGCs, the above is seen to be an efficient approach to the problem, particularly as far as cost and updating speeds and number of terminal units capable of being serviced is concerned. Off-the-shelf control computers are usually small, general-purpose computers or else small, control computers; they do not make for efficient graphics control computers and in addition require added software backup. "INTERGRAPHIC" may be considered a case where software-hardware tradeoffs are optimized. Fig.54 shows the system organization.

2.8. GRAPHICS SOFTWARE

2.8.1 General

2.8.1.1 Software Organization (83,92,102,162)

Software backup for IGCs and for graphics processing occurs at several levels of complexity.

This is logical, as an IGC where graphic commands are interpreted for display etc, is a special purpose peripheral, while the central processor, where user-application programs are processed, is a general purpose device; thus there is a program for the IGC, there is a program for the CPU, and there is a program linking the two. Analogously there are 3 levels of complexity in the languages in which the programs are written. These are: -

(i) The lowest level language, usually written in the assembly language of the controlling computer (either the CPU in the stand-alone configuration or the peripheral computer in other cases), interacts directly with the IGC hardware. It handles and controls:

 (a) CRT display file maintenance.
 (b) Hardware function generation (vector and alphanumeric) for display and input.
 (c) associated interrupt functions for display and input controlled by the interrupt analysis program.
 (d) transfer of data between the IGC sub-systems.

 Being in machine- and IGC-dependent assembly language, subroutines and programs are not interchangeable between machines; consequently little standardization exists, with each IGC manufacturer going his merry way.

(ii) The intermediate level language is the conversion or "interface" language between the hardware control language of the IGC and the application or user program in the CPU. It:

 (a) includes graphic routines which perform data translation between the complex

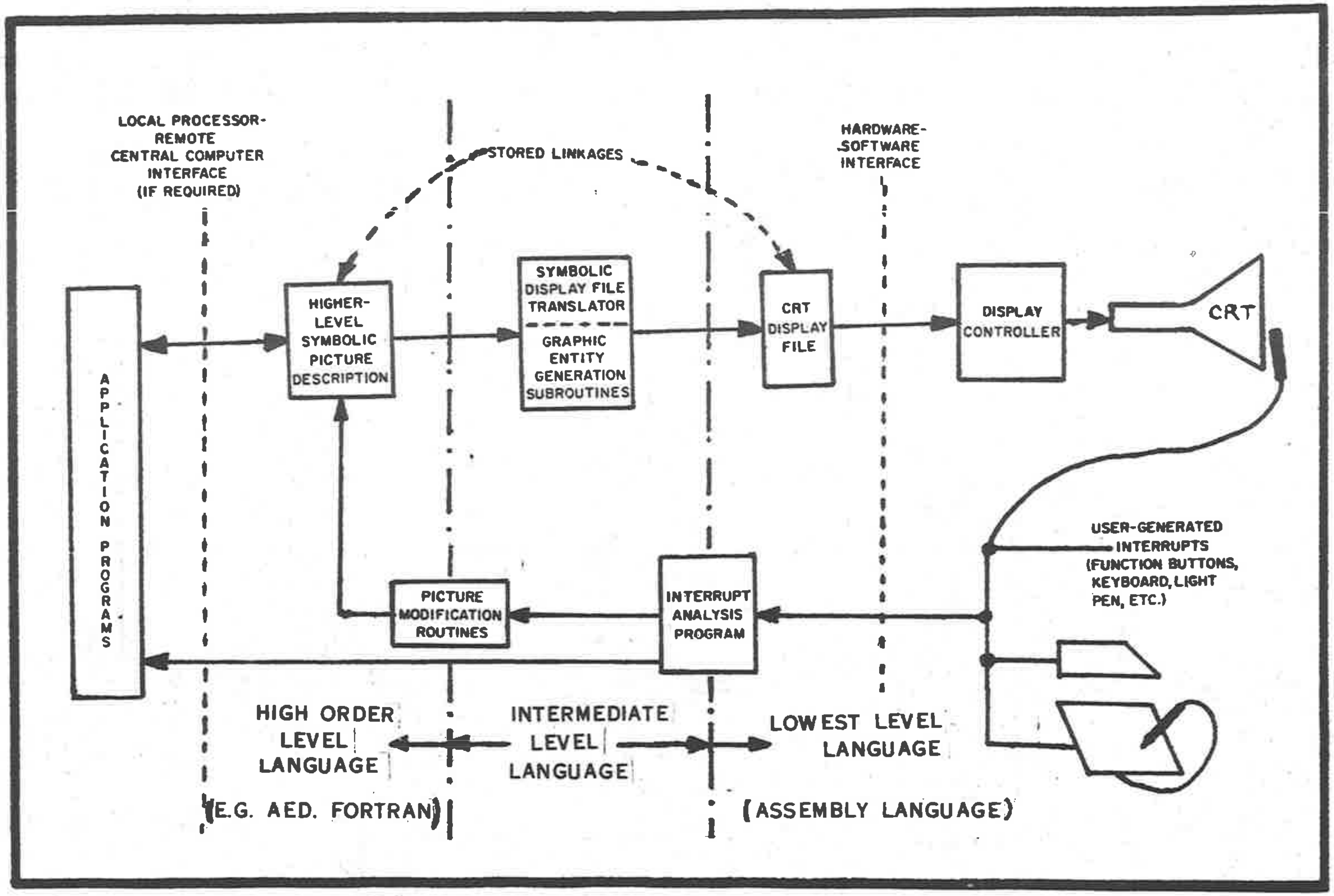


Figure 55. Typical Interactive Graphic Software and Language Level Function Organization. (based on [85])

structure (ring-structure) of graphic data in the applications program to the simple 12-20-bit word structure of graphic elements used within the IGC (see section 2.8.3.).

(b) Being an interface language it has embedded within itself graphic subroutines such as "rotate", "scale", "shift" etc. and various graphic generation subroutines which are on call by the application program and the user. In Fig.55, showing typical graphics software organization, this is called the "Graphic Entity Generation Subroutine" block. It thus has parts in the lower level assembly language and yet still needs be capable of handling graphic data expressed in "ring-structures", i.e. the high level symbolic graphic image representation. The block labelled "Display File Translator" is "where the action is" for the intermediate language level. For example, topological information giving circuit element interconnections will, for the display file, be interpreted and will result in the generation of vectors connecting the individual components or effectively result in the specification of the nodes. For the higher level language picture description, the Translator generated the complex interconnected ring structures giving the symbolic graphic circuit representation.

A start towards standardization of the intermediate language has been made with work done towards defining a "picture calculus" by which canonic and therefore standardized ways of expressing fundamental graphic processing procedures can be formulated (163).

(3) The higher level language is reserved for the user- or applications-program for specific problems. These currently are written in extended versions of high level languages such as FORTRAN, ALGOL, PL-1 (92).

It is clear that such a language must be capable of describing and operating efficiently on graphics data, while at the same time describe the application program (e.g. such as circuit analysis). AED-O mentioned before in section 1.2.3.3. is an extended version of ALGOL. FORTRAN on the other hand, although used "extensively" (relatively, as very few true, interactive graphics systems are available) due to its availability, is very inefficient for MCG; it has no inbuilt ability to handle real-time interrupts (on which after all, MCG is based) nor has it available graphic Input-Output commands. Whenever it is used, it has on call a set of graphic subroutines embedded within the lower level interface language (172).

The above 3 level structure of language has been implemented for example in GINA (section 1.3.4.2)

Until more fundamental work is done on software graphic data representation, and manipulation, it seems that this 3-level language structure will remain the philosophy for the time being.

2.8.1.2 Graphic Data Structure

Experience has shown that to represent graphics information in such a way that explicit topological information and "graphic" features are still retained and easily accessed by user actions such as "pointing" or "picking", it is best represented as interconnected variable-length lists (ring-structures such as the "plex-structures" described in section 1.2.3.3. and shown in figure 3(b)); these are alternatively described as "cross-indexed" or "tagged" block structures (49,120). Languages for graphics need have efficient ability to manipulate these interconnected blocks of data and their descriptors. The available list processing languages have formed a basis for the higher level (macro-language") graphics processing languages. A short review of graphic data structures and the macro-language needed to manipulate them is found in reference (101).

2.8.2. General Purpose Graphics Language

2.8.2.1 General (164,165,166)

What is ideally required is a truly general purpose graphics language, rather than the ad hoc approach taken and implemented thus far. An attempt at such a general purpose language has been made (164). Although only an approximation to a complete language (whose requirements are enumerated below), in the sense that not all the requirements have been met, it is "general", because, even though initially developed for an IBM 360, it has been run on other computers. The system software (compilers, operational executive etc) have been coded in IBM 360 assembly language.

2.8.2.2. Requirements of a General Purpose Language (164)

A general purpose graphics language needs to

(i) describe graphics features

(ii) mainpulate graphics features

(iii) analyse and interpret graphics features

(iv) generate and display graphics features

(v) contain features desirable in a general purpose language.

The data structure should not directly affect the structure and design of the language. One of the variations of the "ring" or multi-indexed list structures can be used.

To <u>describe</u> graphics features, commands must be provided for placing points, vectors (linear, segments of conic sections, or arbitrary curves) and alphanumerics within the graphic image area.

<u>Manipulation</u> of graphics features includes rotation, translation, scaling, rearranging, and deletion of subelements; storing of images for future reference also comes under this category.

<u>Analysis</u> requires that commands be provided to examine and locate special features within the image (e.g. "angle", "square" etc) and relate a subelement with respect to another.

If graphics data is to serve as input to some other CPU (e.g. to a peripheral computer or to a remote IGC), then some form of analysis and check should be made to validate the graphic image prior to exectution.

In addition, the <u>language should be interactive</u> for on-line interactive applications and to indicate any errors of a trivial nature immediately. Procedure and sub-routine capabilities are necessary, not only to augment the capabilities of the system as time goes on, but also for user-defined on-line subroutines.

Finally the <u>language should be "incremental"</u> i.e.

any additions or deletions to a program should be accepted without recompilation; compilation is to occur on a line-by-line basis. With error diagnostics and self-explanatory error comments, trivial errors can be eliminated as they occur.

To achieve the on-line requirements for the language a compiler-compiler ("meta-compiler") is necessary. Two passes are required in this system, but with incremental compilation by the meta-compiler and the output compiler, the two passes can be made serially, line-by-line which can thus be made to appear as a "single" pass, but with the two component passes "staggered" or delayed one program statement line. The meta-compiler accepts the graphics language input statements (which combine syntax and semantics), on a line-by-line basis and outputs code which constitutes the parts of the graphic compiler, not previously included.

The complete system with the emphasis on the User Input and Output System is shown in Fig.56. Referring to the figure, the box labelled "Graphic Compiler" is the output compiler from the meta-compiler (i.e. compiler after one pass).

The library of "Graphics Subroutines and Applications Programs" can be in any language but in an already compiled form compatible with the graphics output language.

The box labelled "Operational Executive" controls compilation, computation, storage and retrieval, and input-output interrupts and determines when execution

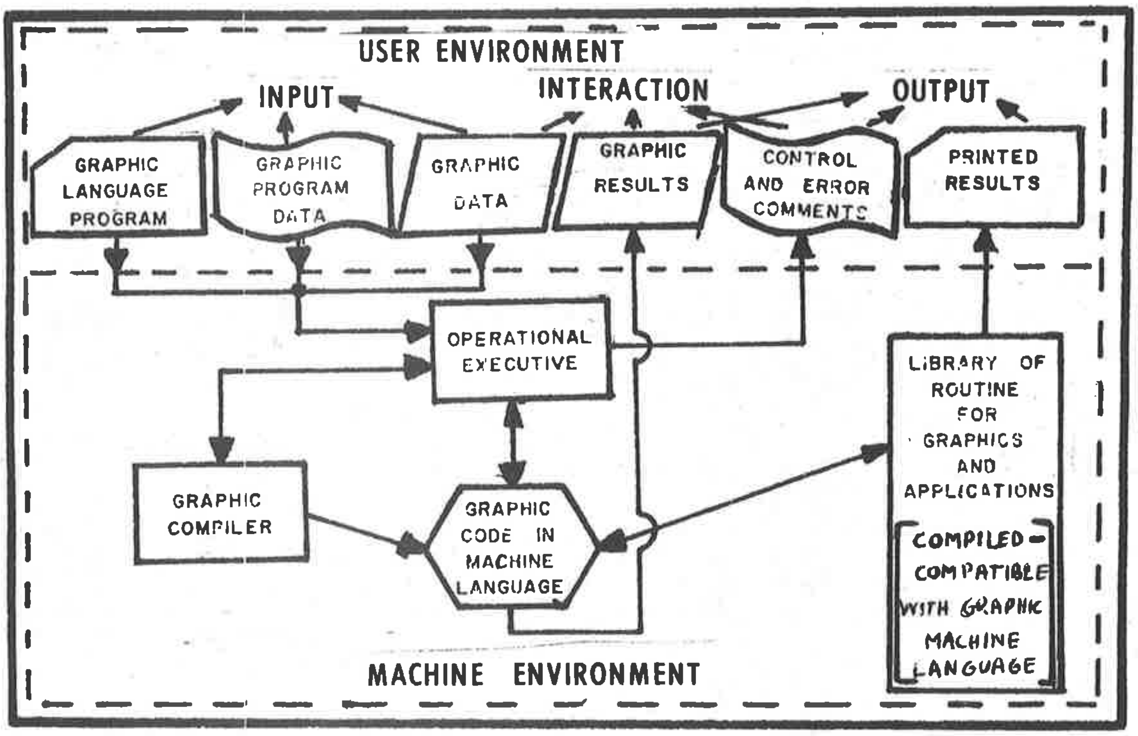


Figure 56. General Purpose Graphics Language System Organization (derived from [164])

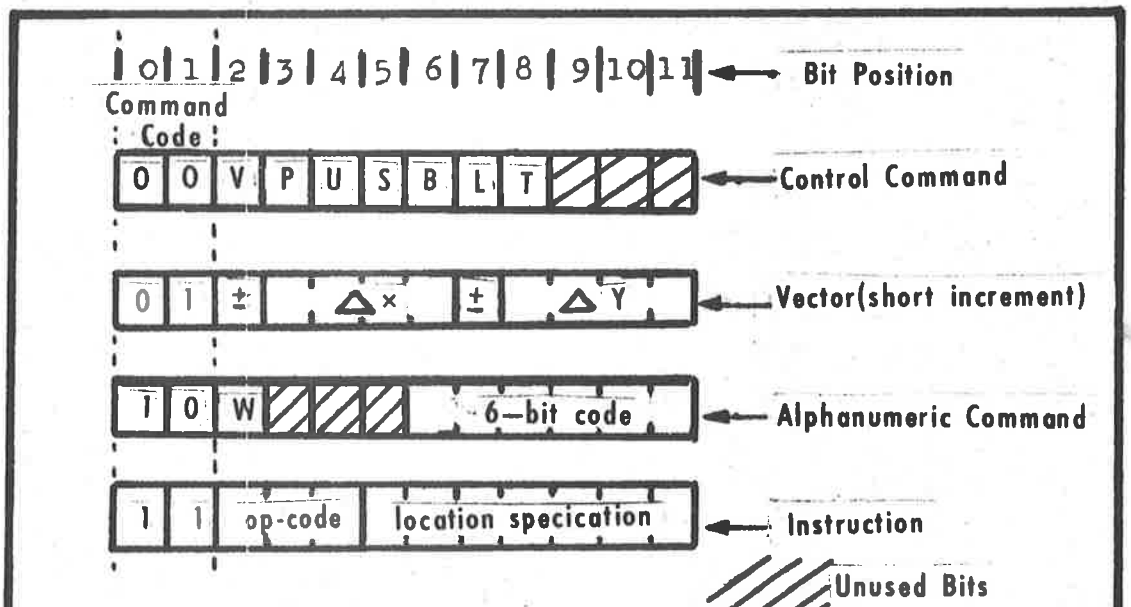


Figure 57. Typical IGC Display Processor Commands for Display Operations

is possible and when error statements are generated; debugging is on the graphics language level. The other boxes are self explanatory. Reference (164) also contains examples of command statements necessary in such a general purpose language, specifying descriptions, manipulations, program control and image analysis.

As stated previously the language described above is a suggested approach; further development have not been forthcoming. The current working systems are still based on the multi-level language approach described in section 2.8.1.

2.8.3 Display Interface Software

2.8.3.1 General (74,76,88,99,165,167,168)

In order that an image be generated on the display-device screen, the display software generates a sequence of binary words i.e. a series of display commands, which correspond to the displayed image; this is called the "display file". This sequence of binary words is generated by the CPU or the peripheral computer and is stored in the display refresh memory; each refresh cycle, the words are injected into, and processed by, the "display processor" which contains logic and control which decode the display commands and generate the direct display control signals to activate the appropriate display hardware circuitry which generate the displayed image (Fig.51). The form and structure of the display commands is usally dictated by the display processor hardware and other display hardware; hardware having being on the scene for longer than software, this is only to be expected. How these display commands are initially generated within the CPU or decoded by the logic hardware within the display process or is outside the scope of this review.

The former is within the province of the user applications programs and systems software, while the latter is similar to the decoding within the control unit in a small computer. Some mention of this was made in section 2.8.1.1.

Of interest here are the type and number of display commands presented to, and processed by, the display processor and the format that such display commands have.

There are 3 basic display commands:

(i) control commands

(ii) instruction commands

(iii) vector generation commands

which may be further subdivided into:

(a) true vector generation commands

(b) alphanumeric or text commands.

Hypothetical examples using command words of 12-bit length will be given. There being four general commands, 2 bits (the first two) in each command word are used to specify the type of command leaving 10 bits for the specific command information.

2.8.3.2. Control Commands

This command word preceeds a sequence of vector commands and the parameters specified within it apply to all the successive vector commands until a new control command appears. The control commands specify whether alphanumerics or vectors follow it ("V" requiring 1 bit). (The letters in brackets refer to these shown in the control command words in Fig57(a)).

If alphanumerics are specified then:

(i) various character sizes may be specified.

(ii) whether they are to be subscripts, superscripts or normal (2 bits)

(iii) brightness level (say 1 bit - 2 levels).

If vectors are specified, then specifications include

(i) point plotting or line plotting ("P", 1 bit)

(ii) line-plotting options, either solid or dashed vectors ("U", 1 bit).

(iii) scaled or normal size ("S", 1 bit) - same bit specifications as for alphanumeric size.

(iv) brightness level ("B", 1 bit) may also be specified.

The several remaining bits may be used to enable pointing-pen interrupt and user actions to be implemented. In Fig.57(a), "L" and "T" stand for Light Pen and Graphic Tablet interrupts. Typical command formats of varying length words are found in references (76,88,99).

2.8.3.3. Vector Generation Commands

These specify vector generating data, which may be specified by absolute location (a sequence of locations equals a vector or else specifies the starting point of a vector) or by specifying "Δx" and "Δy" increments. Depending on the command word length (e.g. 24-bits or 12-bits say) either one or two words may be required to specify the x- and y-coordinates. In the incremental case (for long vectors) 9 bits are used for increment length (a maximum of 512 location spacings) the 10th bit giving the sign (i.e. up or down, or left or right). For short incremental vectors (such as are used to generate small detail) a single 12-bit word may be enough, giving vector lengths of up to 16 locations spacings for both the x and y direction; the other two bits specify signs for the Δx and Δy increments; this is shown in Fig.57(b). Polar coordinates can be analogously specified, both for short and long vector mode.

2.8.3.4 Alphanumeric or Text Commands

With a normal character code of 64 or 128 characters and symbols, 6-7 bits are required to activate the symbol generator. The remaining 3-4 bits can specify the lay-out of the text, such as inter-character spacing (the W bits), word spacing, "carriage return", a new entry within a table etc. A 12-bit alphanumeric command is shown in Fig.57(c).

2.8.3.5 Instruction Commands

The 10 remaining bits are used to specify certain instructions such as "transfer control", "execute", "delete", "halt"; 3 bits can be set aside for this operation code with the remaining 7 bits specifying the location within the command location register where the above instructions reside. Display subroutines are specified this way, e.g. editing or updating display files. This command is shown in Fig.57(d).

2.8.4. User- Interaction Software

2.8.4.1 General (24,51,120,157)

User-interaction during the MCG problem solving phase is through the various input devices associated with the IGC; which of these devices are used, when and how they are used depends on the particular program being reached. Since the user controls the interactions, he needs to know the capabilities of the program and the state of the solution reached. These two requirements are made available to him via the display device, both at the beginning of the problem run when the data and requests are fed in, and during the problem phase whenever user decisions are requirea.

The state of the problem may be indicated in the form of alphanumeric data and labels, plotted graphs etc. and the current "light button" options being displayed. The program capabilities and user requirements are indicated by the function buttons on the "overlays" section 2.4.4.1), the possible "light buttons" which can be displayed during the problem run, and the set of possible commands which can be constructed with the keyboard. It is by knowing these, and seeing them displayed, examining the choices, making the decision and initiating the action via the appropriate input device that the user interacts with the computer and, in conjunction with it, solves his problem.

It was mentioned in Chapter 1, that a user-designer is not usually an expert programmer; at the most he may be moderately able. Thus the user-interaction procedures, and the language in which they are expressed in, need be easily understood and be unambigous. As the user is likely to make mistakes, (particulary during the initial phases) "obvious" wrong commands or actions must be ignored by the display processor and the CPU; also errors should be indicated to the user by being displayed (e.g. the "error" light button should blink-see section 2.2.2.11).

The current philosophy in meeting these requirements is to "guide" ("conducted tour"?) the user through the program, offering him, step by step, a set of plausible moves he can make by displaying these moves on the display and telling him how to make his move; a set of such options is called a "menu" (51,120). The textual question-answer or "menu-select" part of the program is either written in FORTRAN or QUIKTRAN (169) which is a FORTRAN-like language, but with a set of statements to help debugging (statement by statement) and editing the input program sequence, and also a set of operating statements which provide the user with the functions available on the IGC. Reference (170) lists a set of circuit design programs which are conversational in structure; nearly all the ones listed are written in FORTRAN or QUIKTRAN.

A particualr problem, when solved by a computer in conjunction with the user, has its solution specified by the sequence of moves, which is the "solution trace" of the user-selected and computer-selected subroutines and data inputs, through the complete tree of all possible plausible subroutine and data combinations. In the usual off-line batch processing case, the a priori (prior to computer input) defined program, would have required many "IF" and "TEST" statements to cover "all" of the plausible cases to be inserted within the program at locations based on the user's knowledge of the problem and of programming techniques.

2.8.4.2 Interactive Languages

The sequence of choosing and picking items off "menus" and initiating other console functions is a form of language - the "pointing language" (120). Analogous to statements within a normally written program for batch processing, these interactive program "statements" are constructed by selecting the proper sequences of items from the menu, typing — in alphanumeric data, or activating moves on other user input devices.

The language in which the subroutines, activated by keyboard use or function buttons or menu items, are written-in are high-level languages such as PL-1, ALGOL etc. Graphics information written in this language is ring-structured as objects or features pointed to should be recognized at such, without the intervention of the CPU. As mentioned in section 2.3.1.2, ring- or plex-structures are a cross-indexed listing of the graphics features; pointing to any such feature on the display effectively means selecting that feature from the ring-structure representation.

The language provides statements which effectively ask questions of the user or indicates the options on the "menu". Control is then held by that statement until the user makes a response either via his light pen, cursor and joystick, or keyboard; control is passed onto the statement or subroutine specified by the user. This is performed by the block labelled "Picture Modification Routine" in Fig.55. For example, one such statement may be a "WHICH" statement (157). The user is effectively asked "which of the objects displayed on the display screen?" Whatever happens when the user makes his selection depends on which subroutine happens to be in control; thus if a "DELETE" subroutine is in control, the selected item is deleted. Other state-

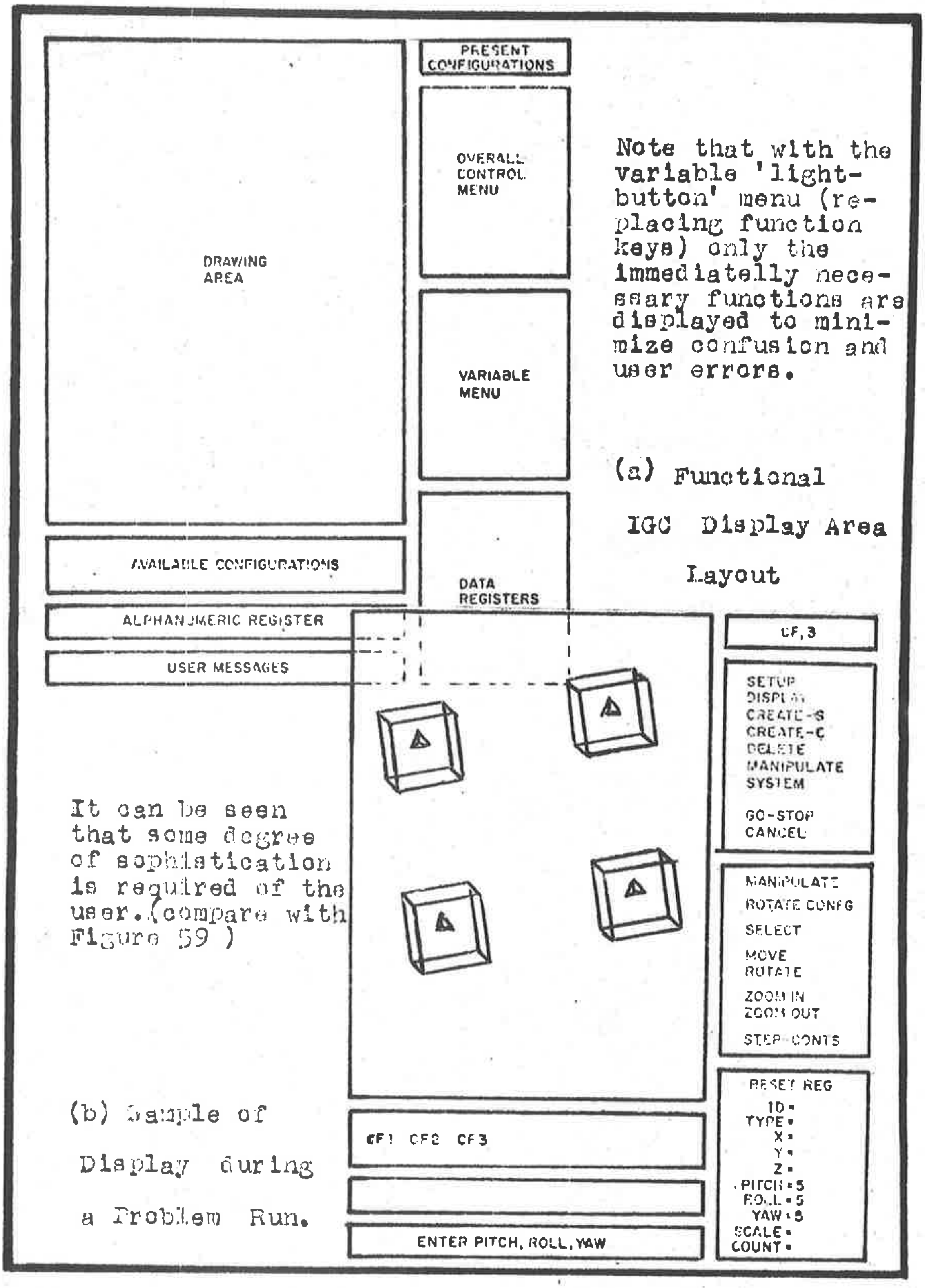


Figure 58. Sophisticated IGC Display Area Layout

(from [120])

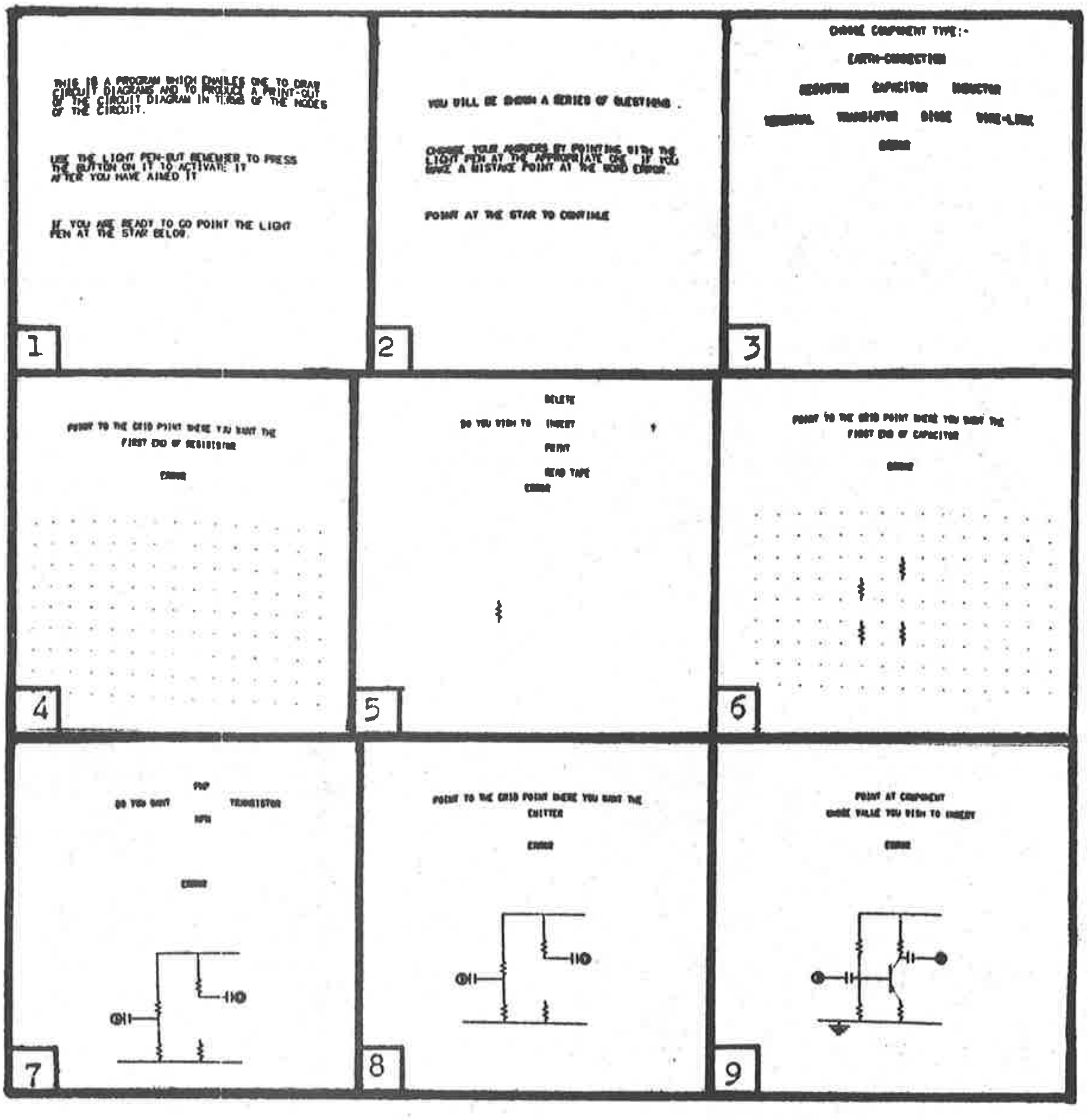
ments in the graphics environment may be the "WHERE" statement, the "GO TO", the "DETACH" and so on. Subroutines in control may be "SCALE", "ROTATE", "SHIFT" etc. Other commands must be provided for describing an object and to change the object.

A most useful feature is to allow the user to construct and specify his own subroutines or "macros" and put them on the "menu" (157). For example, such a subroutine may be a generation subroutine to construct a symbol for a transistor which is not part of the symbol repertoire. This symbol may be drawn large scale by a sequence of moves, then scaled down to the normal component size. A "Macro Definition" light button is provided for this purpose; when activated, the name "TRANSISTOR" is entered. The transistor symbol is then constructed, with the sequence of control commands being stored. The "Terminate Macro" light button is activated. The user-constructed "Transistor" generator subroutine now is part of the menu.

Display control commands such as changing the display brightness or controlling the display grid spacing need also be provided. System procedures, initiated by the user only, may store the present displayed information for future use ("inter-mediate results") or generate hard-copy of the display.

2.8.4.3 Existing Software Systems

Existing systems vary in complexity and organization depending on the application and the degree of user-interaction possible.



Sequence of frames (reverse polarity) of actual displays showing a 'guided' program enabling an inexperienced IGC user to input an arbitrary circuit for later analysis.

Figure 59. Example of Highly Interactive Graphics Program.

(from [51])

Where relatively "straight-forward" interaction occurs, such as the graphical inputting and specification of an electric circuit, an interactive program may be implemented such that a user with no previous programming experience may use it (51). The user is not only instructed via the display in the purpose of the program but also how to use it and how to use the input devices such as the pointing pen. The allowable locations where circuit elements may be positioned are even indicated. The "menu" is very limited to minimize user errors, with an error indicator being also provided. Such a "verbose" or "chatty" program is possible because the very nature of inputting circuit elements (of a limited repertoire) on a specified position grid is fairly predictable. A sequence of display frames taken from an actual example from reference (51) is shown in Fig.59. The GINA system described in section 1.3.4.2 is very similar but not so "verbose" as this; the menu items are in graphic form and are shown in Fig.9(a).

On the other hand, with larger CPUs, and users who are programmers with some experience, or rather have knowledge of the application program structure, menu lists and user-options may be provided at several levels; for example a relatively "unchanging" menu of available subroutines and a "volatile" menu which is displayed in conjunction with a particular subroutine may be simultaneously shown (120). Such adisplay is shown in Fig.58. The number of options being so large, user error probability is increased; hence programming knowledge is required.

The highest form of interactive language in MCG which has thus far been developed is one wherein graphical symbols specify computer procedures (29). This has been briefly illustrated in Fig.5; it is an offshoot of the SKETCHPAD set of programs. Simple computer procedures such as arithmetical or logical functions can be specified by drawing-in and interconnecting graphic symbolic blocks. This is analogous to drawing block- or flow-diagrams. The inputs to these blocks may be algebraic, numerical, or even analog functions such as sinusoids or noise (as in Fig.5). Thus programs, rather than being generated by sequences of the various selected menu items and keyboard actions, are generated by "stringing" together graphic function blocks, the permanent record of the sequence (and hence of the program) being displayed to the user as the completed flow or block diagram.

To describe more fully the software and user interaction languages is beyond this short review. More information concerning software and the mechanics of its implementation can be found in references (171,172,173).

2.9. OTHER GENERAL CONSIDERATIONS (151)

Certain desirable features need always be kept in mind:

(a) Economics both in initial outlay and in operating costs (mainly measured by the time the CPU is tied up during execution of graphics operations and commands) need be considered and compared with the cost of the task as performed without MCG. Hard ware-software trade-offs help to determine this. As regards CPU economy, it must be calculated on CPU cost and effectiveness of use when MCG is present. Studies on CPU economy (117) based on viewing a computer as a memory system connected to a series of processors have been made. Optimum use results when the time to access a given number of lists in a certain order is minimized. This depends (for a fixed memory organization) on the data structure, the applications program and systems programs. A similar analogy to graphics applications can also be made.

(b) Reliability and long service ensure user confidence (and management confidence). Display devices, particularly for precision graphics, cannot be replaced every few weeks; not only is tube replacement costly, but time and expense in measuring and correcting display linearity incurs added costs. The same goes for input devices. At least 1000-2000 hours of IGC use should be obtained without major replacements. Ageing, ambient effects etc, all affect performance, particularly repeatability. Air conditioning may thus be necessary.

(c) Physical size and good maintainability are desirable; modularity, for multi-station configurations is another thing to be considered.

2.10 GENERAL CONCLUSIONS

The enumeration of parameters relevent to graphic display and graphic input devices, and the overview of current devices and techniques which, when assembled, make up the Interactive Graphics Console, not only help to identify the necessary components of such an IGC, but indicate the complexity, and hence the resulting high cost, of commercial IGCs.

For user-comfort and acceptability of use, certain requirements are necessary, such as the quality of the displayed data (resolution, contrast, linearity etc) and the naturalness of use of the graphics input device, and its speed of response.

From the implementation viewpoint, it has been seen that the IGC must consist of

(1) a display device on which displayed features can be directly accessed by the user.

(2) a means of user input of alphanumerics and graphics data, and a means of accessing ("pointing to") displayed data.

(3) means of generating and maintaining, once generated, user-specified and computer-generated alphanumeric and grapics vector data.

Other "basic" or "minimum requirements" occasionally mentioned, such a display refresh stores peripheral control computers etc, are covered by, or are concerned with, implementing the above requirements.

User and application programs, together with the central computer processes and computations do not come under the domain of Man-Computer-Graphics hardware.

In Chapter 1. the power of MCG in solving problems which previously were solved inefficiently using batch-processing, or which could not have been solved by the then available means were clearly shown. The classes of problems and field of applications are increasing all the time. It is not surprising then that a real need exists for a low-cost IGC. A successful solution to this problem is a necessity.

....

PART 4
APPENDICES

1 APPLICATIONS OF MAN—COMPUTER GRAPHICS

2 AVAILABLE CIRCUIT ANALYSIS AND M-C-G-CIRCUIT DESIGN PROGRAMS

3 COMPARISON OF MAJOR CLASSES OF DISPLAYS

4 COMMERCIALLY AVAILABLE "LOW—COST" CRT GRAPHICS TERMINALS AND I.G.Cs

5 GRAPHICS INPUT AND POINTING DEVICES

6 CRT AND VIDICON BEAM DEFLECTION

7 BEAM DISCHARGE IN VIDICONS

8 PHOTOMETRY AND OPTICS

9 PHOSPHORS AND CATHODOLUMINESCENCE

10 PHOTOCONDUCTIVITY

11 MEASUREMENT OF DISPLAY DISTORTION AND VIDICON SCANNED AREA DISTORTION : MOIRE FRINGES

12 TYPICAL TV—CRT AND VIDICON CAMERA SYSTEMS

13 REPRINTS OF PAPERS AND MISCELLANEOUS

REFERENCES

APPENDIX 1

APPLICATIONS OF MAN—COMPUTER GRAPHICS

APPENDIX 1 - TABLE

APPLICATIONS OF MAN—COMPUTER GRAPHICS

FIELD of INTEREST	PRESENTED DISPLAY	USER INTERACTION	RESULTANT DISPLAY	APPLICATION	EXAMPLES WITH LISTED REFERENCES
REAL-TIME APPLICATIONS	Current traffic indications (road, rail port, air), civil and military; Current status of controlled parameters.	Point; evaluate current situations; modify by pointing and pressing "function" buttons.	Controlled traffic situation. Controlled parameters updated.	Traffic control. Tactical evaluation. Petrochemical industry. Process control. Weather displays. Reactor control.	(2),(30), (121),(205) (207).
ALPHANUMERIC PROCESSING	Unedited drafts; files; typewriter inputs; computer cues and options.	Pointing, deletion, typing.	Corrected or updated files. Intermediate alphanumeric results.	File retrieval and file updating; ticket reservations ; text editing; program editing and on-line debugging; computer-aided-instruction; on-line phototypesetting.	(2),(8), (159),(207), (216),(227).
MATHEMATICAL ANALYSIS	Current status of iterative graphical solution, "i-th" iteration; graphs.	Evaluate solution trend; modify coefficiants if unsatisfactory; increase or decrease iteration step size; evaluate occurence of local maxima, minima etc.	"i+1-th" iteration.	On-line iterative solutions of ill-defined solutions; management , trend evaluation.	(2),(8),(31) (32),(205).
GRAPHIC ANALYSIS (i)	Point plots; curve representing beam or linkage deflections.	Specify trial curves; keyboard action; modify initial curves.	Optimum fitted analytic expressions; resultant curves of shear, bending stresses, velocity, acceleration curves.	Curve fitting; graphical differentiation and integration.	(2),(4),(8) (20),(26), (62).
GRAPHIC ANALYSIS (ii)	One and two dimensional patterns to be recognized.	—	Selected tracks or recognized and emphasized patterns; enhanced picture.	Picture processing; pattern recognition; bubble chamber track analysis.	(47),(238) (239).
TOPOLOGICAL ANALYSIS (QUALITATIVE)	Elements to be connected.	Location of elements or flow indication; dimension of elements.	Finished layouts of pipes, circuits; I-C; printed circuits wiring diagram etc.	Circuit design; refinery pipe line layout.	(2),(4), (211),(224), (233),(234).

APPENDIX 1 - TABLE (cont.)

APPLICATIONS OF MAN—COMPUTER GRAPHICS

FIELD of INTEREST		PRESENTED DISPLAY	USER INTERACTION	RESULTANT DISPLAY	APPLICATION	EXAMPLES WITH LISTED REFERENCES
TOPOLOGICAL ANALYSIS (QUANTITATIVE)		Elements to be laid out; size of layout area.	Indicate desired location of certain elements; specify element size.	Optimum layout according to certain criteria.	I.C.Layouts; Plant, office layout; graphic layout (newspaper); garment cutting; molecular studies.	(226),(230), (237),(238).
GEOMETRY in 2-Ds		Precise engineering drawings in simple views.	Rotate to obtain complex views; scale; modify; generate control tapes for N.C. machines, on-line control of machining.	Modified views; progress of N.C machining.	Drafting and design in engineering industries; N.C. machining.	(2),(4),(8), (11),(231), (232).
GEOMETRY in 3-Ds	(i)	3-D projected views.	Rotate, scale, evaluate.	Projections from different angles.	Hidden line problems; architecture, pilot visibility views; driver visibility of highways.	(19),(30), (208),(210).
	(ii)	surfaces generated from 3 coordinates.	Scale, modify, rotate.	Smoothed" surfaces.	Complex-shape, smooth body design(aerospace, ship, automotive industries).	(2),(4),(10), (11),(18), (19).
	(iii)	Generated coordinates representing surfaces which cannot be expressed analytically.	Specify surfaces; rotate; "smooth".	Desired elevation views and smoothed shapes.	3-D field plotting; geology and mining applications.	(2),(8), (210),(229).
GRAPHIC ANIMATION	(i)	2-D views of initial conditions.	Analysis initiation; modification of initial conditions.	Successive views of dynamic situations at specified intervals	Turbulence analysis of fluid flows; kinematics of mechanical linkages.	(2),(8),(30).
	(ii)	———	———	Animated block diagrams or of simple shapes in projected perspective views.	Animated movies; system studies.	(210),(228), (241).
	(iii)	Background scene and initial conditions "i-th" scene.	Erase and modify; store (e.g. arm motion of cartoon character).	"i+1-th" scene.	Animated cartoons.	(210),(228), (241).
GRAPHIC ARTS		———	Modify - "do one's own thing".	Abstract and "Op-art" (?).	Textile design. Entering "Computer Art" Contests; expensive hobby!	(242).

APPENDIX 2

1 AVAILABLE CIRCUIT ANALYSIS COMPUTER PROGRAMS

2 COMPARISON OF SOME AVAILABLE CIRCUIT ANALYSIS PROGRAMS

3 SOME PRESENT M-C-G-CIRCUIT- DESIGN PROGRAMS

APPENDIX 2

AVAILABLE COMPUTER-AIDED-CIRCUIT ANALYIS PROGRAMS

There is a vast amount of literature describing successful efforts (181, 182, 203, 204) of what are called "computer-aided circuit designs"; in fact these are solved by the usual off-line, and often, close-shop methods. The circuits are pre-determined, and data cards etc., are pre-punched, describing circuit interconnections and component values. They are in fact circuit analysis programs for pre-determined circuit configurations; if alterations need be made, new cards need be punched. They are "computer-aided-circuit-analyses' and not "design"! (Admittedly "analysis" is part of the design process. See figure 8.)

Such programs are immensely useful - indeed they are utilized in part or in whole in various M-C-G circuit design systems. The size of such programs is what prohibits them running on nothing except the largest computers, and results in but the very simplest circuits being solved on-line by M-C-G. (A table of the most common programs with some of their features follows.) For example,ECAP hogs up the whole computer (IBM 7094) at instant of execution; but it can handle circuits containing up to 50-odd nodes and about 200 branches. Clearly a trade-off between graphics software and size of circuit capable of being solved is necessary in M-C-G, resulting in the relatively small and simple circuits being solved using this method.

The reason for such large analysis programs is that classic circuit analysis techniques have been used; these are not necessarily the best or most economic methods suitable for computer manipulation. It has been recently stated that "the present state of affairs in (C-A-C-D) has been characterized as 'chaotic' but perhaps 'intense' would be a better word"(183).

It is commonly agreed (183, 184) that three major problem areas exist, hampering the development of the field.

(a) present day network analysis methods such as branch or node methods are not necessarily suitable for computer processing. Intergrated circuits or large circuits may require more than 1,000 equations for their description; matrix inversions poses problems. In fact, "sparse" matrices are what occur and these techniques need be more fully explored. Classifying a network into fixed/variable parameters, linear/non-linear, memory/memory-less elements etc, and "tearing" the resulting matrix (via Kron's method) into specialized smaller sub-networks and obtaining the separate solutions, need be investigated. The "state-space" approach (185) offers some interesting advantages, namely that the "space state" formulation results in the minumum number of equations necessary to describe a network. Symbolic methods and their manipulations (187, 194) offer advantages as far as sensitivity analysis and round-off errors are concerned; numbers are only inserted in the final desired expressions, which were arrived at algebratically.

(b) On-line graphics implementation still leave much to be desired, mainly due to the economics involved. The complaint has been made (184) "that many at present (Nov. 1967) are operating at a loss with large "down" periods and long interaction periods". Efficient dynamic memory allocation programs coping with the hundreds of small and large matrices in various levels of storage need be developed. Compatability between programs is required - most working systems, to be efficient, have parts of the programs in assembly or machine language and thus preclude their implementation on some other computer. No graphic language with adequate generality as FORTRAN, say, exists as yet, but AED (22) and PL/1 (188) seem to be in the right direction.

(c) Applied mathematical programming to generate or select an "optimum" design, where "optimum" is defined by the designer, need be more fully developed. "Optimum" usually refers to some output characteristic to be constant under expected variable conditions such as the expected spread of component values, voltage rail variations etc. Circuit parameters should be adjusted by the computer until the "optimum" is met. With most present-day algorithms, some local optimum condition is reached, only to be relaxed again when some other parameter is changed dramatically. Logic-circuit optimization techniques are quite successful butlinear circuits still pose a problem.

Work is being actively pursued to overcome the above mentioned problems. The so called "second generation" circuit analysis programs (182) seem to indicate that these new programs, rather than be truly general purpose and capable of simultaneously handling AC, DC frequency, sensitivity, least power and other analyses, concentrate on one such analysis (although they can still handle to a lesser degree other analyses). A designer then would have on call a family of circuit analysis programs, each performing the particular analysis required.

. . . .

APPENDIX 2 - TABLE 1

COMPARISON OF SOME AVAILABLE CIRCUIT ANALYSIS PROGRAMS

PROGRAM and ORIGINATOR	CAPABILITY and METHOD of SOLUTION	MAX. CIRCUIT SIZE	MODEL of ACTIVE DEVICES	ON-LINE CAPABILITY	LANGUAGE	SIZE of MEMORY and COMPUTER	COMMENTS	REFERENCE
NET-1, NET-2 LOS ALAMOS SCIENTIFIC LABORITORIES	D.C, A.C (Net II), Transient, Frequency. Numerical; matrix inversion.	100 nodes, 300 branches.	Ebers-Moll (stored or user-specified).	None	FAP	32K IBM 7090-94		(182) (189) (190)
ECAP Electronic Circuit Analysis Program IBM	D.C., A.C, Transient in time domain. Numerical; matrix inversion	50 nodes, 200 branches	Piece-wise linear; user-specified.	Possible	FORTRAN II, IV	32K IBM 7090-94	Non-linear subroutines present.	(182) (190) (191) (192) (193)
NASAP Network Analysis for Systems Application Program N.A.S.A	D.C. , A.C Transient, Frequency Sensitivity. Symbolic, flow-graph methods		Piece-wise linear; user-specified.	Yes	FORTRAN IV	48K 12 tapes IBM 7094	(i) "Symbolic Analysis", excellent for sensitivity (Bode Sensitivity) (ii) User must code circuit into current and Voltage elements.	(182) (186) (187) (194)
SCEPTRE System for Circuit Evaluation Prediction of Transient Radiation Effects. IBM	D.C. , Transient. State-variable methods.	300 nodes, 300 branches.	Ebers-Moll (stored or user-specified.	Possible	FORTRAN IV	32K 12 tapes IBM 7094	User can specify subcircuits of up to 25 terminals (linear and non-line). User can specify own subroutines.	(182) (195) (196)

APPENDIX 2 - TABLE 1 (cont.)

COMPARISON OF SOME AVAILABLE CIRCUIT ANALYSIS PROGRAMS

PROGRAM and ORIGINATOR	CAPABILITY and METHOD of SOLUTION	MAX. CIRCUIT SIZE	MODEL of ACTIVE DEVICES	ON-LINE CAPABILITY	LANGUAGE	SIZE of MEMORY and COMPUTER	COMMENTS	REFERENCE
CALAHAN GENERAL LINEAR NETWORK ANALYSIS PROGRAM U. of Michigan.	D. , A.C, Transient Frequency Domain Topological K-tree method Inverse Laplace.	15-20 nodes max.	Controlled current sources and piece-wise models(?)	yes	FORTRAN II	32K IBM 7094 CDC 3600	Very limited capability in describing branch elements.	(182) (197)
POTTLE Cornell U.	Poles, Zeros & Frequency response State variable Approach		User specified models.	possible	FORTRAN IV			(185) (198)
CIRCUS Boeing	D.C Transient State variable Approach	100 nodes 300 branches.	Beaufoy-Sparkes (for transistor) Zener, FET user-specified.	possible	FORTRAN IV	32K IBM 7094 CDC 3600 Univac 1108	- Non-linear devices accepted. - Graphic Output	(182) (199)
CIRC S.D.S	D.C (A.C. & Transient later).	>50 nodes	Ebers-Moll	yes	FORTRAN II	16K (D.C analysis) SDS 900 series.	Interactive programming	(200)
REDAC Racal (U.K)	D.C, A.C, Transient, Frequency Sensitivity Matrix inversion, symbolic approach.	"Very large" (at least as above)	Eber-Moll (stored) or user-specified.	yes	C.C.L (Fortran based)	Elliot 4130.	- Used as a design service for industry. - Versatile non-linear capability.	(45) (201) (202)

APPENDIX 2 - TABLE 2

SOME PRESENT M-C-G-CIRCUIT- DESIGN PROGRAMS

PROGRAM and ORIGINATOR	CAPABILITY and METHOD of SOLUTION	MAX. CIRCUIT SIZE	MODEL of ACTIVE DEVICES	INPUT	LANGUAGE	SIZE of MEMORY and COMPUTER	COMMENTS	REFER -ENCE
OLCA ON LINE CIRCUIT ANALYSIS BELL TELEPHONE LABS 1966-7	FREQUENCY ANALYSIS (Gain vs frequency) (Phase vs frequency)	16 nodes	EQUIVALENT CIRCUIT + INDEPENDENT SOURCES	- LIGHT PEN - LIGHT BUTTONS. - PROGRAM CUES & OPTIONS	"GRIN" (GRAPHIC I-O LANGUAGE) FAP ? "FORTRAN" (CIRCUIT ANALYSIS)	32K, 8K (PDP5) IBM 7094 + GRAPHIC CONSOLE + PDP5	Experimental program; R,L,C,M, Voltage current source, phase shift, available.	(42)
AEDNET-1 PROJECT MAC 1966	TIME DOMAIN ANALYSIS (one Variable vs. another)	50 elements. (?)	PIECE-WISE LINEAR EQUIVALENT CIRCUIT + INDEPENDENT & DEPENDENT SOURCES	- TYPEWRITER - LIGHT PEN	"AED - 0"	1 + 5 CORES (1 Core = 32K) IBM 7094 + PROJECT MAC CONSOLE	Specifically a simulator for non-linear, time varying networks (can be specified numerically or graphically).	(48) (49)
GINA (GRAPHICAL INPUT for NETWORK ANALYSIS BELL TELEPHONE LABS 1968	FREQUENCY ANALYSIS (GAIN vs FREQUENCY) (CALAHAN PROGRAM)	20 nodes. (?)	PIECE-WISE LINEAR (as above)	- LIGHT PEN - LIGHT BUTTONS. - PROGRAM CUES & OPTIONS	"CALLIGRAPH" (GRAPHIC I-O LANGUAGE)	20 K words IBM 7094 CDC 280	Experimental program limited to R,L,C current source,	(47)
CIRCAL-1 CIRCUIT ANALYSIS PROJECT MAC 1966	TIME DOMAIN ANALYSIS. Matrix inversion for linear network iteration for non linear).	20 nodes, 50 branches	PIECE-WISE LINEAR (Can generate from table or by subroutines)	- TYPEWRITER - LIGHT PEN	"AED"	40,000 words (exceeds capacity) IBM 7094 + PROJECT MAC CONSOLE	Handles non-linear elements; can use subcircuits as elements.	(40)

NOTE - ALL PROGRAMS RECOMPUTE COMPLETE CIRCUIT WHEN ANY COMPONENT IS DELETED.

APPENDIX 3

COMPARISON OF MAJOR CLASSES OF DISPLAYS

APPENDIX 3 - TABLE

COMPARISON OF MAJOR CLASSES OF DISPLAYS

TYPE	COST	CLASS	RESOLUTION SPOT SIZE	LINEARITY	SPEED	MAX. BRIGHTNESS COLOUR	SIZE	LIFE	VOLTAGE	ADVANTAGES	DISADVANTAGES
CATHODOLUMINESCENCE CRTs (DYNAMIC)	LOW < $1000	ANALOG	0·002" ↓ 0·025" DIAMETER	1-3% TYPICAL 0·01% AT BEST.	2-3 µs FAST FOR RANDOM POSITION-ING.	(VARIA-BLE). 1000ft-L -Depends on phos-phor type green or white).	Up to 20"x20" DEPTH 15"	3000 hours	10-15KV	- Can display moving display -Scan mode flexibility. -high informa-tion content. - Intensit hence depth. -Colour -Low cost & fast	- Large size - catastro-phic failure mode (fila-ment fail-ure mode).
DIRECT VIEW STORAGE TUBE DVST CRT (STATIC)	MEDIUM ≈$ 2500	ANALOG	0·008" DIAMETER	2% TYPICAL 0·01% POSSIBLE	1000"/SEC ASYNCHRO-NOUS WRITING MODE.	6ft-L -Green	6"x9" DEPTH 15"	2000 hours	15KV	- Storage inherent. - relatively fast. - moderate cost.	-Incapable of showing moving displays. -Catastrophic failure. -Updating sel-ective erasure requires whole display era-sure.
ELECTROLUMINESCENCE E-L DISPLAYS	MEDIUM HIGH (EXPE-RIMEN-TAL).	DIGITAL	0·040" X 0·040"	BETTER THAN 0·01%	SLOW SWITCHING "ON" AND "OFF" ≈ Millisec-onds.	3-30ft-L -Depends on Phosphor	10"x10" DEPTH 1/8"	1500 - 5000 hours	15KV	- Small size - no catas-trophic failure. - low power	- Limited scanning flexibility - slow erasure slow switch-ing speeds.
GAS LUMINESCENCE NEON PLASMA DISPLAY	MEDIUM HIGH (EXPE-RIMEN-TAL) ≈$ 20K	DIGITAL	0·015" X 0·015"	BETTER THAN 0·01%	MEDIUM SLOW "ON"≈ 20us "OFF"≈ Milliseconds	500-1000 ft-L -Neon orange red.	10"x10"x DEPTH 1/2"	20000 - 50000 hours.	200 - 300 VDC.	- Small size. - No catastro-phic failure. - low power - high bright-ness & simple light pen for direct writing in & erasure.	-Limited Scanning flexability. - Slow erasure - Uneven manu-facturing tolerance.
INJECTION LUMINESCENCE L.E.D. DISPLAY	VERY HIGH (>$10[6])	DIGITAL	0·02" X 0·02" typical	BETTER THAN 0·01%	VERY FAST Several nanosec-onds.	Variable 1300ft-L -Red -Yellow -Green	20"x20" DEPTH 1"	100,000 hours	1.6V	- Fast - High bright-ness. - No catastro-phic failure. - Very low power.	- Prohibitive cost. (Only small matrix arrays for alphanumerics have been implemented

APPENDIX 4

1 REPRESENTATIVE SELECTION OF "LOW—COST" CRT GRAPHICS TERMINALS

2 REPRESENTATIVE SELECTION OF COMMERCIALLY AVAILABLE I.G.Cs

APPENDIX 4 - TABLE 1

REPRESENTATIVE SELECTION OF "LOW—COST" CRT GRAPHICS TERMINALS (mainly from (59))

		TEKTRONIX T-4002	BUNKER—RAMO BR-700	CCI CC-30	SANDERS 720
DISPLAY CHARACTERISTICS	DISPLAY SIZE	6" x 11"	$7\frac{3}{4}$"x$5\frac{1}{2}$"	TV-Monitor -size Optional	$7\frac{1}{2}$" x $8\frac{1}{2}$"
	SPOT SIZE	0·012"DIAM	0·015"DIAM	0·020-0·025"DIAM	0·020"DIAM
	MAX. No. of CHARACTERS	3200	960	≈900	256, 512 or 1024
	GENERATION TECHNIQUE	POINT PLOT	5 x 7 DOT MATRIX	5 x 7 DOT MATRIX	STROKE
	DEFLECTION	RANDOM SCAN	ELECTRO-MAGNETIC	ELECTRO-MAGNETIC	ELECTRO-MAGNETIC
	CONTRAST	6:1	10:1 (?)	Adjustable	Not Specified.
	MEMORY	DVST	DRUM-380K CHARACTER CAPACITY	CORE-1024 CHARACTER CAPACITY	DELAY LINE -1024 CHARA-CTER CAPACITY
SYSTEM CHARACTERISTICS	FUNCTION KEYS	LIMITED	STANDARD	NONE	STANDARD
	KEYBOARD	STANDARD	STANDARD	OPTIONAL	STANDARD
	VECTOR GENERATION	STANDARD	NONE	NONE	NONE
	STAND—ALONE	YES	NO	YES	NO
	MULTI—STATION CONFIGURATION	YES	YES	NO	YES
	MAX. No. of TERMINALS	DEPENDS ON CPU	16	-	12
	COST of CONTROLLER + TERMINALS)	$10K	$96K + 16 TRMNLs + CONTROLLER	$7K	$45K 12 TERMINALS + CONTROLLER
	EDITING FACILITIES	OFF-LINE	MOST EDIT-ING FACIL-ITIES.	VERY LITTLE -"CURSOR SHOWN"	ALL EDIT-ING AVAIL-ABLE.
	GRAPHIC INPUT FACILITIES	JOYSTICK	NONE	LIGHT PEN	NO

APPENDIX 4 - TABLE 2

REPRESENTATIVE SELECTION OF COMMERCIALLY AVAILABLE I.G.Cs

		INCREASING COMPLEXITY ⟶			SPECIAL PURPOSE	
	TYPE	ARDS-100A (144)	IBM 2250 III (62)	IDIIOM (61)	LDS-1 (63)	SEACO Artwork Generator (67)
DISPLAY FACTORS	PRIMARY PURPOSE	DVST GENERAL PURPOSE LOW COST GRAPHICS TERMINAL	GENERAL PURPOSE GRAPHICS IGC.	GENERAL PURPOSE GRAPHICS IGC.	PRIMARILY FOR PERSPECTIVE, ROTATION & ZOOMING & GENERATION OF MOVING PERSPECTIVE OBJECTS	HIGH PRECISION ARTWORK GENERATOR
DISPLAY FACTORS	DISPLAY SIZE	6·4" x 8·2"	10" x 12"	12" x 16"	12" x 12"	9" x 9" and a Monitoring DVST
DISPLAY FACTORS	ADDRESSABILITY	1000 x 1400 addressable points	10 bits for each x and y	10 bits for each x and y	12 bits for each x and y	14 bits for each x and y
DISPLAY FACTORS	SPOT SIZE	0·008" Diameter	0·020"(?) Diameter	0·010"-0·015" Diam.	0·020" Diameter	0·002" Diameter
DISPLAY FACTORS	LINEARITY	± 2%	± 3%	± 3%	"High"-not specified	± 0·006%
DISPLAY FACTORS	DISPLAY GENERATION	- Vector - Characters (96)	- Vector - Characters (64) (two sizes)	-Vectors (4 types) -Circle -Position (for sub-elemnt) -Character (64) (4 sizes)	- Line - Character - Complex function Character	Overlapping point by point vector generators
DISPLAY FACTORS	VECTOR GENERATION	- Point Plot Mode - Incremental Generating Mode	- Point Plot Mode - Incremental Mode - Stroke Generator (±45°, 180±45° Slope)	- Continuous line generation, "constant-time" with intensity compensation	Analog Line Generation-Constant drawing Speed.	Repertoire of subroutines for various geometrical fig. construction
INPUT	GRAPHICS INPUT DEVICES	- SRI "Mouse" or "Joystick" - Keyboard	- Light Pen - Function Keys - Keyboard	- Light Pen - Function Keys - Keyboard	- Graphics Input Tablet - Function Keys	- Keyboard - Magnetic Tapes - Paper Punch + Options - Discs
SYSTEM FACTORS	CONTROLLER or CPU	Can be interfaced with most computers	IBM 2840 - with 32K Store	VARIAN 620 (4k-32K Store)	Own Stored Program Controller	H.P 2116B Minicomputer + 8K Store
SYSTEM FACTORS	STAND-ALONE or MULTI-STATION	Both	Stand Alone or up to 4 units with a IBM 2840 processor	STAND ALONE	STAND ALONE or Controller can handle 4 Displays.	STAND ALONE
SYSTEM FACTORS	COST	$8 500	HIGH (> $50)	HIGH	HIGH	HIGH
SYSTEM FACTORS	SPECIAL FEATURES	Inherent storage (about 15 minutes)	High Cost.	- Rotated Characters -100us Circle Gnrtn. -Intensity Chrtr. size Vector type under user control.	- Rotation - Translation - Scaling (up to 250x) - Replication (all hardware implemented	- Software implemented (high command language) - high precision hardware.

APPENDIX 5

GRAPHICS INPUT AND POINTING DEVICES

1 DESCRIPTION AND SURVEY

2 TABLE

APPENDIX 5

GRAPHICS INPUT AND POINTING DEVICES DESCRIPTION AND SURVEY

A.5.1 GENERAL

The principles on which these devices are based are briefly described; the table at the end of the Appendix tabulates their main characteristics and features.

Three main classifications are used:

(1) Graphic input pads or 'tablets' which may be used in conjunction with a display device but can be used alone without a display.

(2) Graphic pointing pen devices which need be used in conjunction with a display device.

(3) Electromechanical input and cursor positioning devices.

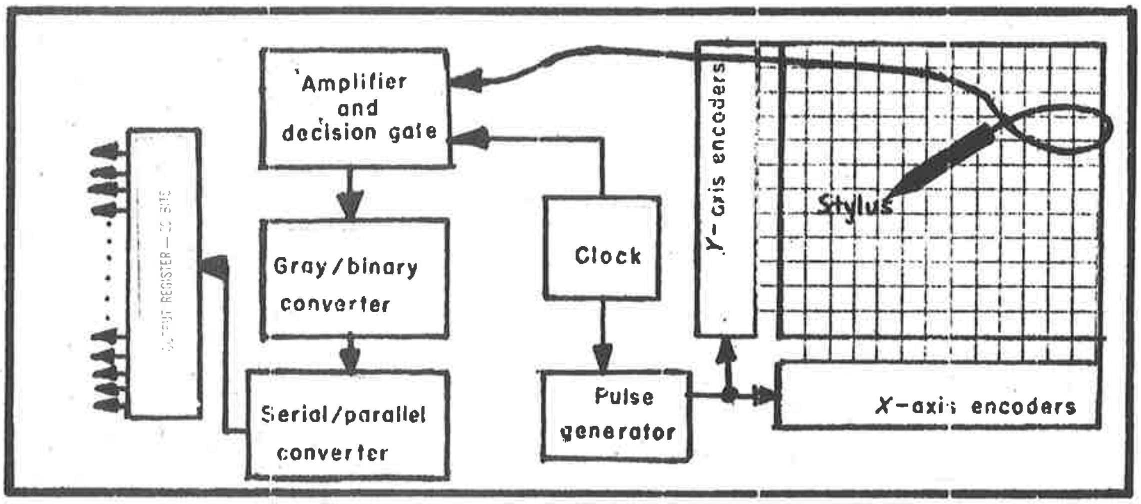
A.5.2 BRIEF DESCRIPTION OF GRAPHIC INPUT DEVICES.

A.5.2 (1) GRAFACON (60,125) - see figure A.1

Along a set of parallel "x" and "y" etched conductors, pulse trains are injected via special encoders such that at any location (the intersections of the "x" and "y" conductors), a unique pulse train of 20 pulses (10 bits for each "x" and "y") can be detected by the stylus tip, detection being capacitive. Grey code is used to minimize incorrect readings, since in this case, adjacent locations differ by only 1 bit.

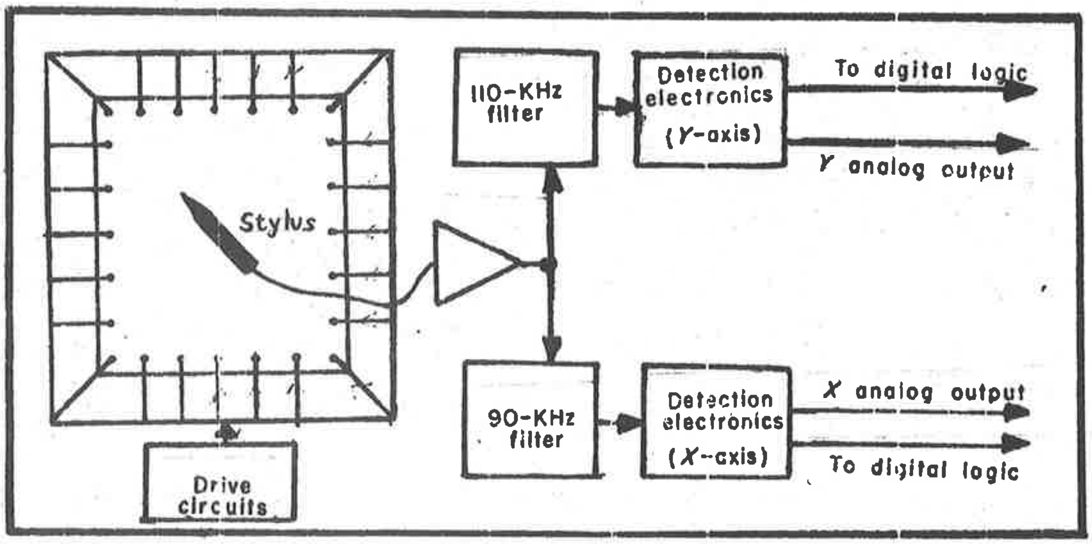
A.5.2. (2) SYLVANIA Data Tablet (60, 126) - see figure A.2

An analog system, this consists of a transparent conductive plane inserted between two sheets of glass (hence superposition over a display is possible). One Khz. sine and cosine signals modulated by 90 Khz. and 110 Khz. are repeatedly launched parallel to each "x" and "y" axis, their phases (1 Khz signal) being linear function of its position on the tablet



Grey code is used as then adjacent locations differ by only one bit. An error-checking block (not shown) compares each 20-bit x and y position word with the previous word.

Figure A.1 GRAFACON Graphic Tablet



The 'Detection Electronics' consists of filters, Synch. detectors, ramp generators, and thresh-hold detectors.

Figure A.2 SYLVANIA Data Tablet

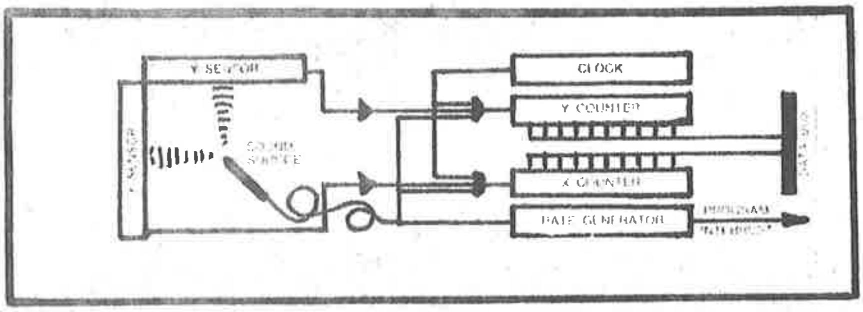
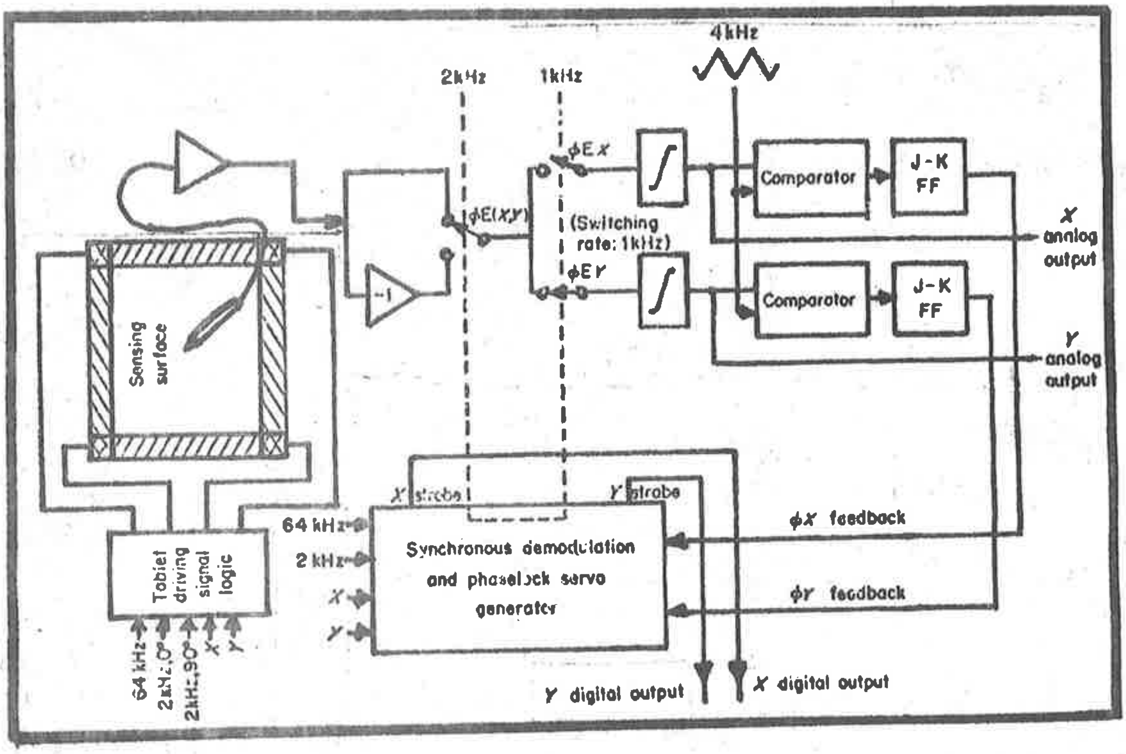


Figure A.3 GRAF-PEN Graphic Tablet



Compare with the Sylvania Data Tablet- both use phase information to indicate position, but the Ecricon, unlike the Sylvania Data Tablet, alternativly launches x and y waveforms.

Figure A.4 ECRICON Graphic Tablet (based on [60]) surface. The stylus tip capacitively detects these signals (and hence the phase) which are then demodulated and detected in a system of filters, synch. detectors, ramp generators and thresh-hold detectors. The two modulating frequencies make simultaneous "x" and "y" phase (and hence position) detection possible.

A.5.2. (3) GRAF-PEN Graphic Tablet (60, 127) - see figure A.3

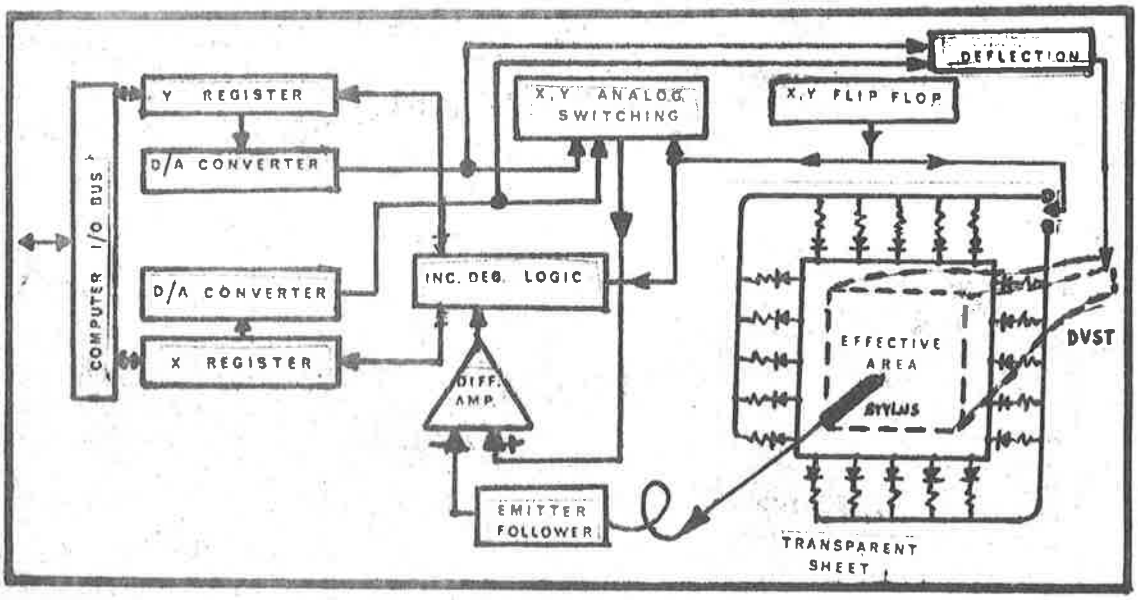
A small spark-gap at the pen-tip generates fast rise time sound pulses which are detected by strip microphones at the "x" and "y" axis boundaries. The time delay between spark generation and edge detection is proportional to the stylus tip distance from the edges.

A.5.2 (4) ECRICON Graphic Tablet (60,128) - see figure A.4

In principle this is similar to the Sylvania Data Tablet, in that position information is determined from phase information of sinusoidal waves launched between the "x" and "y" boundaries; detection by Stylus tip is again capacitive. Unlike the Sylvania Data Tablet, "x" and "y" waves are launched alternatively and hence have a common 64 Khz carrier.

A.5.2 (5) VOLTAGE PEN (Rose) (123) - see figure A.5

A transparent durable conductive sheet is placed over a display screen (either CRT or a DVST), and uniform voltage gradients are switched alternatively between the "x" and "y" directions. Position on the conductive area can be associated with the voltage at that point, the position-voltage relationship being linear. The stylus is a hard graphite pencil, (to provide a fine tip) which provides the feed from the detected voltage and the detection circuitry, and at the same time does not appreciably affect the quality of the conducting surface. The block diagram shown also shows the circuitry to generate the manually drawn points on



The writing area is superimposed over a DVST screen.

Figure A.5 Voltage Pen Tablet (Rose)

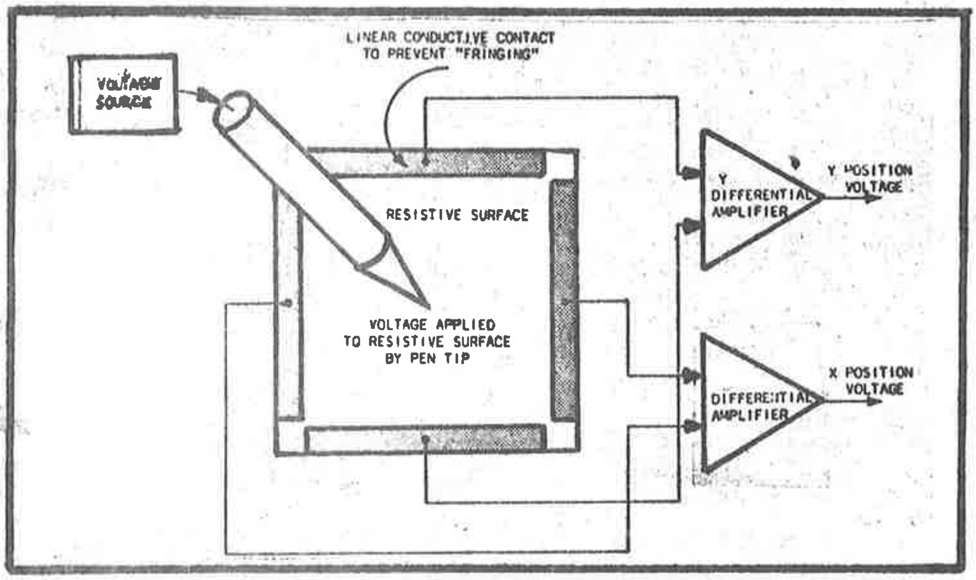


Figure A.6 Voltage Pen and Resistive Paper Input Tablet (from [129]) the display screen.

A.5.2 (6) Voltage Pen and Resistive Input Tablet (129) - see figure A.6

An input tablet of about 10" x 10" is made of a even resistive surface, occassionally of "resistive paper" ("Teledeltos" paper). A voltage is introduced via the pen tip at the desired locations. The resistive surface acts as a voltage divider in both the "x" and "y" direction and two sets of differential voltages appear at both pairs of the tablet boundaries, where, picked off by strip conductors and fed into two differential amplifiers (one each for "x" and "y" coordinates), they generate difference voltages which are linear analog representation of the "x" and "y" coordinates.

A.5.2 (7) Magnetically-coupled Pen and Tablet (174)

The pen contains a small airgapped ferrite core whose windings are excited periodically, generating a localized magnetic field. The tablet consists of twenty planar layers on which are parallel, finely spaced, interconnected conductors, 10 layers aligned longitudinally, the other 10 layers vertically aligned; their "intersections" specify the input tablet locations. Using thin film deposition the thickness of the total layers is no more than 1 mm. The magnetic field induces voltages in all the twenty windings at the location nearest to the pen (the field producing core is aligned 45° to the lines so that equal voltages are induced in both "x" and "y" lines). The lines are so interconnected that the output at the tablet sensors is a 20-bit Grey code specifying the tablet location digitally. The unit has recently (175) been released commercially.

A.5.2 (8) "Touch sensitive" X - Y positioner (176)

Piezo-electric strip transducers at the edges of the input area generate and radiate pulse-modulated "surface waves". An object such as a passive pen tip, or finger, reflects these waves which are sensed at the transducers at the edges. The time delay between transmitted and received echoes is proportional to the distance from the pen tip to the edge transducers and thus directly proportional to the "x - y" coordinates of the pen or finger.

A.5.2. (9) Surface wave Detecting Pen (177)

Two strip piezo-electric transducers, one for each coordinate, generate surface waves which are directly detected by the pen probe, which contains its own transducer. As in the above type of "surface-wave" device, "x, y" coordinates are proportional to the time delay between send and detect pulses.

A.5.2. (10) "Pressure Paper" Device - Analog (119)

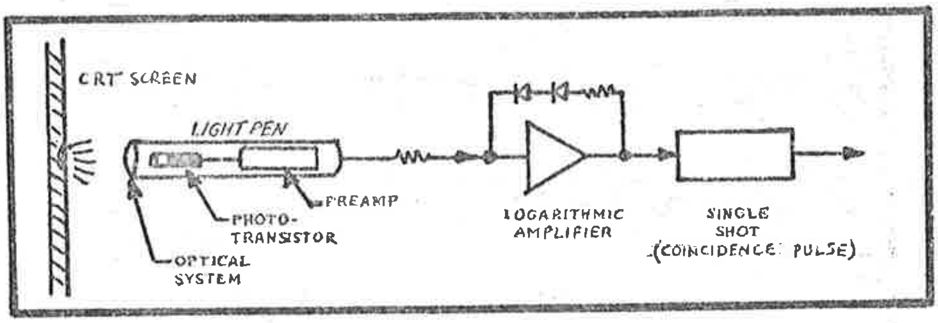
An electric potential gradient is set up across two separate resistive sheets, one in the "x", and the other in the "y" direction. The sheets are separated by two insulated sheets, one side which has a conductor of negligible resistance on it, which is connected to a D/A converter. The pressure of a passive pen tip causes contact between the conducting sheets and the corresponding resistive sheets, making voltages appear at the D/A converters, which due to the uniform gradients present, is proportional to the distance from a "x" or "y" edge.

A.5.2. (11) "Pressure Paper" Device - Digital (129)

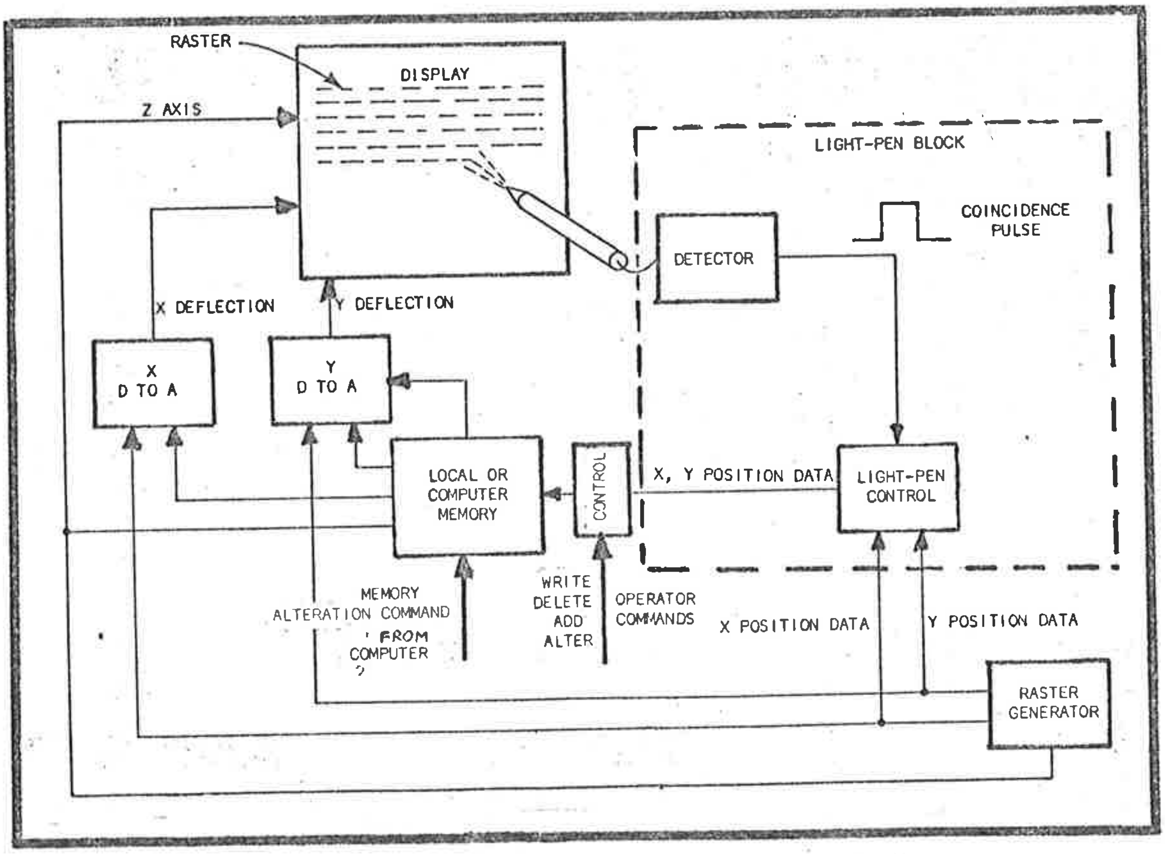
Identical to the analog device except that the sheets are "grooved", so that pressure contact is only within the area of the particular groove and not over a larger area; resolution (but not necessarily linearity) is thus improved.

A.5.2. (12) "Touch-wire" Device (178)

Strictly not a graphic input device but a "touch-



(a) Block Diagram of Light Pen



(b) Light Pen System Block Diagram (based on [129])

Figure A.7 LIGHT PEN

type" function keyboard, whereby items off a displayed "menu" can be selected. Within the display area associated with menu items display, passive copper wires are attached of spacing corresponding to menu item spacing and adjacent to them. The wires form one arm of a balanced bridge; when touched by a finger, the hand capacitance unbalance the corresponding bridge, which becomes unbalanced, generating a signal (100:1 S/N ratio). About 10 touch wires per display seem typical.

A.5.2. (13) Light Pen or Light Gun - see figure A.7 (25, 129, 130, 131, 179)

A photo-transistor in the light pen tip detects the initial flash ("luminescence" phase of phosphor excitation) of the CRT phosphor. (In some cases two phosphors are used, with the light-pen being sensitive to the fast decay rate phosphor). The resultant pulse is processed and generates a coincidence pulse, which when AND-gated with the raster generator timing pulses, indicate the detected initial flash location.

The block diagram shows the light pen operating in conjunction with a raster scanned display.

A.5.2. (14) Light Pen for Plasma Display Panel (113)

A tentative design to be used with a Plasma Panel Display, enabling direct input and erasure of information. Under certain gas-cell wall-voltage conditions, light emitted from the light pen liberates photo-electrons which reduce cell wall-voltages. Certain logic functions and switching voltages enable all such illuminated cells to be ignited.

A.5.2. (15) Beam Pen (180)

A pickup metal probe pen held as closely to the CRT screen as possible will be capacitively coupled to the

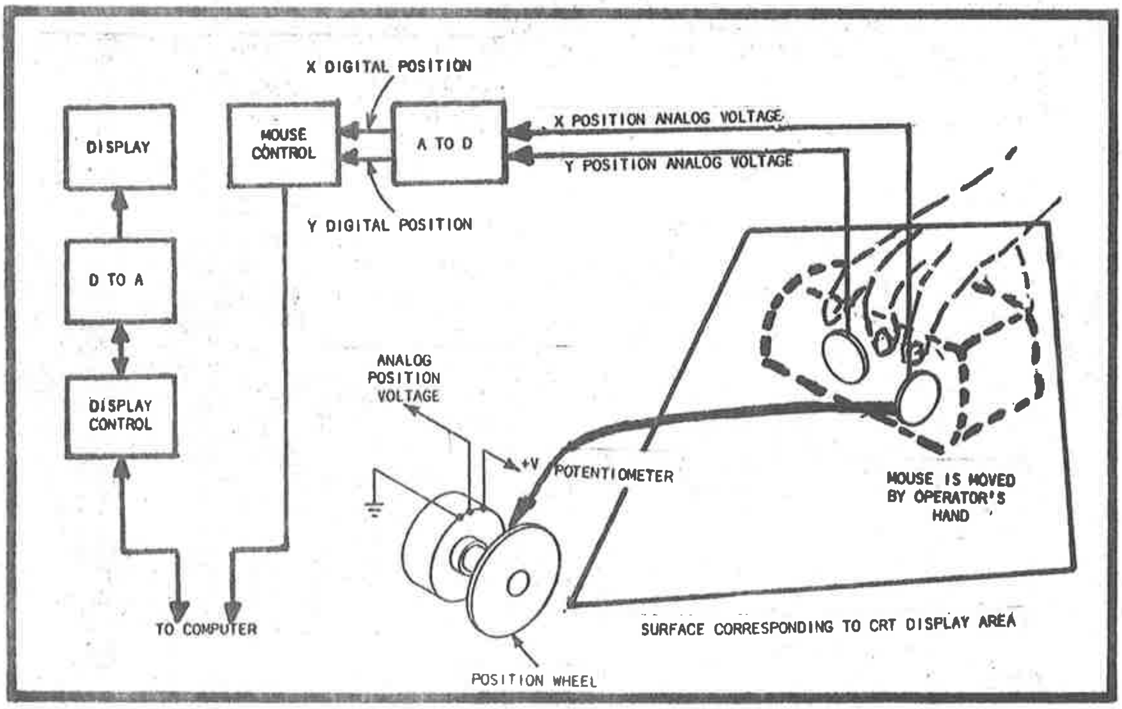
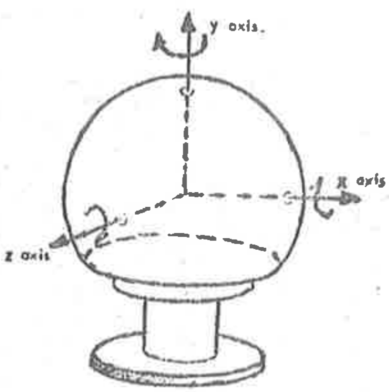


Figure A.8 Stanford Research Institute ' MOUSE ' Positioning Device



FOR EACH x, y, z, AXIS				
CODE g_1	g_2	g_3	g_4	MEANING
1	0	0	0	Max neg. rate
1	1	0	0	Med. neg. rate
0	1	0	0	Min. neg. rate
0	0	0	0	Reset zero rate
0	0	1	0	Min. pos. rate
0	0	1	1	Med. pos. rate
0	0	0	1	Max. pos. rate

Figure A.9 MIT ' CRYSTAL BALL ' for Rotating Displayed Images electron beam when it is in the immediate neighbourhood (as the electron beams striking the phosphor screen create charge variations in the phosphor). The electron beam is modulated at 10 Mcps to increase sensitivity and hence detection.

A.5.2. (16) <u>Lincoln Wand (132)</u>

Four ultrasonic generators (one for each x,y, z coordinate, the fourth for error checking) emit sound waves which are picked up and detected by a probe, the "wand". The time delay between "send" and "receive" is a measure of the probe-tip spatial coordinates. Working space is 4' x 4' x 6'.

A.5.2. (17) <u>Stanford Research Institute "Mouse" (119,129)-</u>
<u>see figure A.8</u>

A hand-held box, the "Mouse" is moved about on a plane surface. Two wheels, orthogonal to each other (ie "x" and "y" directions) rotate as it is moved; they are connected to two potentiometers, the terminals of which are to a voltage. The movable wiper has as its out-put a DC voltage proportional to wiper distance moved, and hence to present coordinates from a given origin.

A.5.2. (18) <u>"Joystick" or "Trackball" (119, 121)</u>

A movable joystick (angular movement), or a hand rotatable ball within an enclosure, has attached to its axis shaft encoders, the output of which are signals indicating the "x" and "y" coordinates, proportional to joystick movement or "trackball" rotation. Potentiometers as in the "Mouse" can also be used instead of shaft encoders.

A.5.2. (19) <u>MIT "Crystal Ball" (25)-see figure A.9</u>

A "three dimensional" globe or "Crystal ball" can be used to rotate displayed three dimensional images in any desired way; the amount of twist in any axis provides a <u>rate</u> of rotation (binary coded), with the total rotation being the time the sphere was held in that particular position multiplied by the particular rate of rotation selected.

A.5.2. (20) "Datacoders", X- Y Digitizers (119)

These are simple "x-y" electomechanical devices, with a pen or probe connected to "x" and "y" sliding arms, the probe position along each arm causing a voltage proportional to its position to be output each time a "plot point" button is pressed. They can have the appearance of small (or large) hard-copy plotters with the required drawing to be input placed in the working area, or else they can be mounted in front of the CRT display and the probe moved along the screen (or just above it) for free hand input or selection of display features.

A.5.2. (21) "Rho-Theta" Transducer (119)

A cantilevered arm having radial extension motion and rotation can be guided with a pen over some tracing or drawing. The movement is transmitted via universal joints to potentiometers whose terminals are connected to voltages - the movable pointer (of each potentiometer) has, as its output, a voltage proportional to the angular or radial motion of the pen tip.

. . . .

APPENDIX 5 - TABLE

GRAPHICS INPUT AND POINTING DEVICES

		DEVICE	TYPE	DATA RATE	RESOLUTION	ACCURACY	SIZE	FEATURES	COST	ORIGINATOR
GRAPHICS INPUT PAD OR TABLET DEVICES USING PENS	COMMERCIAL	'GRAFACON'	DIGITAL Capacitive pick-up from tablet to pen.	4500 points per second	10 bits for each "x" and "y"	0·1% absolute	10·24" x 10·24" or 20·48" x 20·48"	1. Tracing capability 2. Error checking in-built. 3. Can be superimposed over display.	$9000	DAVIS & ELLIS 1964 (60,125)
		'ECRICON'	ANALOG - Capacitive pick-up from tablet to pen.	2000 points per second	10 bits for each "x" and "y"	Line linearity 0.5% . 2% long term position accuracy.	11"x11"	1. "Ball-point" hard-copy capability.	$2000	SHINTRON Co. 1968-69 (60,128)
		'GRAPH-PEN'	ANALOG - pen generates sound picked up at tablet edges.	200 points per second	10 or 11 bits for each "x" and "y"	0·2% position. 0·1% in reproducability.	14"x14"	1. "Ball-point" hard copy capability. 2. Can be superimposed on display.	$2800	S.A.C 1968-69 (60,127)
		SYLVANIA DATA TABLET	ANALOG - Capacitive pick-up from tablet to pen.	2000 points per second	12 bits for each "x" and "y"	±1% in position for each "x" and "y".	11"x11"	1. Ball point hard-copy capability. 2. Pen height information (3 bits) available. 3. Can be superimposed over display.	$6900	TEIXERA & SALLEN 1968 (60,126)
	EXPERIMENTAL	VOLTAGE PEN 1	ANALOG - Resistive input surface acts as potential divider for Voltage generating pen.	1000 points per second	0.1%	±1% for each "x" and "y"	10"x10"	1. Resistive surface becomes worn and performance deteriorates.	?	(129)

APPENDIX 5 - TABLE (cont.)

GRAPHICS INPUT AND POINTING DEVICES

GRAPHICS INPUT PAD OR TABLET DEVICES USING PENS — EXPERIMENTAL

DEVICE	TYPE	DATA RATE	RESOLUTION	ACCURACY	SIZE	FEATURES	ORIGINATOR
VOLTAGE PEN 2	ANALOG - VOLTAGE GRADIENT PICK UP BY PEN.	Limited by writing speed (?)	9 bits for each x and y	0·1% of display height	3"x4" DVST	1. Developed for DVSTs 2. "Pencil" is ordinary lead pencil. 3. DAC-1 Voltage pencil very similar (no performance data given).	ROSE 1964 (123)
MAGNETICALLY COUPLED PEN	DIGITAL - MAGNETICALLY COUPLED PEN & PAD.	Limited by writing speed.	10 bits for each x and y	0·1% absolute	10" x 10"	1. 32 x 32 position tablet built of discrete wires (size 4" x 4") can be overlaid over display.	LEWIN 1965 (175)
"TOUCH-WIRE" Device	DIGITAL - FINGER TOUCHING UNBALANCES BRIDGE VIA USER'S CAPACITANCE.	-	Very Coarse.	——	Over small area of display surface.	Used as a "touch-select" device to pick items of a menu presented on CRT-Linear Array of up to say 10 "Touch Positions."	JOHNSON 1965 (175)
"TOUCH-SENSITIVE" x-y Positioner	ANALOG - ECHO RANGING USING SURFACE WAVES.	1000 points per second	0·4% contact area of $\frac{1}{4}$" required to activate	0·5% (Tentative)	10" x 10"	1. "Passive-pen" or finger can be used. 2. Can be overlaid over display.	HLADY 1969 (176)
"SURFACE WAVE" Detector	ANALOG - PEN DETECTION OF SURFACE WAVES.	1000 points per second	0·05% (Tentative)	0·1% (?)	20" x 20" (Tentative)	1. Tentative Proposal. 2. Hard Copy Possible.	WOO 1963 (177)
PRESSURE PAPER Device 1	ANALOG - PASSIVE PEN PRESSURE ENABLES D.C. CONTACT.	Limited by writing speed	±1% (Tentative)	±1% (?)	10" x 10" (?)	1. Prone to mechanical failure.	(119)
PRESSURE PAPER Device 2	DIGITAL - PASSIVE PEN PRESSURE ENABLES D.C CONTACT.	1000 points per second	32 lines per inch	within resolution capability	3" x 5"	1. Prone to mechanical failure (breaks in wires x conductors).	TUNIS & SIMER 1964 (119)

APPENDIX 5 - TABLE (cont.)

GRAPHICS INPUT AND POINTING DEVICES

		DEVICE	TYPE	DATA RATE	RESOLUTION	ACCURACY	SIZE	FEATURES	COST	ORIGINATOR
POINTING AND INPUT DEVICES USED IN CONJUNCTION WITH A DISPLAY	COMMERCIAL	LIGHT PEN, LIGHT GUN	LIGHT DETECTOR	1. Limited by scanning rate of CRT. 2. Light detector response lag is ≈1 us.	0·1" (Can be extended by scaling up display image)	within 0·1" - see resolution.	CRT Screen Size	no hard copy possible	$300 excluding back-up software	M.I.T(?) LINCOLN LABS. 1958-9 (25,130)
	EXPERIMENTAL	LIGHT PEN	LIGHT EMITTING PEN	"Very slow" as light beam must supply adequate photoelectrons to switch on cells		Display-screen parameter dependent		TENTATIVE PROPOSAL - in certain display mode can erase using light pen.		SLOTTOW & BITZER (1967) (113)
		BEAM PEN	ANALOG-capacitive coupling between scanning electron beam and pen	1. Limited by scanning rate of CRT 2. No response lag as in light detector pen	± 0·7" on CRT Screen	see resolution	CRT Screen Size	1. hard copy possible 2. does not operate with Aluminized-screen CRTs.		HARING & ARNOLD 1965 (180)
		"LINCOLN WAND"	ANALOG - pen detects ultrasonic waves generated at display edges.	25 points per second.	12 bits for each x, y, z	0·3% absolute	4'x4'x6'	1. "z"-dimension can be used as a scaler for component values displayed on the CRT. 2. No hard-copy or 3-D objects can be placed within "working area"		ROBERTS 1966 (132)
ELECTROMECHANICAL INPUT AND DISPLAY TRACKING DEVICES	COMMERCIAL	S.R.I. MOUSE	ANALOG-Rotary Wheels Drive potentiometers and thus Position from origin	"FAST" (Used for Cursor Positioning)	0·1%	±0·5%	4"x3" Device (Input area no limit)	1. Positions Cursor on CRT 2. Has simple Graphic input feasability. 3. No hard copy possible	$400	S.R.I 1965 (119,129)
		JOYSTICK TRACKBALL	ANALOG - Movable "Ball" or "Joystick" moves Shaft encoders thus Generating "Position"	"FAST" (Used for Cursor Positioning)	0·1%	±0·5%	6"x4"x2" Display screen area.	1. Positions Cursor on CRT 2. Extremely Difficult to generate graphics 3. No hard copy possible	$360 JOYSTICK	M.I.T 1963 (121)
		"DATACODER"	ANALOG - Electro-mechanical device Analogous to x-y Plotter	"SLOW" (30-100 points per second in "trace mode")	0·1%	0·1%	10"x10" to 6'x6'	1. Mainly for tracing from hard copy. 2. Cumbersome	$1000	mid 1950's (119)
		"RHO-THETA" Transducer	ANALOG - Electro-mechanical device drives 2 Potentiometers Generating Angular & Radial Coordinates	MODERATELY FAST. (Several ins of Trace Sec)	0·4%	± 0·04" within working area	10" Radius x 100° Rotation.	1. Accomodates large drawing areas. 2. May require coordinate transformation to Cartesian coords.	"Low" - several $ 100's	BOLT BERANEK & NEWMAN (119)

APPENDIX 6

CRT AND VIDICON BEAM DEFLECTION

1 BEAM LENGTH INCREASE DURING SCANNING

2 VOLTAGE DRIVEN CRT V-COILS

3 CRT DEFLECTION DISTORTION

4 VIDICON DEFLECTION DISTORTION

APPENDIX 6

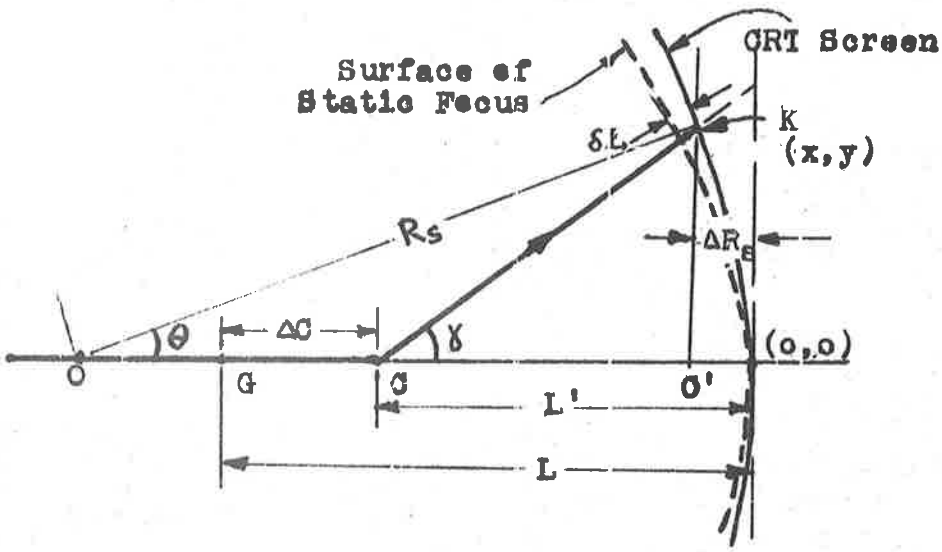
CRT AND VIDICON BEAM DEFLECTION

A.6.1 DERIVATION OF "δL" THE ELECTRON BEAM LENGTH INCREASE DUE TO BEAM DEFLECTION

For CRTs the deflection coils are nearer the screen than the focus coil.

Call the centre of the focus coil region "G", "L" units distant from the screen. Let the centre of beam deflection be "C", '$L - \Delta C = L'$ " units from the screen, where "ΔC" is the separation between "C" and "G". Let the screen radius of curvature be "R_s" units, and centered at O, with

$R_s > L$ (this is usually true).



For the beam incident on some screen location K(x,y) (with the undeflected beam located centrally on the screen at x = 0, y = 0), then from the diagram

$$\Delta R_s = R_s(1 - \cos\theta) \qquad \text{A.6.1.1}$$

where "ΔR_s" is the screen deviation from flatness.

Using the approximation

$$\cos\theta = 1 - \frac{\theta^2}{2} \quad \text{(accurate to 1 \% for } \theta < 25^\circ\text{)}$$

where $\theta = \arcsin\frac{\sqrt{x^2+y^2}}{R_s} \approx \frac{\sqrt{x^2+y^2}}{R_s}$ (accurate to some 3% for $\theta < 25^\circ$).

The maximum angle of deflection "γ" in a projection CRT is some 25°, and it can be seen from the diagram that $\gamma > \theta$. Hence the above approximation is valid for the maximum CRT deflection angle.

Hence $$\Delta R_s = R_s \cdot \frac{\theta^2}{2} = \frac{x^2+y^2}{2R_s} \quad \ldots\ldots\ldots\ldots\ldots \text{A.6.1.2}$$

From triangle CO′K in the figure

$$(L' + \delta L)^2 = x^2 + y^2 + (L' - \Delta R_s)^2 \quad \ldots\ldots\ldots \text{A.6.1.3}$$

and solving the quadratic for "δL" results in

$$\delta L \approx \frac{x^2+y^2}{2L'}\left(1 - \frac{L'}{R_s}\right)$$

$$\therefore \quad \delta L \approx \frac{x^2+y^2}{2(L - \Delta C)}\left(1 - \frac{L - \Delta C}{R_s}\right) \quad \ldots\ldots\ldots\ldots \text{A.6.1.4}$$

as it was assumed that the term $(\Delta R_s)^2$ was negligible with "$x^2 + y^2$" and "$L'\Delta R_s$".

6.2 FEASIBILITY OF RETAINING VOLTAGE DRIVEN VERTICAL DEFLECTION COILS IN CRTs

In commercial TV-CRTs, the vertical deflection coils are very often fed by ramp voltage scanning waveforms rather than ramp current waveforms; this is because at the V-scanning rates of 50-60 Hz. the impedance of the V-coils is primarily resistive, the coil inductance introducing only a slight delay at the start of scanning. (For H-coils, with scanning frequencies 300 times higher, voltage scanning waveforms are out of the question).

The possibility arises of using the existing ramp voltage scanning circuits for vertical deflection in the projection CRT, with correction accomplished also by voltage waveforms, rather than the more difficult to implement current scanning waveforms, and still obtain the required display linearity. The expected current scanning waveform delays will still be present due to the associated RL time constant, but remembering that the initial segment of the display (the topmost part of the CRT display) and the first segment of each horizontal scan-line need not be linear, these present delays etc. could be tolerated. Hence economy of cost will result if existing circuits are used.

The V-deflection coil will be taken as a series RL element driven by the voltage ramp or staircase; the requirement is to find the transient and steady state solution - hence Laplace methods are used. Typically for the 14 inch PYE TV-monitor on hand $R_V = 6{\cdot}6\,\Omega$, $L_V = 5{\cdot}2$ mH.

For the ramp voltage "V(t)" of slope "m",

$$V(t) = m.t \qquad \ldots\ldots\ldots \text{A.6.2.1}$$

the current is given by

$$i(s) = \frac{V(s)}{L\left(s + \frac{R}{L}\right)} \quad \ldots\ldots\ldots\ldots \text{A.6.2.2}$$

where $\mathcal{L}v(t) = V(s) = \frac{m}{s^2}$

Thus

$$i(s) = \frac{m}{s^2} \cdot \frac{1}{L} \cdot \frac{1}{\left(s + \frac{R}{L}\right)} \quad \ldots\ldots\ldots\ldots \text{A.6.2.3}$$

which on solving gives

$$i(t) = \frac{mt}{R} + \frac{m}{R} \cdot \frac{L}{R}\left(e^{-\frac{t}{\tau}} - 1\right) \quad \ldots\ldots\ldots\ldots \text{A.6.2.4}$$

$$= \frac{m}{R}(t - \tau) + \frac{m}{R}\,\tau . e^{-\frac{t}{\tau}} \quad \ldots\ldots \text{A.6.2.4(a)}$$

where $\frac{\text{"}mt\text{"}}{R}$ is the required ramp current

$\text{"}\frac{m}{R} \cdot \frac{L}{R}\left(e^{-\frac{t}{\tau}}\right)$ is the unwanted transient term

"τ" $= \frac{L}{R}$, the time constant.

For the transient term to be less than 1% of the steady state term (and hence contribute less than 1% of distortion of scanning waveform), it is required that

$$\frac{e^{-\frac{t}{\tau}}}{\frac{t}{\tau} - 1} \leqslant 0 \cdot 01 \quad \ldots\ldots\ldots\ldots \text{A.6.2.5}$$

For $t \geqslant 3 \cdot 6\,\tau$, this is satisfied.

From the given typical values of L,R, giving $\tau = 0 \cdot 79$ms, this corresponds to a "t" of about 2·8ms. As only the central $\frac{2}{3}$ of the display height is used, a vertical strip of $\frac{1}{2}.\frac{1}{3}.18\,ms = 3ms$, corresponding to the first 3ms of the vertical scanning interval, is not required for display purposes.

Hence during the active display area, the current scanning waveform is better than 1% linear if a linear voltage waveform is input.

For the actual case a staircase current is required; feeding a staircase voltage, the above expressions are modified somewhat.

Each step height is generated within 10 μs (the H-scan blanking time) and is constant for the remaining 54 μs of the 64 μs scan line length. The step slope within the region of 10 μs is thus 6·4 times greater than previously, for the same overall scanning waveform amplitude

The staircase is of the form shown here.

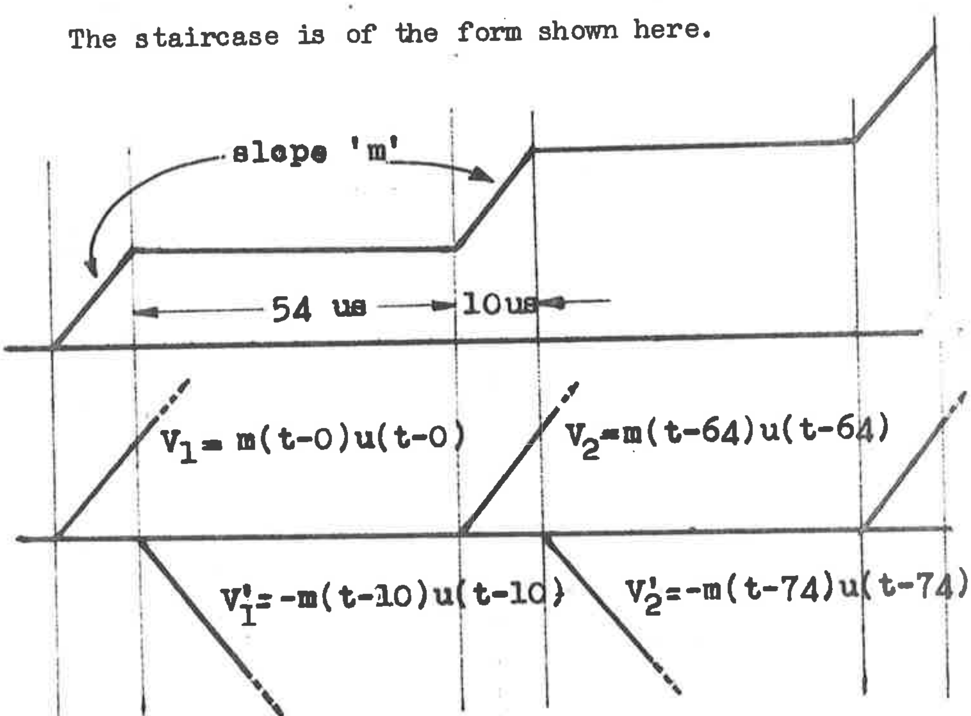


Figure A. Voltage Staircase defined for Laplace Analysis.

Each step is formed from 2 ramps with slopes of opposite signs, and the origin of one with respect to the other is displaced by $t = 10\,\mu s$. In the analysis the voltage, current and transient terms, all contain these summed ramps modified by step functions to indicate their starting times. The analysis, though not difficult, is somewhat tedious and very lengthy

However the expressions are wholly analogous to the ramp voltage case in expression 4(a).

The analysis shows that current staircase instead of being asymptotic to a line shifted by "τ" secs. from the time origin (see Figure B) it is asymptotic to a line shifted by $\tau(1 + \frac{x-a}{x})$ seconds,

where 'x' = stair length (64 μs)

"a" = step ramp length (10 μs)

(For a = x, i.e. a ramp, it reduces back to the former case). For the values shown this is "1·84 τ" and using expression (5), the 1% deviation off linearity is reached by 3·9 τ or just over 3 ms. Hence to correct for this added distortion (greater than 1% (on top of existing 0·5-1% distortion) as the distortion here is due to scanning current waveform linearity) more stages in the correction circuits are required.

More importantly the analysis shows that this added delay of "0·84 τ" is at the expense of the current waveform being unable to follow the voltage staircase.

For our purpose thus voltage controlled CRT Vertical scanning coils are out of the question.

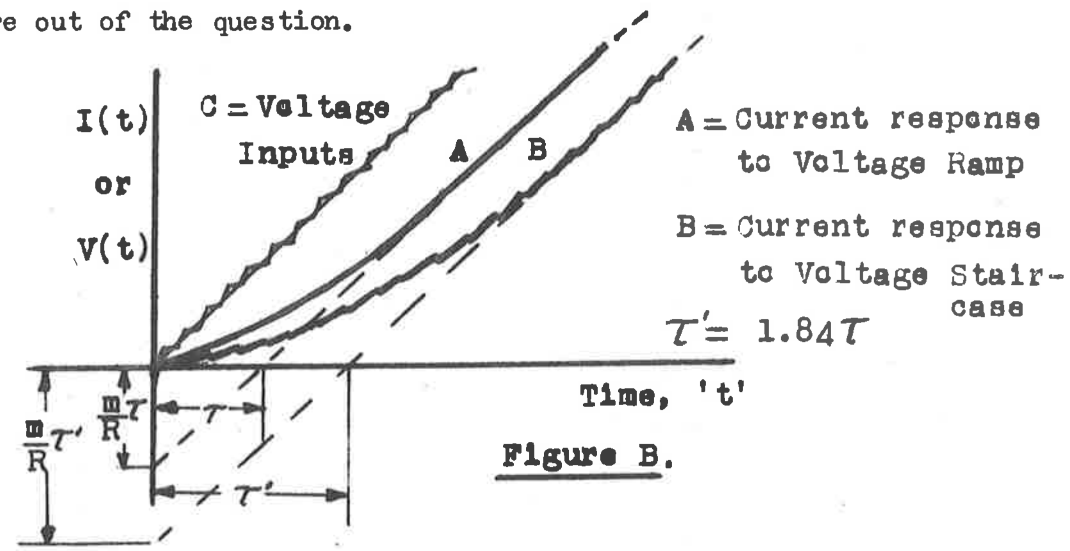


Figure B.

A.6.3 DERIVATION OF CRT BEAM DEFLECTION AND DISTORTION

PINCHUSION DISTORTION IN MAGNETICALLY DEFLECTED C.R.T's WITH FLAT AND NON-FLAT SCREENS

by

Victor C. Sobolewski, Student Member

Submitted to "INFORMATION DISPLAY"

The author is with the Electrical Engineering Department, University of Adelaide, Adelaide, South Australia, and the Weapons Research Establishment, Salisbury, South Australia.

1. INTRODUCTION

Expressions indicating pincushion distortion present in magnetic deflection display tubes have been known for some time[1,2]. Thus the deflection, say in the y-direction, may be expressed in terms of linear coil deflection currents[2] such as

$$y = k_1 Iy \left\{ 1 + \underbrace{k_2(Ix^2 + Iy^2)}_{\epsilon_1} + \underbrace{k_3(Ix^2 + Iy^2)^2}_{\epsilon_2} + \cdots \right\} \quad \ldots\ldots (1)$$

where "y" is the deflection

"Iy" is the linear y-coil deflection current

"Ix" is the linear x-coil deflection current

"k_1", "k_2", "k_3" are constants.

"ϵ_1", "ϵ_2" are the "deviation-from-linearity" terms. Or else it may be expressed in terms of the ideal deflections "y_o", "x_o", [1] such as

$$y = y_o(1 + \underbrace{k(x^2 + y^2)}_{\epsilon_1}) \quad \ldots\ldots (2)$$

where "k" is some constant.

In the first case, for example, feeding into the deflection coil a current of the form

$$Iy' = Iy(1 - k_2(Ix^2 + Iy^2)) \quad \ldots\ldots (3)$$

results in $y = k_1 Iy$ nearly exactly, giving a linear deflection for a linear sweep current. This current Iy′ may be quite easily generated with analog squarers and adders, or multipliers, or even quite closely approximated by certain sections of sinusoids.

Analogous expressions of course hold for the x-deflection. Recently, display distortion, of which pincushion distortion contributes a large part, has become more of interest as the need for high quality linear TV-mode displays has grown, not only using flat-faced screens with small deflection angles, but also those with curved screens and large deflection angles. Particularly in the latter case, the usual expressions for display distortion in one of the above forms are less accurate, as will be seen when the assumptions made in the derivation of these expressions are closely examined.

The usual derivation of expressions (1) and (2) is based on the equating of the force acting on an electron moving in a magnetic field which is normal to its direction of motion, with the centripetal force. Thus

$$B_x . e . v = \frac{m . v^2}{Ry} \qquad (4)$$

where "B_x", is the magnetic flux density producing the y-deflection, assumed to be constant at any instant of electron transit,

"e, m" are the charge and mass of the electron, respectively

"v" is the velocity of the electron,

"Ry" is the radius of curvature of the circular path.

The velocity of the electron is given by

$$v = \sqrt{\frac{2 . e . V}{m}} \qquad (5)$$

where "V" is the final accelerating anode voltage, of the order of 10KV. The deflecting field "B_x" is usually considered to be constant (for a given deflection) within the coil region of length "L", and is zero outside this region.

From Figure 1(a) which shows the deflection and the associated geometry, it can be seen from simple geometrical considerations that

$$\varphi = \theta, \text{the deflection angle} \quad \ldots\ldots\ldots\ldots\ldots\ldots (6)$$

In deriving expression (1) and (2) above, it is considered that the electron beam on deflection has, as its origin, the point "C," called the "center of deflection", which is considered to be constant and located at the center of the deflecting magnetic field region.

The total deflection is then considered to be

$$y = \left(S - \frac{L}{2} + CM\right) \tan\theta$$

$$= S.\tan\theta \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots (7)$$

Expanding "$\tan\theta$" into a series and manipulating expressions (4), (6), (7) in both x- and y-dimensions leads to the results (1) and (2).

In actual fact, the distance "CM", taken as constant and equal to "$\frac{L}{2}$" in the above type of derivation, varies with the angle of deflection quite significantly, particularly when second-order effects "ϵ_1" and "ϵ_2", which indicate deviation from linearity, are taken into account. From the figure

$$CM = L - OC$$

$$= L - R\tan\frac{\theta}{2}$$

$$= \frac{L}{2}\left[2 - \sec^2\frac{\theta}{2}\right] \qquad \text{as } L = R\sin\theta$$

Expanding "$\sec^2\frac{\theta}{2}$" as a series and collecting terms gives

$$CM = \frac{L}{2}\left(1 - \left(\frac{\theta}{2}\right)^2 - \frac{2}{3}\left(\frac{\theta}{2}\right)^4 - \cdots\right) \ldots\ldots\ldots\ldots\ldots (8)$$

For a maximum deflection angle of say 90° (across a diagonal), θ is $\frac{\pi}{4}$, resulting in "CM" deviating by about -17% from its nominal value.

From Figure 2 it is seen that if the "origin", C, of the electron beam is shifted by

$$\delta L = -\frac{L}{2}\left(\left(\frac{\theta}{2}\right)^2 + \frac{2}{3}\left(\frac{\theta}{2}\right)^4 + \cdots\cdots\right)$$

the deflection "y" on some flat-faced screen, "S" away from the center "C", is shifted by some "δy". From similar triangles C'KY', CKY

$$\frac{S - \delta L}{S} = \frac{y - \delta y}{y} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (9)$$

or

$$-\frac{\delta y}{y} = -\frac{\delta L}{S} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (9a)$$

For a hypothetical C.R.T. where "S" is say 12", "L" is 2", then, for a diagonal deflection angle of 90°, a square raster where $y_{max} \approx 8{\cdot}5"$, δL is 0·17", gives,

$$-\frac{\delta y}{y} \approx -1{\cdot}4\%$$

This is of opposite sign to pincushion distortion and usually smaller in magnitude, and is solely due to the shifting of the "center of deflection" "C" as the deflection angle is varied. It is therefore expected that any formulation for pincushion distortion where the assumption that the position of "C" is constant is not made, will give pincushion distortion of smaller magnitude then the expressions in (1) and (2). This is indeed what has been found below.

The method adopted here is based on classical electron optics, using the Principle of Least Action for a moving electron acted upon by an electric accelerating field and a magnetic deflecting field. The resulting differential equations need some slight approximations so that an analytic solution may easily be obtained.

The solutions of the differential equations can be shown to be identical to the first few terms of the expressions (1) or (2), but the coefficients of the terms of the form "$Ix(Ix^2 + Iy^2)^n$" are slightly reduced, being a result of the length "CM" varying with the deflection.

The main advantages of the formulation adopted is that in principle, solutions for non-constant (within the length "L") magnetic field distributions can be solved for; also, as the solution is in terms of "z", the distance along the Z-axis, then "$\frac{dy}{dz}$", corresponding to "$\tan\theta$" in expression (7) can easily be found. For that reason the solution for flat-faced screens, can easily be extended to screens which are sections of spheres, of radius greater than "S", or indeed of screens of any known geometric surface.

2. PRINCIPLE OF LEAST ACTION AND DERIVATION OF EQUATIONS OF MOTION

For a conservative dynamical system, for example an electron acted upon by some electrical and magnetic field and exchanging potential energy for kinetic energy, the "action integral", "A", defined by

$$A = \int_{P_1}^{P_2} \frac{\partial L}{\partial \underset{\sim}{v}} \cdot \underset{\sim}{v}\, dt \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (10)$$

has some stationary value when evaluated between 2 positional states, P_1 to P_2[3].

"$\frac{\partial L}{\partial \underset{\sim}{v}}$" is the generalized momentum of mass "m" (in this case the electron of mass "m") and "L" is the Lagrangian function of the electron, which links the kinetic energy to the potential energy of the electron. For electric and magnetic fields present,

$$L = mc^2(1 - \sqrt{1 - \beta^2}) + eV - e(\underset{\sim}{A} \cdot \dot{\underset{\sim}{x}}) \quad \ldots\ldots\ldots\ldots (11)$$

where $\beta^2 = \frac{1}{c^2}\sum \dot{x}_i{}^2$

"c" is the velocity of light in free space

"V" is the potential of the accelerating electric field

"$\underset{\sim}{A}$" is the vector potential of the magnetic field

"$\dot{\underset{\sim}{x}}_i$" $= \frac{d\underset{\sim}{x}_i}{dt} = \underset{\sim}{v}_i$, the velocity of the electron along the x_i-th coordinate direction i.e. along, the x-, y-, or z-axis.

For small velocities (corresponding to V < 5KV, where the error is <1%), the relativistic expression

$$mc^2(1 - \sqrt{1 - \beta^2})$$

reduces to $\frac{1}{2}m\sum \dot{x}_i{}^2$. Using this, the expression (11) for the Lagrangian becomes

$$L = \tfrac{1}{2}m\sum \dot{x}_i{}^2 + eV - e(\underset{\sim}{A} \cdot \dot{\underset{\sim}{x}}) \quad \ldots\ldots\ldots\ldots (12)$$

Now

$\frac{\partial L}{\partial \underset{\sim}{v}} = \frac{\partial L}{\partial \dot{\underset{\sim}{x}}_i}$ differentiated over each coordinate axis,

gives

$$\frac{\partial L}{\partial \underset{\sim}{v}} = m\underset{\sim}{v} \cdot \underset{\sim}{u} - e\underset{\sim}{A} \cdot \underset{\sim}{u} \quad \ldots\ldots\ldots\ldots (13)$$

where $\underset{\sim}{u}$ is the unit vector.

This further reduces to the following when the substitution $v = \sqrt{\frac{2eV}{m}}$ is made:

$$\frac{\partial L}{\partial \underset{\sim}{v}} = m\sqrt{\frac{2eV}{m}} - e\underset{\sim}{A} \cdot \underset{\sim}{u} \quad \ldots\ldots\ldots\ldots (13a)$$

The relativistic expression corresponding to this becomes

$$\frac{\partial L}{\partial \underset{\sim}{v}} = m\sqrt{\frac{2eV}{m} + \left(\frac{e\ V}{m\ c}\right)^2} - e\underset{\sim}{A} \cdot \underset{\sim}{u} \quad \ldots\ldots\ldots\ldots (13b)$$

Substituting (13a) into the Action Integral (10), results in

$$\int_{P_1}^{P_2} \left(m\sqrt{\frac{2eV}{m}} - e\underset{\sim}{A} \cdot \underset{\sim}{u}\right) v.dt = K \quad \ldots\ldots\ldots\ldots (14)$$

where "K" is its stationary value.

It must be noted that the above holds only for "monochromatic" rays - i.e. electrons of constant initial velocity.

Because the above integral has a stationary value it may be changed to a variational equation of the form.

$$\delta \int_{P_1}^{P_2} \left(m\sqrt{\frac{2eV}{m}} - e\underset{\sim}{A} \cdot \underset{\sim}{u}\right) v.dt = 0 \quad \ldots\ldots\ldots\ldots (15)$$

where "δ" refers to the variational symbol.

This is another way of stating that from the same initial conditions, and the same electric and magnetic fields, only one electron path can be followed from P_1 to P_2. This path of interest is the one for an electron entering a deflecting magnetic field at some $z = P_1 = z_0$ with a velocity corresponding to $v = \sqrt{\frac{2eV}{m}}$, and leaving it at some $z = P_2 = z_1$. The integration must then be with respect to "z" and the appropriate change of variables must be made.

As $$v = \left|\frac{d\underset{\sim}{s}}{dt}\right| = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2 + \left(\frac{dz}{dt}\right)^2}$$

then

$$v dt = ds = v\frac{dt}{dz} . dz = \sqrt{\left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2 + 1} . dz \quad \ldots\ldots\ldots (16)$$

Also

$$(e\underset{\sim}{A} . \underset{\sim}{u}) v dt = e\underset{\sim}{A} . d\underset{\sim}{s}$$

$$= e\left(A_x \frac{dx}{dz} + A_y \frac{dy}{dz} + A_z\right) dz \quad \ldots\ldots\ldots (17)$$

Substituting these expressions into the variational expression (15), then

$$\delta \int_{z_0}^{z_1} \left(m\sqrt{\frac{2eV}{m}} . \sqrt{\left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2 + 1} - e\left(A_x \frac{dx}{dz} + A_y . \frac{dy}{dz} + 1\right)\right) dz = 0 \quad (18)$$

which is of the form

$$\delta \int_{z_0}^{z_1} F\left(x, y, z, \frac{dx}{dz}, \frac{dy}{dz}\right) dz = 0 \quad \ldots\ldots\ldots (19)$$

which must obey the variational principle i.e. $F\left(x, y, z, \frac{dy}{dz}, \frac{dx}{dz}\right)$ obey the Euler-Lagrange equations which are

$$\frac{\partial F}{\partial y} - \frac{d}{dz}\left(\frac{\partial F}{\partial\left(\frac{dy}{dz}\right)}\right) = 0 \quad \ldots\ldots\ldots (20a)$$

$$\frac{\partial F}{\partial x} - \frac{d}{dz}\left(\frac{\partial F}{\partial\left(\frac{dx}{dz}\right)}\right) = 0 \quad \ldots\ldots\ldots (20b)$$

Obtaining the particular form of these equations where

$$F = m\sqrt{\frac{2eV}{m}} . \sqrt{\left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2 + 1} - e\left(A_x \frac{dx}{dz} + A_y \frac{dy}{dz} + 1\right) \quad \ldots\ldots\ldots (21)$$

results in the differential equations whose solutions give the required equations of motion of the electron (beam) in the X-Y plane of the display screen.

Only the expression (20a) is solved here giving the y-deflection. The x-deflection obtained from (20b) (actually both must be solved simultaneously) is of an analogous form.

Now differentiating (21) according to (20a), and transposing terms gives

$$\frac{m}{e}\sqrt{\frac{2eV}{m}} \cdot \frac{d}{dz}\left(\frac{\frac{dy}{dz}}{\left\{\left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2 + 1\right\}^{\frac{1}{2}}}\right) = \frac{dAy}{dz} - \frac{\partial Az}{\partial y} \quad \ldots\ldots\ldots (22)$$

Making the approximation that $\frac{dAy}{dz} = \frac{\partial Ay}{\partial z}$ and remembering that

$$\frac{\partial Ay}{\partial z} - \frac{\partial Az}{\partial y} = -B_x \quad \ldots\ldots\ldots (23)$$

which is the magnetic flux density component in the x-direction (i.e. the magnetic flux producing y-deflection), and also assuming that

$$\left(\frac{dx}{dz}\right)^2 << 1, \left(\frac{dy}{dz}\right)^2 << 1$$

(i.e. small deflections only) results in

$$\frac{d^2y}{dz^2} = -\sqrt{\frac{e}{2mV}} \cdot B_x \quad \ldots\ldots\ldots (24a)$$

A similar expression results for the x-deflection namely

$$\frac{d^2x}{dz^2} = \sqrt{\frac{e}{2mV}} \cdot B_y \quad \ldots\ldots\ldots (24b)$$

(The difference in sign arises because of the way $\underset{\sim}{A}$, the magnetic potential was chosen and the convention in defining B_x, B_y etc. from $\underset{\sim}{A}$).

The expressions (24a,b) are the 1st order approximation differential equations defining the motion of each electron.

Integrating (24a) twice results in

$$y = y_0 + y_0' (z_1 - z_0) - \sqrt{\frac{e}{2mV}} \int_{z_0}^{z_1} \int_{z_0}^{z_1} B_x . dz . dz \quad \ldots\ldots\ldots\ldots (25)$$

where $y_0' = \frac{dy}{dz}$ at point (x_0, y_0, z_0),

with a corresponding expression giving the x-deflection. If the origin for deflection $(y_0, z_0, x_0) = (o,o,o)$, and $z_1 = L$, the length of the magnetic field, and for a beam entering the deflecting beam normally, resulting in $y_0' = o$, $x_0' = o$ then

$$y = -\sqrt{\frac{e}{2mV}} \int_{o}^{L} \int_{o}^{L} B_x . dz . dz \quad \ldots\ldots\ldots\ldots (25a)$$

and

$$x = \sqrt{\frac{e}{2mV}} \int_{o}^{L} \int_{o}^{L} B_y . dz . dz \quad \ldots\ldots\ldots\ldots (25b)$$

If the magnetic flux density distribution B_x, and B_y is known in the z-direction, then the above can be integrated directly.

The usual conditions taken when using the geometrical method described in Section 1, are

(1) a constant magnetic flux density (at some constant deflection) within the interval $z = o$ to $z = L$.

(2) the electron beam is projected onto a plane screen at $z = S + \frac{L}{2}$.

and (3) the y-deflection on such a screen is then given as

$$y = y\Big]_{z=L} + \frac{dy}{dz}\Big]_{z=L} \left(S - \frac{L}{2}\right) \quad \ldots\ldots\ldots\ldots (26)$$

Then expression (25a) reduces to

$$y = -\sqrt{\frac{e}{2mV}} \cdot B_x \cdot \frac{L^2}{2} - \left\{\sqrt{\frac{e}{2mV}} \cdot B_x \int_0^L dz\right\}\left(S - \frac{L}{2}\right) \quad \ldots\ldots\ldots (27)$$

$$= -\sqrt{\frac{e}{2mV}} \cdot B_x \cdot L \cdot S \quad \ldots\ldots\ldots (27a)$$

which is the usual expression obtained by equating the centripetal force with the magnetic field force as in Section 1.

3. PINCUSHION AND OTHER DISTORTIONS

To derive the expressions giving pincushion distortion etc., the approximations used before, namely that $\left(\frac{dx}{dz}\right)^2 << 1$, $\left(\frac{dy}{dz}\right)^2 << 1$ can no longer be made, when differentiating expression (22). Proceeding with the differentiation gives

$$\frac{\frac{d^2y}{dz^2}}{\left\{\left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2 + 1\right\}^{\frac{1}{2}}} - \frac{\frac{d^2y}{dz^2} \cdot \left(\frac{dy}{dz}\right)^2 + \frac{d^2x}{dz^2} \cdot \frac{dy}{dz} \cdot \frac{dx}{dz}}{\left\{\left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2 + 1\right\}^{\frac{3}{2}}} = -\sqrt{\frac{e}{2mV}} \cdot B_x \quad \ldots (28)$$

Bringing the L.H.S. over the common denominator $\left\{\left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2 + 1\right)^{\frac{3}{2}}$ and making the approximation

$$\left\{\left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2 + 1\right\}^{\frac{3}{2}} = \frac{3}{2}\left(\frac{dx}{dz}\right)^2 + \frac{3}{2}\left(\frac{dy}{dz}\right)^2 + 1$$

and collecting terms gives

$$\frac{\frac{d^2y}{dz^2}\left\{\frac{3}{2}\left(\frac{dx}{dz}\right)^2 + \frac{3}{2}\left(\frac{dy}{dz}\right)^2 + 1\right\}}{\left\{\frac{3}{2}\left(\frac{dx}{dz}\right)^2 + \frac{3}{2}\left(\frac{dy}{dz}\right)^2 + 1\right\}} - \frac{\frac{d^2x}{dz^2} \cdot \frac{dy}{dz} \cdot \frac{dx}{dz} + \frac{1}{2}\frac{d^2y}{dz^2} \cdot \left(\frac{dx}{dz}\right)^2 + \frac{3}{2}\frac{d^2y}{dz^2} \cdot \left(\frac{dy}{dz}\right)}{\left\{\left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2 + 1\right\}^{\frac{3}{2}}}$$

$$= -\sqrt{\frac{e}{2mV}} \cdot B_x \quad \ldots\ldots\ldots (29)$$

and again making the approximation that

$$\left\{\left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2 + 1\right\}^{-\frac{3}{2}} = 1 - \tfrac{3}{2}\left(\frac{dx}{dz}\right)^2 - \tfrac{3}{2}\left(\frac{dy}{dz}\right)^2$$

then

$$\frac{d^2y}{dz^2} = -\sqrt{\frac{e}{2mV}}\cdot B_x + \underbrace{\tfrac{3}{2}\frac{d^2y}{dz^2}\left(\frac{dy}{dz}\right)^2 + \tfrac{1}{2}\frac{d^2y}{dz^2}\cdot\left(\frac{dx}{dz}\right)^2 + \frac{d^2x}{dz^2}\cdot\frac{dy}{dz}\cdot\frac{dx}{dz}}_{\epsilon_1}$$

$$\underbrace{-\tfrac{3}{2}\left\{\left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2\right\}\left(\tfrac{3}{2}\cdot\frac{d^2y}{dz^2}\left(\frac{dy}{dz}\right)^2 + \tfrac{1}{2}\frac{d^2y}{dz^2}\left(\frac{dx}{dz}\right)^2 + \frac{d^2x}{dz^2}\cdot\frac{dy}{dz}\cdot\frac{dx}{dz}\right)}_{\epsilon_2} \quad \ldots\ldots\ldots (30)$$

which is of the same form as (24a) except that the terms "ϵ_1" and "ϵ_2" are present on the R.H.S. These cause the pincushion distortion and other geometrical distortions.

The analytic solution of the above presents a problem. However knowing that "ϵ_1" and "ϵ_2" when integrated, provide only a few percent of the total solution (that being the order of pincushion and geometric distortion) the following approximations can be made and substituted into the above

$$\left.\begin{aligned} \frac{d^2y}{dz^2} &= -\sqrt{\frac{e}{2mV}}\cdot B_x, \quad \frac{dy}{dz} = -\sqrt{\frac{e}{2mV}}\cdot B_x\cdot z \\ \frac{d^2x}{dz^2} &= \sqrt{\frac{e}{2mV}}\cdot B_y, \quad \frac{dx}{dz} = \sqrt{\frac{e}{2mV}}\cdot B_y\cdot z \end{aligned}\right\} \quad \ldots\ldots\ldots (31)$$

This results in

$$\frac{d^2y}{dz^2} = -\sqrt{\frac{e}{2mV}}\cdot B_x - \left(\sqrt{\frac{e}{2mV}}\right)^3\cdot\tfrac{3}{2}B_x\cdot\left(B_x^2 + B_y^2\right)z^2$$

$$-\left(\sqrt{\frac{e}{2mV}}\right)^5\cdot\tfrac{9}{4}B_x\left(B_x^2 + B_y^2\right)^2 z^4 \quad \ldots\ldots\ldots (32)$$

Again for the usual conditions of constant deflecting field within the interal $z_0 = o$ and $z_1 = L$, and the electron beam projected onto a plane screen at $z = S + \frac{L}{2}$ and using

$$y = y\Big]_{z=L} + \frac{dy}{dz}\Big]_{z=L}\left(S - \frac{L}{2}\right) \quad \ldots\ldots\ldots\ldots\ldots (26)$$

gives

$$y = -\sqrt{\frac{e}{2mV}} \cdot B_x \cdot L \cdot S - \tfrac{1}{2}\left(\sqrt{\frac{e}{2mV}}\right)^3 B_x\,(B_x^{\,2} + B_y^{\,2})\,L^3 \cdot \left(S - \frac{L}{4}\right)$$

$$- \tfrac{3}{8}\left(\sqrt{\frac{e}{2mV}}\right)^5 B_x\,(B_x^{\,2} + B_y^{\,2})^2 L^5 \cdot \tfrac{6}{5}\left(S - \frac{L}{3}\right) \cdot (33)$$

with a similar expression for the x-deflection.

This expression must be compared with an expression obtained from a geometric method such as given in reference [2].

When the appropriate constants are inserted into that expression to be anlogous to expression (32) it becomes

$$y = -\sqrt{\frac{e}{2mV}} \cdot B_x \cdot L \cdot S - \tfrac{1}{2}\left(\sqrt{\frac{e}{2mV}}\right)^3 B_x\,(B_x^{\,2} + B_y^{\,2})\,L^3 \cdot S$$

$$- \tfrac{3}{8}\left(\sqrt{\frac{e}{2mV}}\right)^3 B_x\,(B_x^{\,2} + B_y^{\,2})^2\,L^5 \cdot S \quad \ldots\ldots (34)$$

Comparing these two expressions it can be seen that, as expected, the deflection derived in this paper is less than that obtained using the geometrical method; this is in accord with the fact that the "center of deflection", nominally at $z = \frac{L}{2}$, moves towards the direction of the screen as the deflection is increased. Quite clearly from the two expressions, the further the distance of the screen from the deflecting region, the closer do the two expressions become as then $\left(S - \frac{L}{4}\right) \to S$ etc.

If a periodic sawtooth scanning current "I_x" is fed into the y-deflection coils to produce the linear deflecting magnetic field "B_x", then with no saturation effects being present,

$$B_x = k_x I_x \quad \cdots\cdots\cdots\cdots (35a)$$

$$B_y = k_y I_y \quad \cdots\cdots\cdots\cdots (35b)$$

Calling $\sqrt{\frac{e}{2mV}} \cdot L = C$

And substituting into (32) gives

$$y = -c.k_x I_x . S - \tfrac{1}{2} c^3.k_x{}^3 I_x \left(I_x{}^2 + \left(\frac{k_y}{k_x}\right)^2 I_y{}^2\right)\left(S - \frac{L}{4}\right)$$

$$- \tfrac{3}{8} c^5.k_x{}^5 I_x \left(I_x{}^2 + \left(\frac{k_y}{k_x}\right)^2 I_y{}^2\right)^2 \tfrac{6}{5}\left(S - \frac{L}{3}\right) \quad \cdots\cdots (36)$$

indicating that if a current of the form

$$I_x' = I_x\left(1 - \tfrac{1}{2}c^2 k_x{}^2 \left(1 - \tfrac{1}{4}.\frac{L}{S}\right)\left(I_x{}^2 + \left(\frac{k_y}{k_x}\right)^2 I_y\right)^2\right) \quad \cdots\cdots\cdots (37)$$

is fed into the y-deflection coils, a deflection linear to the 2nd degree (of the order of 0·1%) is obtained. More complex current wave-forms, containing the higher degree terms, will of course give better linearity. It must be noted that using more exact substitutions for the terms in (31) and more terms in the expansion of $\left\{\left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2 + 1\right\}^{\frac{3}{2}}$ merely contribute higher order terms in "z" and "B_x" and "B_y" in the expressions for x- and y-deflection; however their contribution is negligible compared with the "ϵ_1" and "ϵ_2" terms.

4. NON-FLAT SCREENS

For screens which are non-flat, or as is the case in most TV-mode display equipment, those which consist of a section of a sphere, the y-deflection expression becomes

$$y = y\Big]_{z=L} + \frac{dy}{dz}\Big]_{z=L}\left(S - \frac{L}{2} - \Delta S(x,y,z)\right) \quad \ldots\ldots\ldots\ldots (26a)$$

where "$\Delta S(x,y,z)$" is the normal distance between the actual screen and a flat-screen at the point where the deflected beam falls; it is easily obtained knowing the geometry of the screen with reference to the origin "O"(o,o,o) i.e. the point of entry of an undeflected beam into the deflection region.

Usually "R_s", the radius of curvature of the screen, is such that

$$R_s \geq 2 \rightarrow 3(S)$$

Figure 3 shows the geometry of the screen (only in 2 dimensions) with respect to the deflection region.

For a sphere of radius "R_s",

$$x^2 + y^2 + z^2 = R_s^{\,2}$$

For this case, taking $z = \frac{L}{2}$ as the centre of deflection as usual,

$$z^2 = (R_s - \Delta S(x,y,z))^2 \quad \ldots\ldots\ldots\ldots\ldots (38)$$

Hence "solving" for ΔS gives approximately

$$\Delta S(x,y,z) \simeq \frac{x^2}{2R_s - \Delta S} + \frac{y^2}{2R_s - \Delta S}$$

To evaluate this the approximations obtained from (26) are used.

These give

$$x \simeq kB_y\frac{L^2}{2} + kB_y.L.\left(S - \frac{L}{2}\right) \quad \ldots\ldots\ldots\ldots\ldots (39a)$$

$$y \simeq -kB_x.\frac{L^2}{2} - kB_x.L.\left(S - \frac{L}{2}\right) \quad \ldots\ldots\ldots\ldots\ldots (39b)$$

Taking only the first two significant terms of $\left.\frac{dy}{dz}\right]_{z=L}$

and of "ΔS" in (26a) result in

$$y = -kB_x.\frac{L^2}{2} - k^3.\frac{L^4}{8} . B_x(B_x^2 + B_y^2)$$

$$- \left(kB_x.L + k^3.\frac{L^3}{2}.B_x(B_x^2+B_y^2)\right)\left\{S - \frac{L}{2} - k^2.L^2 \frac{\left(S-\frac{L}{2}\right)^2}{2R_S - \Delta S} . B_x(B_x^2+B_y^2)\right.$$

$$\left. - k^3.L^3. \frac{\left(S-\frac{L}{2}\right)}{2R_S - \Delta S} B_x(B_x^2 + B_y^2)\right\} \quad \ldots\ldots\ldots\ldots\ldots (40)$$

On collecting common terms, this reduces to

$$y = -kB_x . L . S - k^3.B_x(B_x^2+B_y^2)\frac{L^3}{2}.S.\left[1 - \frac{S}{R_S - \frac{\Delta S}{2}} - \frac{L}{4S}\right] \quad \ldots\ldots\ldots (40a)$$

Studying this expression it can be seen that

(1) it is identical in form to expression (33) giving the y-deflection on a flat-faced screen, except that the 2nd order term is reduced by a factor

$$\frac{\left(1 - \frac{S}{R_S - \frac{\Delta S}{2}} - \frac{L}{4S}\right)}{\left(1 - \frac{L}{4S}\right)} \quad \ldots\ldots\ldots\ldots\ldots (41)$$

implying that the pincushion distortion is decreased - as is expected.

(2) the distortion term is a function of "$\Delta S(x,y,z)$", and hence of the deflection; the greater the deflection the greater is the distortion.

Two special cases are of interest.

(1) When $R_s \to \infty$, the screen becomes flat-faced and this expression reduces exactly to the one obtained previously.

(2) When $R_s = S$, the screen has its center at the center of deflection. This case is usually quoted in the literature as the one which gives a linear deflection for a linear input deflection current. It has already been indicated previously that the movement of the center of deflection precludes this.

Substituting $R = S$ into (40a), and for the hypothetical tube with a 90° deflection angle across the diagonal, the factor (41) is approximately 0·2, implying that the pincushion and geometric distortion although not absent has been decreased by a factor of 5 over the flat-faced screen tube.

The necessary deflection coil current for linear deflection can be obtained by an expression analogous to (3).

5. REMARKS AND CONCLUSION

It is instructive to go over the assumptions which have been made in deriving the above equations. These include:-

(1) a deflecting magnetic field distribution along the z-axis such that within the interval $z = 0$ to $z = L$, the flux density is constant, and zero outside this interval.
In practice this is not achieved, but the equations are of such a form that if the actual non-ideal magnetic field distribution is known along the z-axis, the deflections can still be obtained.

(2) even if the magnetic field distribution along the z-axis is known, for large deflections, the field distribution off-axis may be different, particularly due to fringing effects at the extremities of the deflecting region.

The effects due to this however are much smaller than the above; they are the "3rd order abberations".

(3) the approximations in deriving the final expressions for deflections, namely the approximations in (31), and in expanding $\left\{ \left(\frac{dx}{dz}\right)^2 + \left(\frac{dy}{dz}\right)^2 + 1 \right\}^{\frac{3}{2}}$ and in "$\Delta S(x,y,z)$", only contribute to higher order terms which may be neglected.

(4) the equations hold only for "mono-chromatic electron beams" - i.e. single velocity electrons.

Numerical comparisons of actual TV tube distortions with calculated ones using the above expressions are difficult to make. Not only do the above points need be taken into account but the deflection coils are so wound and input currents may be pre-shaped so as to partially or (hopefully) wholly compensate for the pincushion distortion. However with "simple" coil design observing the relationship

$$B_x = k_x I_x ,$$

terms corresponding to "ϵ_1" and "ϵ_2" of (33) and (40a) can be quite easily generated with a network of adders and multipliers to eliminate distortion. In addition these networks can be used to generate some additional deflection functions.

REFERENCES

1. Eds. T. Soller, M.A. Starr, G.E. Valley, "Cathode Ray Tube Displays", pp. 303-305, 338-340, Vol.22, Radiation Lab. Series, McGraw Hill: New York, 1948.

2. B.F. Wadsworth "Deflection Non-Linearity and Distortion in Pepr's CRT Display", Pepr Engineering Note No.19, Aug. 1963 Lab. for Nuclear Science, M.I.T.

3. P.A. Sturrock, "Static and Dynamic Electron Optics", pp. 6-8 Cambridge University Press, 1955.

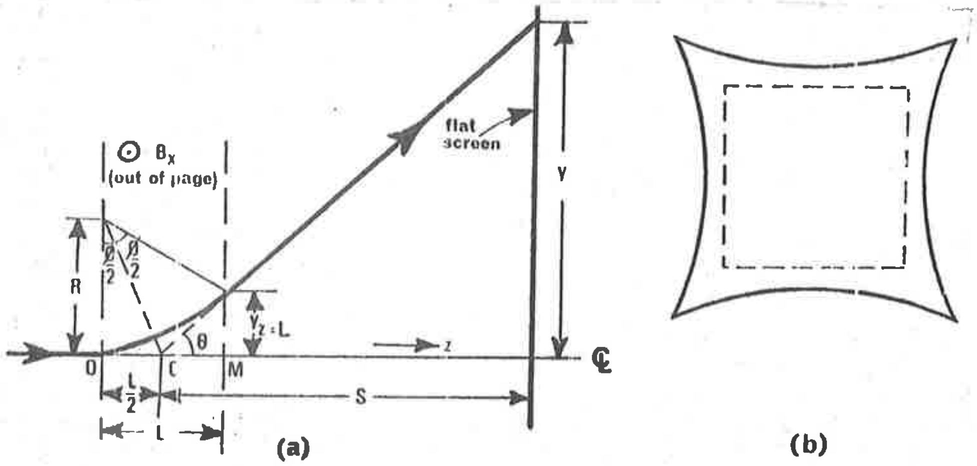


Figure 1. Geometry of the deflected electron beam (arrowed) when passing through a deflecting magnetic field B_x present in the region of length "L". (b) is the resultant Pincushion Distortion (exaggerated) of a square raster shown dotted.

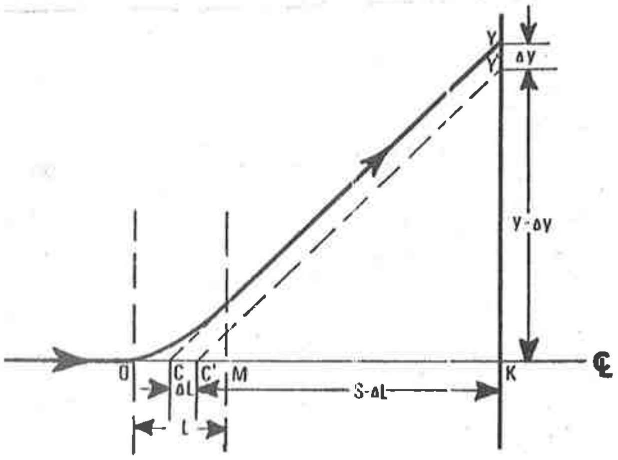


Figure 2. Effect on deflection when the " Centre of Deflection , C " is shifted by some " δL ".

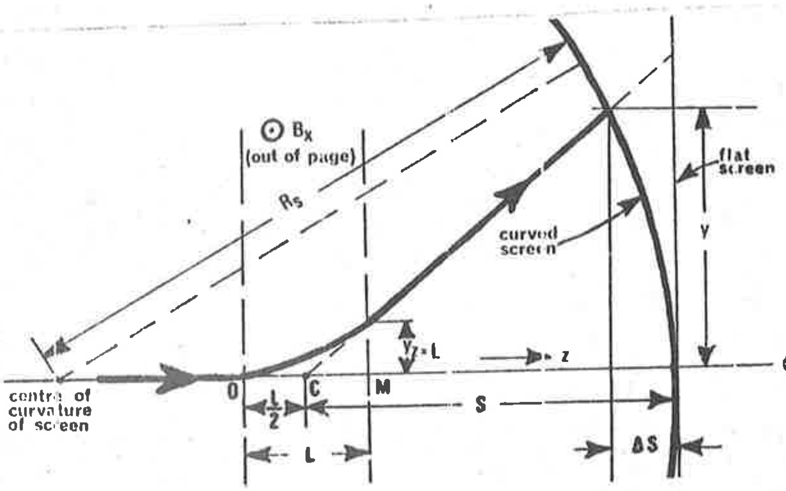


Figure 3. Geometry of the deflected electron beam (arrowed) when deflected onto a curved screen.

4 DERIVATION OF VIDICON BEAM DEFLECTION AND DISTORTION

The force $\underset{\sim}{F}$ acting on an electron moving at a velocity $\underset{\sim}{v}$ is given by the well-known Lorenz equation:

$$\underset{\sim}{F} = m\frac{dv}{dt} = -e(\underset{\sim}{E} + \underset{\sim}{v} \times \underset{\sim}{B}) \quad \ldots\ldots\ldots\ldots\ldots (1)$$

where "$\underset{\sim}{F}$" is the force

"m, e" are the mass and charge of the electron

"$\underset{\sim}{E}$" is the electric field

"$\underset{\sim}{B}$" is the magnetic field

Also, "$e\underset{\sim}{E}$" produces acceleration and hence work and thus a change in energy, and $e(\underset{\sim}{v} \times \underset{\sim}{B})$ is always perpendicular to the direction of motion of electrons. Hence no work is done, no change in energy occurs, only a change in direction of motion.

In Cartesian coordinates the above expression becomes

$$m\ddot{x} = -e.E_x - e\dot{y}.B_z + e\dot{z}.B_y \quad \ldots\ldots\ldots\ldots\ldots (2)$$

$$m\ddot{y} = -e.E_y - e\dot{z}.B_x + e\dot{x}.B_z \quad \ldots\ldots\ldots\ldots\ldots (3)$$

$$m\ddot{z} = -e.E_z - e\dot{x}.B_y + e\dot{y}.B_x \quad \ldots\ldots\ldots\ldots\ldots (4)$$

where E_x, E_y, E_z are the electrical field components

B_x, B_y, B_z are the magnetic field components.

The determination of the "x" and "y" deflections in the Vidicon is based on the solution of the above expressions.

Assumptions:

(i) No electric field $\underset{\sim}{E}$ exists within the deflecting field region,

i.e. $E_x = E_y = E_z = 0$

(ii) The initial velocity of the electrons are given by

$$v_{oz} = \sqrt{2\frac{m}{e}V_A} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots (5)$$

where "V_A" is the accelerating anode potential ($\approx$ 300V) and the velocity is along the axis in the z-direction (this is derived from expression (5)).

Within the deflecting region of the Vidicon three magnetic fields exist:

(1) the axial focussing magnetic field $\underset{\sim}{B}_z$.

(2) the magnetic field in the x-direction, $\underset{\sim}{B}_x$.

(3) the magnetic field in the y-direction, $\underset{\sim}{B}_y$.

As "B_z" is constant, while during the trajectory transit time, "B_x" and "B_y" can be considered to be quasi-static, the solution of the above expressions will be solved for x and y, with the time variable entering within the expressions by the transformation,

$$v_z \cdot t = z \quad \ldots\ldots\ldots\ldots\ldots\ldots \quad (6)$$

The equations requiring solution are then:

$$m\ddot{x} = -e\dot{y}\,B_z + e\dot{z}\,B_y \quad \ldots\ldots\ldots\ldots \quad (2a)$$

$$m\ddot{y} = -e\dot{z}\,B_x + e\dot{x}\,B_z \quad \ldots\ldots\ldots\ldots \quad (3a)$$

$$m\ddot{z} = -e\dot{x}\,B_y + e\dot{y}\,B_x \quad \ldots\ldots\ldots\ldots \quad (4a)$$

Taking Laplace Transforms with initial condition terms gives

$$X(s) \cdot S^2 - S.x_o - \dot{x}_o = -\frac{e}{m}\left(S.Y(s) - y_o\right) B_z + \frac{e}{m}\left(S.Z(s) - z_o\right) B_y \quad .. \; (2b)$$

$$Y(s) \cdot S^2 - S.y_o - \dot{y}_o = -\frac{e}{m}\left(S.Z(s) - z_o\right) B_x + \frac{e}{m}\left(S.X(s) - x_o\right) B_z \quad .. \; (3b)$$

$$Z(s) \cdot S^2 - S.z_o - \dot{z}_o = -\frac{e}{m}\left(S.X(s) - x_o\right) B_y + \frac{e}{m}\left(S.Y(s) - y_o\right) B_x \quad .. \; (4b)$$

The initial conditions are

$$\left.\begin{array}{l} \dot{x}_o = 0,\, x_o = 0 \\ \dot{y}_o = 0,\, y_o = 0 \end{array}\right\}$$ as the beam enters the magnetic field region axially

and $\dot{z}_o = v_{oz} = \sqrt{\dfrac{2eV_A}{m}},\, z_o = 0.$

The origin is thus made along the tube axis (aligned with the scanned area centre), starting at the deflecting field region.

Substituting for these initial conditions and collecting terms results in:

$$S^2.X(s) + \frac{e}{m}.B_z.S.Y(s) - \frac{e}{m}.B_y.S.Z(s) = 0 \quad \ldots\ldots\ldots\ldots (7)$$

$$-\frac{e}{m}B_z.S.X(s) + S^2.Y(s) + \frac{e}{m}.B_x.S.Z(s) = 0 \quad \ldots\ldots\ldots\ldots (8)$$

$$\frac{e}{m}.B_y.S.X(s) - \frac{e}{m}.B_x.S.Y(s) + S^2.Z(s) = v_{oz} \quad \ldots\ldots\ldots\ldots (9)$$

or

$$\begin{vmatrix} S^2 & \frac{e}{m}B_z S & -\frac{e}{m}B_y.S \\ -\frac{e}{m}B_z S & S^2 & \frac{e}{m}B_x.S \\ \frac{e}{m}B_y S & -\frac{e}{m}B_x S & S^2 \end{vmatrix} \cdot \begin{vmatrix} X(s) \\ Y(s) \\ Z(s) \end{vmatrix} = \begin{vmatrix} 0 \\ 0 \\ v_{oz} \end{vmatrix} \quad \ldots (10)$$

Solving for X(s) gives

$$X(s) = \frac{v_o\frac{e}{m}.S^2\left(S.B_y + \frac{e}{m}.B_x.B_z\right)}{S^2\left(S^2 + \left(\frac{e}{m}\right)^2\left(B_x^2 + B_y^2 + B_z^2\right)\right)}$$

$$= \frac{v_o\left(\frac{e}{m}\right)^2 B_x B_z}{S^2\left[S^2 + \left(\frac{e}{m}\right)^2 \sum_i^{x,y,z} B_i^2\right]} + \frac{v_o\frac{e}{m}.B_y}{S\left[S^2 + \left(\frac{e}{m}\right)^2 \sum_i^{x,y,z} B_i^2\right]} \quad \ldots (11)$$

as $\mathcal{L}^{-1}\left(\frac{1}{S^2(S^2 + a^2)}\right) = \left(\frac{t}{a^2} - \frac{\sin(at)}{a^3}\right)$

and $\mathcal{L}^{-1}\left(\frac{1}{S(S^2 + a^2)}\right) = \left(\frac{1}{a^2} - \frac{\cos(at)}{a^2}\right)$

then

$$x(t) = \frac{v_o \cdot \left(\frac{e}{m}\right)^2 B_x \cdot B_z \cdot t}{\left(\frac{e}{m}\right)^2 \sum_i^{x,y,z} B_i^2} - \frac{v_o \cdot \left(\frac{e}{m}\right)^2 B_x \cdot B_z \cdot \sin\left(\left(\frac{e}{m}\right)\sqrt{\sum_i^{x,y,z} B_i^2} \cdot t\right)}{\left(\frac{e}{m}\right)^3 \left(\sum_i^{x,y,z} B_i^2\right)^{\frac{3}{2}}}$$

$$+ \frac{v_o \cdot \left(\frac{e}{m}\right) \cdot B_y}{\left(\frac{e}{m}\right)^2 \sum_i^{x,y,z} B_i^2} - \frac{v_o \cdot \left(\frac{e}{m}\right) \cdot B_y}{\left(\frac{e}{m}\right)^2 \sum_i^{x,y,z} B_i^2} \cos\left(\frac{e}{m}\sqrt{\sum_i^{x,y,z} B_i^2 \cdot t}\right) \quad \ldots (11a)$$

and as $v_o t = Z$ and simplifying

$$x(t) = Z \cdot \frac{B_x}{B_z\left(1 + \left(\frac{B_x}{B_z}\right)^2 + \left(\frac{B_y}{B_z}\right)^2\right)}$$

$$- \frac{v_o \cdot B_x \cdot \sin\left(\frac{e}{m} B_z \left(1 + \frac{1}{2}\left(\frac{B_x}{B_z}\right)^2 + \left(\frac{B_y}{B_z}\right)^2\right)^{\frac{1}{2}} \cdot t\right)}{\frac{e}{m} \cdot B_z^2 \cdot \left(1 + \left(\frac{B_x}{B_z}\right)^2 + \left(\frac{B_y}{B_z}\right)^2\right)^{\frac{3}{2}}}$$

$$+ \frac{v_o \; B_y}{\frac{e}{m} \cdot B_z^2 \cdot \left(1 + \left(\frac{B_x}{B_z}\right)^2 + \left(\frac{B_y}{B_z}\right)^2\right)}$$

$$- \frac{v_o \cdot B_y \cdot \cos\left(\frac{e}{m} B_z \left(1 + \frac{1}{2}\left(\frac{B_x}{B_z}\right)^2 + \frac{1}{2}\left(\frac{B_y}{B_z}\right)^2\right)^{\frac{1}{2}} \cdot t\right)}{\frac{e}{m} B_z^2 \left(1 + \left(\frac{B_x}{B_z}\right)^2 + \left(\frac{B_y}{B_z}\right)^2\right)} \quad \ldots (11b)$$

A similar expression holds for "$y(t)$", except that "B_x" and "B_y" are interchanged.

Certain additional substitutions and approximations can be made:

(1) Since it is known from the simple case that electrons follow the magnetic field lines (Section 9.3), then from Figure 126 the x-deflection is given by

$$\frac{x}{z} = \frac{B_x}{B_z} = \tan\varphi \quad \ldots\ldots\ldots\ldots \quad (12)$$

where "φ" is the deflection. Since $\varphi_{max} \approx 7^0$ for a Vidicon then the assumption that

$$B_x, B_y \ll B_z$$

and more so,

$$B_x^2, B_y^2 \ll B_z^2 \qquad \text{holds}$$

(2) From the usual expression of equating the force due to a magnetic field on an electron, with the resulting centripetal force,

$$B\,ev = \frac{mv^2}{R}$$

the equivalence can be made,

$$B_z \frac{e}{m} \cdot \frac{1}{v_{oz}} = \frac{1}{R} \quad \ldots\ldots\ldots\ldots \quad (13)$$

where "R" is the radius of the spiralling electrons.

In Vidicon, "B_z" is usually of the order of 40-50 gauss (say, 40 gauss = 40.10^{-4} Webers/m²).

While $v_{oz} = \sqrt{\frac{2e}{m}.\, V_A} = \sqrt{2.1\cdot76.10^{11}.300} = 1\cdot03.10^7$ metre/sec.

Hence $R = \dfrac{1\cdot03.10^7}{1\cdot76.10^{11}\ 4.10^{-3}} = 1\cdot46$ cm.

(3) As "x" and "y" are measured on the photoconductor target, then the time at which "x(t)" and "y(t)" are evaluated at, in expression (11b), is obtained from

$$Z(t) = v_{oz} \cdot t \qquad \ldots\ldots\ldots\ldots (14)$$

or

$$T_s = \frac{Z_T}{v_{oz}} \qquad \ldots\ldots\ldots\ldots (14a)$$

where "Z_T" is the axial distance of the target from the beginning of the magnetic deflecting region. In Vidicons it is of the order of 8-10 cm,

and "T_s" is the transit time for "Z_T".

As a first approximation then, with the above factors (1), (2), (3) taken into account, expression (11b) reduces to:

$$x(t) = v_o . \tan \varphi_x . t - R . \tan \varphi_x . \sin\left(\frac{v_{oz} . t}{R}\right) + R \frac{B_y}{B_z}$$

$$- R . \frac{B_y}{B_z} . \cos\left(\frac{v_{oz} . t}{R}\right) \qquad \ldots\ldots\ldots\ldots (11c)$$

and at the target where $v_{oz} . T_s = Z_T$

$$x(T_s) = Z_T . \tan \varphi_x \left(1 - \frac{R}{Z_T} \sin\left(\frac{Z_T}{R}\right)\right) + R \frac{B_y}{B_z}\left(1 - \cos\left(\frac{Z_T}{R}\right)\right) \quad \ldots (11d)$$

Similarly

$$y(t) = v_o . \tan \varphi_y . t - R \tan \varphi_y \sin\left(\frac{v_{oz} . t}{R}\right) + R \frac{B_x}{B_z}$$

$$- R . \frac{B_x}{B_z} . \cos\left(\frac{v_{oz} . t}{R}\right) \qquad \ldots\ldots\ldots\ldots (11e)$$

also $Z(t) = v_o . t$

As the argument of the "sin" and "cos" terms is a function of "t," and hence a function of distance along the z-axis, the "x" and "y" coordinates describe cyclic or spiral motions along a path in the direction of "$v_0 \tan \varphi_x . t$" i.e. "$z \tan \varphi_x$" or "$z \tan \varphi_y$" i.e. in the directions of the magnetic lines of force of the resultant "x" and "y" direction deflecting fields. These are the spiral or helical motion of electrons qualitatively described in textbooks.

When the region is reached where "$B_x = B_y = 0$," a new set of equations, similar to (2b), (3b), (4b) need be solved with "B_x" and "B_y" set to zero, and new initial conditions set which are the terminal conditions of the above case. This can only be done if magnetic field data and deflecting coil dimensions are known accurately. Here, it is assumed that the above deflections are operative at the target, i.e. the magnetic fields "B_x" and "B_y" are present at the target surface. The approximation does not introduce much error; whatever errors are introduced and not evident in the resulting expressions, can be covered by the remanent distortion correction described in Chapter 8.

It is noticed that in Vidicons, the deflection of the beam is directly in the direction of the deflecting field. Thus a magnetic field directed say from the top of the page to the bottom produce a downward vertical beam deflection. (See figure 126(a)).

In CRTs, on the other hand, such a directed magnetic field will produce horizontal beam deflection, out of the page, towards the reader.

Consequently corresponding H- and V-deflection coils are rotated by 90° in the Vidicon when compared with the CRT coils alignment.

Another point to notice is that rotation of the scanned area occurs because of the "sin" terms (and to a lesser degree due to the "cos" term).

From the form of the expression (11d) it is seen that for all "tan φ_x" and hence "B_x" (from expression (12)), the "sin" terms are constant. Hence all points on the raster are rotated by the same amount. However the term involving the "cos" term depends on "B_y", which may vary for a constant "B_x", implying that locations with the nominally identical "x" coordinate have slight asymmetrical rotation about the origin due to the "cos" term, as "B_y" varies from $\pm B_{y\,max}$. Small differential raster rotation thus is present in addition to the constant rotation. (This is analogous to the spiral magnetic distortion mentioned in Section 9.3.7.1).

To eliminate this constant and differential rotation, the "sin" and "cos" terms need be eliminated. This is achieved if the arguments of these terms is set to 2π, 4π, ... (setting the argument to zero is precluded as "t" needs to be finite).

In this case then, expression (11d) reduces to

$$x(t) = v_o(t)\frac{B_x}{B_z} + R\frac{B_y}{B_z}(1 - \cos 2\pi)$$

$$= v_o(t)\frac{B_x}{B_z} \quad \ldots\ldots\ldots\ldots\ldots\ldots (11f)$$

which is the required condition as then $x \propto B_x$.

In the Vidicons, the argument is set to "2π" by appropriately selecting the focussing magnetic field "B_z", and the accelerating anode voltage "V_A", to give a helix radius (expression (13)) with the dimensions of the tube, specifically "Z_T", the distance between the origin of the magnetic deflection region and the photoconductor target.

It is thus required that "B_z", "V_A" and "Z_T" are chosen such that

$$\frac{Z_T}{R} = 2\pi \qquad \text{(from (11d))}$$

"R" was calculated to be 1·46 cm, thus requiring Z_T = 9·2 cm, which correspond to the usual tube dimension "Z_T", usually being between 8-10 cm. Thus the focussing field is adjusted until one complete spiral or "pitch distance" of the helical motion of electrons results. (Adjusting this focussing field, one can see the image being rotated when observing the result on a CRT.)

The focussing and deflecting magnetic fields in the Vidicon are thus intimately linked.

What holds for the approximate case, as given by (11c), (11d) etc., also holds for the more detailed expression (11b).

Rewriting expression (11b) and substituting, using the expressions (13) and (14), and evaluating at

$$t = T_s = \frac{Z_T}{v_{oz}} \qquad \text{i.e. on the photoconductor target}$$

then

$$x(T_s) = Z_T \cdot \frac{B_x}{B_z} \cdot \frac{1}{\left(1 + \left(\frac{B_x}{B_z}\right)^2 + \left(\frac{B_y}{B_z}\right)^2\right)} - R \cdot \frac{B_x}{B_z} \sin\left(\frac{Z_T}{R}\left(1 + \tfrac{1}{2}\left(\frac{B_x}{B_z}\right)^2 + \tfrac{1}{2}\left(\frac{B_y}{B_z}\right)^2\right)\right)$$

$$+ R \cdot \frac{B_y}{B_z} \cdot \frac{1 - \cos\left(\frac{Z_T}{R}\left(1 + \tfrac{1}{2}\left(\frac{B_x}{B_z}\right)^2 + \tfrac{1}{2}\left(\frac{B_y}{B_z}\right)^2\right)\right)}{\left(1 + \left(\frac{B_x}{B_z}\right)^2 + \left(\frac{B_y}{B_z}\right)^2\right)} \qquad \ldots\ldots(15)$$

Expanding the "sin" and "cos" term and substituting

$$\frac{Z_T}{R} = 2\pi$$

and noting that the first order deviations due to $\left(\left(\frac{B_x}{B_z}\right)^2 + \left(\frac{B_y}{B_z}\right)^2\right)$

result in small remanent arguments in the "sin" and "cos" terms, the approximations

$$\sin\theta = \theta \quad \text{and} \quad \cos\theta = 1 - \frac{\theta^2}{2} \quad , \theta << 1$$

can be made. For the terms in the denominators in "$\frac{B_x}{B_z}$", "$\frac{B_y}{B_z}$", the approximations that "$\frac{1}{1+k} = 1 - k$" for $k << 1$ can be made for, as noted previously,

$$B_x^2, B_y^2 << B_z^2$$

The expression (15) then reduces to

$$x(T_s) = Z_T \cdot \frac{B_x}{B_z} \cdot \left(1 - \left(\frac{B_x}{B_z}\right)^2 - \left(\frac{B_y}{B_z}\right)^2\right) - R \cdot \frac{B_x}{B_z} \cdot \frac{Z_T}{R} \cdot \left(1 + \tfrac{1}{2}\left(\frac{B_x}{B_z}\right)^2 + \tfrac{1}{2}\left(\frac{B_y}{B_z}\right)^2\right)$$

$$\left(+ R \cdot \frac{B_y}{B_z} \cdot \frac{Z_T}{R} \cdot \frac{\left(\tfrac{1}{2}\left(\frac{B_x}{B_z}\right)^2 + \tfrac{1}{2}\left(\frac{B_y}{B_z}\right)^2\right)^2}{\left(1 + \left(\frac{B_x}{B_z}\right)^2 + \left(\frac{B_y}{B_z}\right)^2\right)} \text{, which can be neglected}\right)$$

$$= \left(\frac{Z_T}{B_z}\right) \cdot B_x - \frac{3}{2}\left(\frac{Z_T}{B_z^{\,3}}\right) \cdot B_x \cdot (B_x^{\,2} + B_y^{\,2}) \quad \ldots\ldots\ldots\ldots (15a)$$

Similarly

$$y(T_s) = \left(\frac{Z_T}{B_z}\right) \cdot B_y - \frac{3}{2}\left(\frac{Z_T}{B_z^{\,3}}\right) \cdot B_y \cdot (B_x^{\,2} + B_y^{\,2}) \quad \ldots\ldots\ldots\ldots (15b)$$

<u>These are the x and y deflection terms expressed in terms of deflection magnetic fields.</u>

The bracketed terms can be neglected.

Since $\frac{B_x}{B_z} \approx \tan\varphi \approx 7^{\circ}$ maximum

then $\frac{B_x}{B_z} \approx \frac{B_y}{B_z} \approx 0{\cdot}12$

Hence the ratio of the bracketed term to the remanent terms in "B_x'', "B_y'' is

$$\approx \frac{\frac{B_y}{B_z} \cdot Z_T \cdot \frac{1}{4}\left(\left(\frac{B_x}{B_z}\right)^2 + \left(\frac{B_y}{B_z}\right)^2\right)^2}{\frac{3}{2} \cdot Z_T \cdot \frac{B_x}{B_z} \cdot \left(\left(\frac{B_x}{B_z}\right)^2 + \left(\frac{B_y}{B_z}\right)^2\right)}$$

$$\approx \frac{1}{6}\frac{B_y}{B_x} \cdot \left(\left(\frac{B_x}{B_z}\right)^2 + \left(\frac{B_y}{B_z}\right)^2\right)$$

$$\approx \frac{1}{3} \cdot (0{\cdot}12)^2$$

which is less than 1% of the retained distortion term.

APPENDIX 7

BEAM DISCHARGE IN VIDICONS

1 THE ELECTRON SCANNING BEAM

2 SCANNED AREA EDGE FLICKER

3 INTERVAL BETWEEN ERROR SIGNALS

APPENDIX 7

BEAM DISCHARGE IN VIDICONS

A.7.1 NATURE OF ELECTRON SCANNING BEAM

The axial velocity of electrons in the electron beam (i.e. the velocity in the direction of the beam) is not constant but has a velocity distribution determined by the thermal emission velocity from the emitting cathode(251,252,253).

When the electron beam is accelerated and then decelerated by various potentials and then is incident on some target surface, (such as the photoconductor), the beam electronswill be incident so long as at the surface, they have finite axial velocity. From the energy relationship between electric field energy and the kinetic energy

$$eV_A = \tfrac{1}{2} m v^2$$

where "e", "m" are the charge and mass of the electron

"V_A" is the net potential accelerating the electrons resulting in a velocity "v".

If "V_K" is the emission cathode potential

"V_T" is the final target potential

then clearly for electrons reaching the target, the velocity v > o, and thus

$$V = (V_T - V_K) > 0$$

or

$$V_T > V_K$$

Clearly the limiting case for beam electrons landing is when

$$V_T = V_K$$

as then the axial velocity, $v = 0$.

For an insulating target surface, at some initial $V_T > V_K$ electrons will land, lowering the potential of the target until $V_T = V_K$. Thus an electron beam "stabilizes" a target to its own emission cathode potential.

During that interval, all the beam electrons land, and thus

$$i_T = i_b \quad , \quad V_T > V_K$$

i.e. the charge current within the target equals the beam current.
This is called the "normal range".
This of course assumes the secondary emission from the target is negligible; this is true for slow incident electrons i.e. where $V_T > V_K$ by several volts.

However due to the initial distributions of electron velocities when emitted at the cathode (initially assumed to have had zero velocity), some beam electrons have sufficient axial velocity to be incident on the target even if $V_T \leqslant V_K$. The target current "i_T" under these conditions is given by

$$i_T = i_b \exp\left(\frac{e(V_T - V_K)}{kT}\right) \quad , \quad V_T < V_K$$

where "e" is the electron charge ($1\cdot6.10^{-19}$ coloumbs)

"k" is Boltzmann's constant ($1\cdot38.10^{-23}$ Joules $^{o}K^{-1}$)

"T" is the effective cathode temperature (usually taken about $1100^{o}K$)

This is called the "intrinsic current range".
Thus when evaluating,

$$i_T = i_b \exp\left(10\cdot5\,(V_T - V_K)\right) \quad , V_T < V_K$$

$$= i_b \exp\left(a(V_T - V_K)\right)$$

In practice $V_K = 0$, and a "contact potential" exists between the beam electrons and the target (in our case, this is equal to +2V). To become carriers within the photoconductor, the electron must have a minimum energy of 2eV (see Section A8.2.3). Consequently the above expressions become

$$i_T = i_b \dots \text{ for } V_T > 2V \qquad \dots\dots\dots\dots \quad 1.a$$

$$i_T = i_b \cdot \exp(a.V_T) \dots \text{ for } V_T < 2V \qquad \dots\dots\dots \quad 2.a$$

It has been found that "a", the "constant" is itself a function of beam current, and decreases with increasing "i_b" and with "V_T" (253) (see for example the curves given therein). "a" was taken as "3.5" in Section 5.2.3.

A.7.2 BEAM IMPEDANCE AND BEAM DISCHARGE TIME CONSTANT

Within the intrinsic region, the target current being a function of "V_T", an impedance can clearly be associated.

Thus
$$\frac{d\, i_T}{d\, V_T} = a.i_b.e^{a.V_T} \qquad \dots\dots\dots\dots \quad A.7.2.1$$

or
$$\frac{\Delta V_T}{\Delta i_T} = R_b = \frac{1}{a.i_T} \qquad \dots\dots\dots\dots \quad A.7.2.2$$

For a = 5-10, and $i_T \approx 0{\cdot}01 - 0{\cdot}05\mu A$ (the range where $V_T \lesssim 2$), $R_b \approx 10^6 - 10^7\,\Omega$.

An equivalent circuit can be drawn for the beam landing on a display element during the intrinsic region which is shown in figure 72(b).

The associated time constant during the beam discharge during the "intrinsic range" is

$$T.C_B' = c_m.R_b = \frac{C}{m}.\frac{1}{a.i_T} \qquad \dots\dots\dots \quad A.7.2.3$$

as
$$c_m = \frac{C}{m}$$

where "c_m" is the capacitance associated with each target element (of the order of 3.10^{-15} F)

"C" is the total photoconductor capacitance (of the order of 1500 pF)

"m" is the total number of display element (of the order of $4.5.10^5$).

The display element resistance "R_d", shown on figure 72(b), is of the order of $10^9\,\Omega$ and hence can be neglected with the series/parallel combination of "c_m" and "R_b".

This is a "time constant" associated with a beam that accesses a location for a dwell time "T_d" in every frame time "T_F",

where

$$m.T_d = T_F$$

For the time constant for "continuous access time", or for the effect which is evident over several dwell times

$$TC_B = m\,TC_B' = C\,\frac{1}{a.i_T} \qquad \text{A.7.2.4}$$

This result has been obtained separately by Kiuchi and Reddington(252) and in a different form by Van Polder who gave the result (253) as

$$T.C_B = \frac{T_F}{a.\,\delta V_T} \qquad \text{A.7.2.5}$$

where "δV_T" is the potential charge due to beam discharge while "T_F" and "a" are as above.

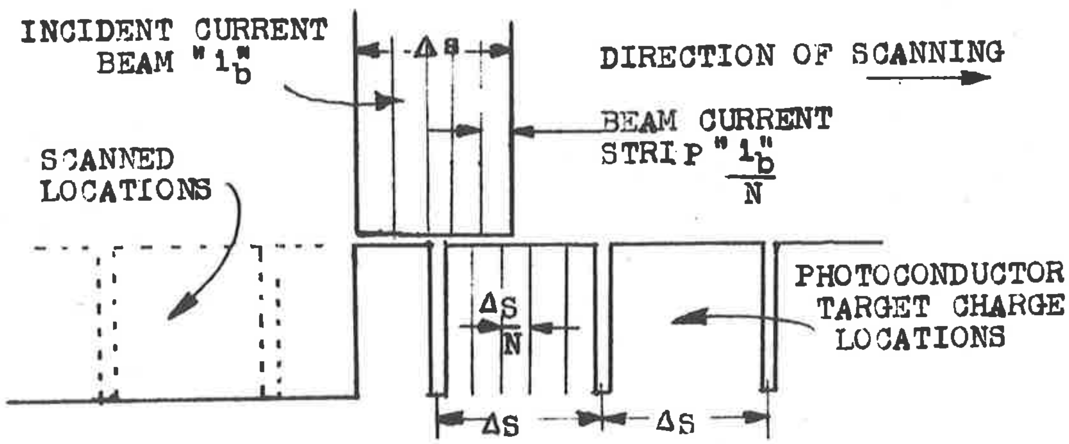
As $$i_s = i_T = \frac{C.\,\delta V_T}{T_F} \quad \text{(from 4.19)}$$

$$\therefore \quad TC_B = \frac{C}{a.i_T} \quad \text{as above.}$$

A.7.3 VALIDITY OF USING A STATIONARY BEAM WITH DWELL TIME "T_d" WITH THE SCANNING BEAM

In the calculation for the output signal due to the electron beam, it was assumed that the beam is incident on any location for the dwell period "T_d", and as if it stepped from location to location in zero time. In fact the beam sweeps across each element with a constant velocity "$\frac{\Delta s}{T_d}$", where "Δs" is the width of an element, and $T_d = \frac{T_F}{m}$.

The beam itself has a finite cross section area at target incidence, of approximately "$\Delta s \times \Delta s$."



If each element is divided into "N" strips each of "$\frac{\Delta s}{N}$" width and similarly the current beam is divided into "N" strips of "$\frac{\Delta s}{N}$" width; with $N \to \infty$ then each current strip is "$\frac{i_b}{N}$" amps.

With each element strip, a capacitance "$\frac{C_m}{N}$" is associated.

With a scanning velocity "$\frac{\Delta s}{T_d}$", the time spent on each element strip is "$\frac{T_d}{N}$". Assume the beam is stepped across the element by "$\frac{\Delta s}{N}$" in zero time and remains stationary for "$\frac{Td}{N}$" between each step. As $N \to \infty$ this approaches continuous scanning.

The charge "q" supplied by each current strip "$\frac{i_b}{N}$" to each element strip "$\frac{\Delta s}{N}$" in time "$\frac{T_d}{N}$" is related by

$$q = i\Delta t = \frac{i_b}{N} \cdot \frac{T_d}{N} \quad \ldots\ldots\ldots\ldots \quad A.7.3.1$$

As the beam scans the element, each element strip is accessed successively by all of the current strips, each supplying "$\frac{i_b}{N} \cdot \frac{T_d}{N}$" charges. Each element strip thus is provided with $N\left(\frac{i_b}{N} \cdot \frac{T_d}{N}\right)$ charge - and for all elements $\left(\frac{i_b}{N} \cdot \frac{T_D}{N}\right)$ charge is supplied which is the same charge supplied by a beam current "i_b" on element "Δs" width in time "T_d".

It is clear that for a given scanning velocity, "$\frac{\Delta s}{T_d}$", the time that the beam interacts between any given element is "$2T_d$" and thus the output pulse signal is spread over "$2T_d$" seconds leading to a limit in resolution. This is the aperture effect of resolution limitation (see (313)). A scanning beam of narrower width ("aperture") but higher current density is clearly desirable. But the reduction in the "a" exponent with increasing beam current, beam spreading due to mutual electron repulsion, focussing limits etc. make the achievement of such a thin strip scanning beam difficult to achieve.

For our purpose here anyway, from beam discharge considerations, the assumption that the beam can be considered stationary on an element for the dwell time "T_d" is valid.

A.7.4 EVALUATION OF BEAM DISCHARGE EXPRESSIONS

For the evaluation of the curves in Fig.73 the beam acceptance curves need be solved, to relate the change in potential due to beam discharge, with the dwell time T_d.

The target current "i_T" is related to the change in charge potential "ΔV_T" by the usual expression

$$-i_T = c_m \cdot \frac{d(\Delta V_T)}{dt} \qquad \ldots\ldots\ldots\ldots \text{A.7.4.1}$$

where "c_m" is the element capacitance $\approx \frac{C}{m}$.

For the "normal range",

$$i_T = i_b, \quad \Delta V_T > 2 \qquad \text{(expression 5.3)}$$

then

$$\Delta V_T(t) = \frac{-i_T \cdot t}{c_m} = \frac{-i_b \cdot t}{c_m}, \quad \Delta V_T > 2 \qquad \text{(expression 5.3(a))}$$

For the "intrinsic range",

$$i_T = i_b e^{a\Delta V_T}, \quad \Delta V_T < 2$$

then using expression A.7.4.1

$$\frac{-i_b . dt}{c_m} = d\Delta V_T . e^{-a\Delta V_T} \qquad \ldots\ldots\ldots\ldots \text{A.7.4.2}$$

or

$$\frac{-i_b \cdot t}{c_m} = -\frac{1}{a} \cdot e^{-a\Delta V_T} + c \qquad \ldots\ldots\ldots\ldots \text{A.7.4.3}$$

or

$$\frac{-i_b(t - to)}{c_m} = -\frac{1}{a}\left(e^{-a\Delta V_T} - e^{-a\Delta V_o}\right). \qquad \ldots\ldots \text{A.7.4.4}$$

(expression 5.4)

where "t_o" is the time when $\Delta V_T = 2V$.

If the coordinate transformation is made such that

$$t' = t - to$$

$$\Delta V' = \Delta V_T - \Delta V_o$$

then expression 5.3 becomes

$$\Delta V_T'(t) = \frac{-i_b \cdot t'}{c_m}, \text{ for } t' < o, \Delta V_T' > o. \quad \ldots\ldots\ldots \quad 5.3(a)$$

and expression 5.4 becomes

$$\frac{-i_b \cdot t'}{c_m} = -\frac{1}{a}\left(e^{-a\Delta V_T'} - 1\right), \quad t' > o, \Delta V_T' < o \quad \ldots\ldots \quad 5.4(a)$$

or $$\left(\frac{i_b \cdot a}{c_m}\right) \cdot t' = e^{-a\Delta V_T(t)} - 1 \qquad t' > o, \Delta V_T' < o \quad \ldots \quad 5.4(b)$$

and $$a\Delta V_T'(t) = -\left(\frac{i_b \cdot a}{c_m}\right) t'$$

$$t' < o, \Delta V_T' > o \quad \ldots\ldots \quad 5.3(b)$$

For "a = 3·5", say, $c_m = 3{\cdot}3.10^{-15}$ F, and various values of "i_b", say 0·25 μA, 0·5 μA, 1 μA, 2 μA, 5·4(a) and 5·4(b) can be plotted.

As the equations now stand, they are most amenable for evaluation as both "t′" and "$\Delta V_T'$" are scaled by the same factor for both current ranges. The origin is centered at $\Delta V_o = 2V$, the contact potential.

The above is a modification of similar calculations in (253).

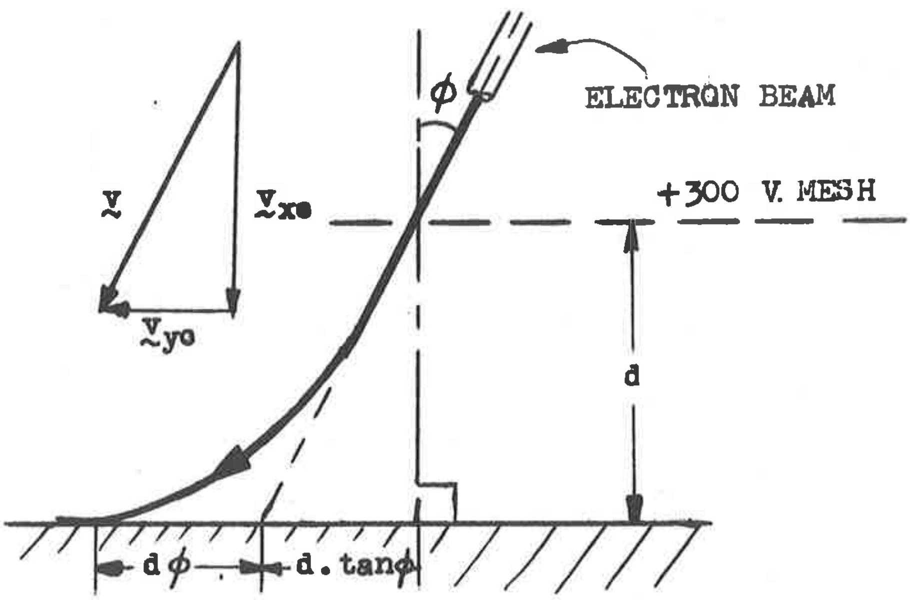
A.7.5 BEAM BENDING DUE TO MESH RETARDING FIELD

The relation between the force due to an electric field "$\underset{\sim}{E}$" on an electron of charge "e" and mass "m", and its resultant kinematic action is

$$-e\,\underset{\sim}{E} = m.\frac{d\underset{\sim}{v}}{dt} \qquad \ldots\ldots\ldots\ldots \text{A.7.5.1}$$

where "$\underset{\sim}{v}$" is the resultant velocity of the electron.

The retarding electric field normal to the photoconductor surface between the mesh at same $V_A >> \Delta V_T$ the potential at the target is $-\underset{\sim}{E}_R = \frac{V}{d}$, where "d" is the mesh-to-photoconductor spacing.



PHOTOCONDUCTOR TARGET

The initial velocity of the beam electrons on reaching the mesh due to their being accelerated by "V_A" (the emission cathode being at zero potential) is given from,

$$eV_A = \tfrac{1}{2}\,m\,v^2$$

and on deflection, at same angle "φ" to the normal, these initial velocities (and hence the initial conditions to the retarding electric field $\underset{\sim}{E}$),

$$v_{xo} = \sqrt{\frac{eV_A}{m}} \cdot \text{Cos } \varphi \qquad \ldots\ldots\ldots \text{A.7.5.2}$$

$$v_{yo} = \sqrt{\frac{eV_A}{m}} \cdot \text{Sin } \varphi \qquad \ldots\ldots\ldots \text{A.7.5.3}$$

Using expression 5.1 and integrating then

$$v_x(t) = -\frac{e}{m}\underset{\sim}{E} \cdot t + \sqrt{\frac{2eV}{m}} \cdot \text{Cos } \varphi \qquad \ldots\ldots \text{A.7.5.4}$$

Since there is no field acting in the y-direction, then integrating w.r. to time, and making the point where the electron beam reaches the mesh, the origin (i.e. (x,y) = (o,o)), then

$$y = v_{\delta} t = \sqrt{\frac{2ev}{m}} \text{Sin } \varphi \cdot t \qquad \ldots\ldots\ldots \text{A.7.5.5}$$

Assuming that on reaching the target the axial velocity "$v_x(t)$" is zero or close to zero (i.e. complete or near complete retardation), then for $V_x(t) \approx o$,

$$t = \sqrt{\frac{2ev}{m}} \text{Cos } \varphi \cdot \frac{m}{e.\underset{\sim}{E}_R} \qquad \ldots\ldots\ldots \text{A.7.5.6}$$

Substituting for "t" in A.7.5.5 gives

$$y \approx \sqrt{\frac{2ev}{m}} \text{Sin } \varphi \text{ Cos } \varphi \cdot \frac{m}{e.\underset{\sim}{E}_R} \qquad \ldots\ldots\ldots \text{A.7.5.7}$$

at the target.

As $\underset{\sim}{E}_R \approx \frac{V_A}{d}$

$$y = 2d.\sin\varphi.\cos\varphi$$

$$\approx 2d\varphi \quad (\text{for } \varphi < \pm 6^{o}) \quad \ldots \ldots \text{A.7.5.8}$$

As within "$\underset{\sim}{E_R}$", the deviation from the normal landing approach is equal to "$d\varphi$" (see Figure 77), the retarding field "$\underset{\sim}{E_R}$" contributes an added deflection "$d\varphi$".

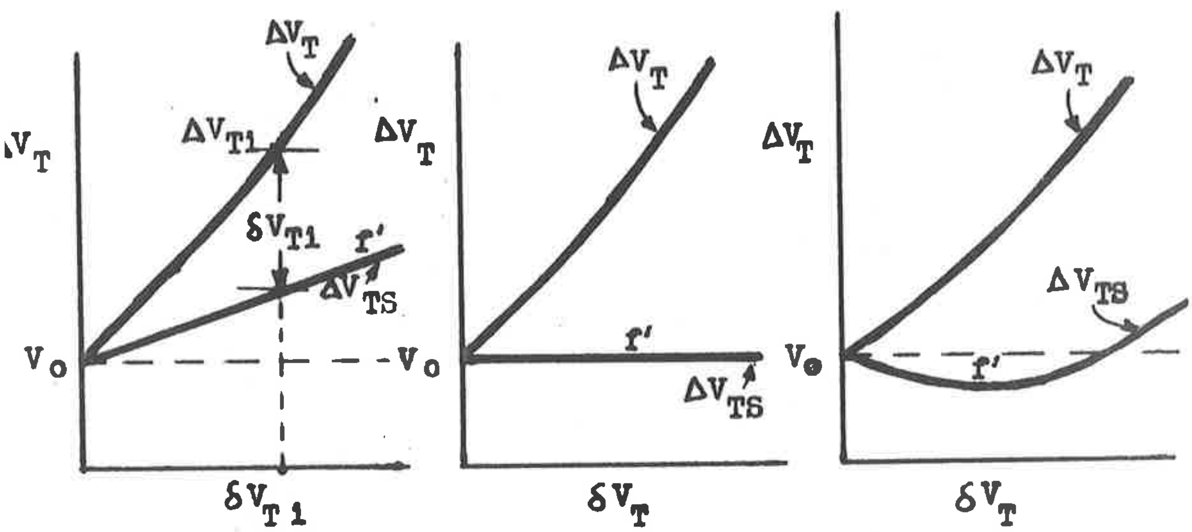
The angle of approach to the photoconductor target is ideally parallel to the target as:

$$\gamma = \frac{\Delta x}{\Delta y} = \frac{\Delta x}{\Delta t} \cdot \frac{\Delta t}{\Delta t}, \quad \text{where} \frac{\Delta x}{\Delta t} = v_x(t) = o.$$

Hence $\gamma \rightarrow o.$ A.7.5.9

A.7.6 PREDICTION OF EDGE-FLICKER FROM STABILIZATION CURVES

The general form of curves giving the target charge potential "ΔV_T" vs. the discharge voltage "δV_T" and the stabilization potential "ΔV_{ST}" is as shown:



(a) $f' > 0$ (b) $f' = 0$ (c) $f' < 0$

It is clear that

$$\Delta V_T = \Delta V_{ST} + \delta V_T \quad \ldots\ldots\ldots\ldots \text{A.7.6.1}$$

The following analysis assumes a linear "ΔV_{ST}" relationship for simplicity. It is an extension and development of results given by Van Polder(254).

Assume that the "ΔV_{ST}" curve has the form,

$$\Delta V_{ST} = V_o + f' \cdot \delta V_T \quad \ldots\ldots\ldots \text{A.7.6.2}$$

where "f'" is the slope of the "ΔV_{ST}" curve.

Hence the following expressions hold.

(a) for Stabilization Potential

$$\Delta V_{ST_i} = V_o + f' \cdot \delta V_{T_i} \quad \ldots\ldots\ldots \text{A.7.6.1a}$$

(b) Potential Prior to Discharge

$$\Delta V_{Ti} = \delta V_{Ti+1} + V_o + f' \cdot \delta V_{Ti+1}$$

$$= \delta V_{Ti+1}(1 + f') + V_o \quad \ldots\ldots\ldots \text{A.7.6.2}$$

(c) Output Potential Change

$$\delta V_{Ti+1} = \frac{\Delta V_{Ti} - V_o}{1 + f'} \quad \ldots\ldots\ldots \text{A.7.6.3}$$

giving

$$\text{Output Signal}, i_s = \frac{cm}{T_d} \cdot \delta V_{i+1} \quad \ldots\ldots\ldots \text{A.7.6.4}$$

Assume that initially the potential prior to the first discharge "D_1" is "ΔV_{To}"

Thus at discharge "D_1"

$$\delta V_{T1} = \frac{\Delta V_{To} - V_o}{1 + f'} \quad \ldots\ldots\ldots \text{A.7.6.5}$$

resulting in the stabilization potential ΔV_{ST1}

$$\Delta V_{ST1} = V_o + f' \cdot \delta V_{T1}$$

$$= V_o + \frac{f'}{1+f'} \cdot (\Delta V_{T_o} - V_o) \quad \ldots\ldots\ldots \text{A.7.6.6}$$

During the frame time internal "T_{F1}", a potential rise "ΔV_D" results, due to a "dark current."

Hence immediately prior to discharge "D_2"

$$\Delta V_{T1} = \Delta V_{ST1} + \Delta V_D \quad \ldots\ldots\ldots \text{A.7.6.7}$$

At discharge "D_2"

$$\delta V_{T2} = \frac{\Delta V_{T1} - V_o}{1 + f'}$$

$$= \frac{\Delta V_{ST1} + \Delta V_D - V_o}{1 + f'}$$

and substituting for ΔV_{ST1}

$$\delta V_{T2} = \frac{f}{(1+f)^2} \cdot \Delta V_{TO} - \frac{f}{(1+f)^2} \cdot V_o + \frac{\Delta V_D}{1+f} \quad \ldots\ldots \text{A.7.6.8}$$

The stabilization voltage is

$$\Delta V_{ST2} = V_o + f' . \delta V_{T2}$$

$$= V_o + \left(\frac{f'}{1+f'}\right)^2 \cdot \Delta V_{To} - \left(\frac{f'}{1+f}\right)^2 \cdot V_o + \left(\frac{f'}{1+f'}\right) \cdot \Delta V_D \quad \ldots \text{A.7.6.9}$$

Similarly after the 3rd discharge "D_3"

$$\delta V_{T3} = \frac{(f')^2}{(1+f')^3} \cdot \Delta V_{To} - \frac{(f')^2}{(1+f')^3} \cdot V_o + \frac{f'}{(1+f')^2} \cdot \Delta V_D + \frac{1}{(1+f')} \cdot \Delta V_D$$

and

$$\Delta V_{ST3} = \left(\frac{f'}{1+f'}\right)^3 \cdot \Delta V_{To} - \left(\frac{f'}{1+f'}\right)^3 \cdot V_o + \left\{\left(\frac{f'}{1+f}\right)^2 + \left(\frac{f}{1+f'}\right)\right\} \cdot \Delta V_D + V_o$$

Hence the general terms after n-discharges is

$$\delta V_{Tn} = \frac{(f')^{n-1}}{(1+f')^n} \cdot \Delta V_{T_o} - \frac{(f')^{n-1}}{(1+f')^n} \cdot V_o + \sum_{n=2}^{n} \frac{(f')^{n-2}}{(1+f')^{n-1}} \cdot \Delta V_D \qquad \text{A.7.6.10}$$

and

$$\Delta V_{STn} = \left(\frac{f}{1+f'}\right)^n \cdot \Delta V_{To} - \left(\frac{f'}{1+f'}\right)^n \cdot V_o + \sum_{n=1}^{n} \left(\frac{f'}{1+f'}\right)^{n-1} \cdot \Delta V_D + V_o \qquad \text{A.7.6.11}$$

Consider the following three cases and consider the current signal "i_s" output which is proportional to δV_{Tn} (see expression A.7. 6.4)

(a) $f' > 1$ hence $\frac{f'}{1+f'} < 1$

and $\frac{d}{d\delta V_T}(\Delta V_T) > f'$

Hence $$\delta V_{Tn} = \frac{\Delta V_D}{1+f'} \sum_{n=2}^{n} \left(\frac{f'}{1+f'}\right)^{n-2} + \underbrace{\left(\frac{f'}{1+f'}\right)^{n-1} \cdot \frac{(\Delta V_{To} - V_o)}{1+f'}}_{=\,o} = o.$$

The first term is a geometric series of common ratio $\frac{"f'"}{1+f'}$.

Evaluating this gives

$\delta V_{Tn} = \Delta V_D$ as expected (see Section 5.3.1)

(b) $f' = o$ hence $\frac{f'}{1+f'} = o$ and $\frac{d(\Delta V_T)}{d(\delta V_T)} > f'$

Hence $\delta V_{Tn} = \Delta V_D$ after the first and successive cycles.

(c) $f' < o$.

Consider three cases

(c.1) $o > f' > -\frac{1}{2}$, hence $\left|\frac{f'}{1+f'}\right| < 1$

and always negative.

Hence "δV_{Tn}" is decreasing in an oscillatary fashion.

(c.2) $-\frac{1}{2} > f' > -1$

then $\frac{f'}{1+f'} > 1$ and always negative, giving an oscillating increasing and decreasing output, because in the two terms, the series and the terms containing "ΔV_{To}-Vo", the powers differ by 1; their sum, giving the resultant output "δV_{Tn}", is the difference of their magnitudes. The "δV_{Tn}" increases until the region where $-\frac{1}{2} > f' > -1$ is crossed, thus decreasing the term $(\Delta V_{To} - V_o)$, and hence "δV_{TD}", until "δV_{Tn}" decreases and returns to the interval where $[-\frac{1}{2} > f' > -1]$.

(c.3) $f' < -1$, then $\frac{f'}{1+f'}$ is positive and always > 1, while $\frac{1}{1+f'}$ is negative. From the expression A.7.6.10 the output "δV_{Tn}" is -ve and unbounded; clearly it is an impossible case.

A.7.7 DERIVATION OF TIME INTERVAL BETWEEN ERROR SIGNALS

Assuming that each display location is scanned within the dwell time "T_d" ≈ 0·1 μs, generating an output pulse of width "T_d", then in the presence of a noise current "i_n", the worst case noise current signal is when it occurs at the boundary between two locations, with an amplitude

$$i_n > i_{L1},$$

such that in both locations, the "i_{L1}" Level Detector will be enabled "on", indicating a spurious display location.

Wider noise pulses may generate a larger number of error signals, but this is covered by the fact that noise may occur at each display location time interval independently of what occur at other locations. Two "superimposed" noise signals at the same location, each generating an error location, will still cause one error signal.

If the probability that the noise causes an error signal is "$P_{(i)}$", then the probability of this occurring at any of the $4{\cdot}5 \cdot 10^5$ possible locations is

$$4{\cdot}5 \cdot 10^5 \cdot P(i) \quad \ldots\ldots\ldots \quad A.7.7.1$$

Since these pulses are detected at the Level Detector "i_{L1}" once a frame time, the probability of error per frame time is still $4{\cdot}5 \cdot 10^5 . P(i)$. For 40 ms frame repetition rate, the error occurrence probability per second is

$$25 \cdot 4{\cdot}5 \cdot 10^5 . P(i) \quad \ldots\ldots\ldots \quad A.7.7.1(a)$$

If "T" is the minimum allowable time between unwanted signals causing errors, then

$$T = \frac{1}{25 \cdot 4{\cdot}5 \cdot 10^5 . P(i)} \quad \ldots\ldots\ldots \quad A.7.7.2.$$

"T" could at worst be, say, 1 minute, that being the time for the user to detect on his display screen an "obviously wrong" signal, either by it being an obviously "stray" display location, or from an understanding of the problem being solved. He could then erase such a display location.

A more useful "T" is the maximum length of continuous operation, in the worst case, say of 8 hours (or $\approx 3.10^4$ seconds), that being the maximum time between switching the unit off.

For P(i) to result in $T \geqslant 3.10^4$, it is required that

$$P(i) \leq \frac{1}{3.10^4 \, . \, 4.5.10^5 \, . \, 25} \qquad \ldots\ldots \text{A.7.7.3.}$$

$$\leq 3.10^{-12}$$

P(i) can be found from the Signal-to-Noise ratio at the output of the Vidicon Video amplifier, thus:

$$\frac{S}{N} \text{ ratio} = \frac{\text{Peak required current Signal}}{\text{RMS current noise}}$$

Vidicons are quoted in the literature as having $\frac{S}{N} \simeq 300$ (262), but with "aperture correction", this reduces it at the tube output to about 100:1. The video amplifiers (mainly the first stage of the amplifier) reduce this further to a $\frac{S}{N} \simeq 30\text{-}50$.

The noise current expected has a Gaussian distribution of current amplitudes.

The probability density "p(i)" for the noise current amplitude to lie between "i" and "i+di" is given by

$$p(i)di = \frac{1}{\sigma\sqrt{2\pi}} . \exp\left(\frac{-(i^2)}{2\sigma^2}\right) . di \qquad \ldots\ldots \text{A.7.7.4.}$$

where "σ" is the RMS noise current (265).

For pulses of amplitude "I+i" the probability density function of "I+i" is given by

$$p(I+i)di = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{-[I+i]^2}{2\sigma^2}\right) di \qquad \ldots\ldots \text{A.7.7.5.}$$

The cumulative probability "P(I+i)" of a signal exceeding some signal level "i_{L1}" (say the switching Current Level Detector level) is

$$P(I+i) = \int_{-\infty}^{i_{L1}} P(I+i)di \qquad \ldots\ldots \text{A.7.7.6.}$$

Values of P(I+i) are usually tabulated (e.g.(266)).

Four cases of error signals may be considered:

(i) a noise signal of amplitude $i_n > i_{L1}$ giving a spurious display location.

(ii) a noise signal superimposed on the largest unwanted signal "i_{see}", giving a spurious display location.

(iii) a noise signal of large enough amplitude subtracting from the required worst case signal "i_{Si}", being less than "i_{L1}", and thus not being displayed.

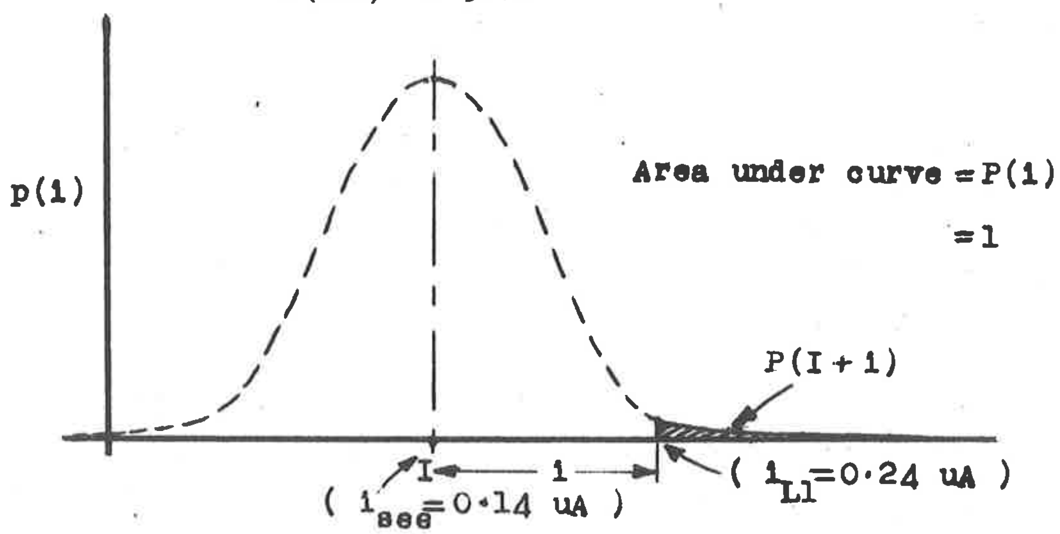
(iv) a noise signal subtracting from the steady state signal "i_{ss}" being less than "i_{L1}" and not being displayed.

Case (ii) covers case (i), while case (iii) covers case (iv).

Case (ii) of spurious signals being displayed

For a signal output "i_{ss}" of say 0·45 μA, and an "i_{see}" equal to 0·14 μA (see Fig.68) and "i_{L1}" set to 0·24 μA, then from the figure it is required that

$$P(I+i) \leq 3.10^{-12}$$



From tables (266), it is found that

for $\frac{i_{L1} - i_{see}}{\sigma} = 7$ then $P(I+i) = 1{\cdot}3 \,.\, 10^{-11}$

and for

$$\frac{i_{L1} - i_{see}}{\sigma} = 8 \quad \text{then} \quad P(I+i) \approx 6{\cdot}3 \,.\, 10^{-15}$$

thus

$$i_{L1} - i_{see} \geqslant 7{\cdot}2\,\sigma \quad \text{for} \quad P(I+i) \leqslant 3 \,.\, 10^{-12}$$

As $i_{L1} - i_{see} \approx 0{\cdot}1\ \mu A$

then

$$\sigma \simeq \frac{0{\cdot}1}{7{\cdot}2} \simeq 0{\cdot}014\ \mu A$$

Thus

$$\left.\frac{S}{N}\right)_{min} \quad \text{for the above fault-free time}$$

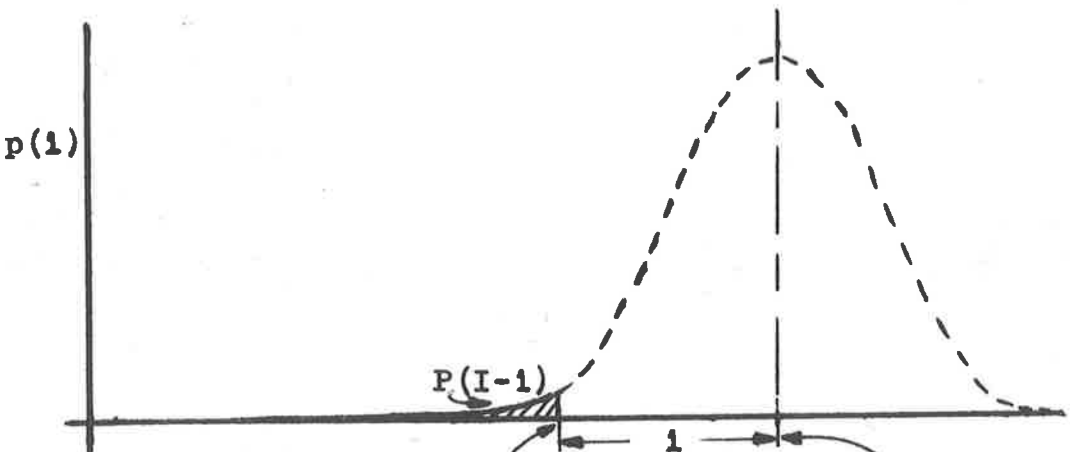
is

$$\left.\frac{S}{N}\right)_{min} = \frac{0{\cdot}45}{0{\cdot}014} \approx 32$$

As the expected "$\frac{S}{N}$" in a Vidicon and its Video amplifier is > 30, then the above fault free period requirements are met.

Case (iii) of required Signals being erased

For the worst case required (input) signal "i_{si}" of about 0·35 μA (see Fig.68) (whereas "i_{ss}" is 0·46 μA), and the threshold signal "i_{L1}" still set at 0·24 μA, the fault free conditions are very similar to the above, since the "p(i)" curve is symmetrical about "i_{see}" and "i_{si}" in this case. The cumulative probability "P(I-i)" is nearly identical to the previous case.



Hence this fault free condition is also met.

If the actual "$\frac{S}{N}$" at the Vidicon output is less than 30, the threshold "i_{L1}" need be adjusted only slightly (P(i) being an exponential with a squared argument) to satisfy the above requirements.

APPENDIX 8

PHOTOMETRY AND OPTICS

1 PHOTOMETRY

2 SPECTRAL SENSITIVITY OF VIDICONS

3 OPTICS

APPENDIX 8

PHOTOMETRY AND OPTICS

A.8.1. PHOTOMETRY

A.8.1.1. Photometric Quantities - Introduction

The use of Photometric quantities often gives rise to much confusion due to the abundance of the number of systems of units currently in use, based either on the M.k.s., E.g.s., or F.p.s. system (or a mixture of these), and on whether radiating sources are "point directional" sources or "diffusing (Lambert)" sources (see section A.8.1.4.)

The following is a brief outline of the quantities used in this report; conversions are facilitated by Tables A.8.T.1 and A.8.T.2.

A.8.1.2. Luminous Flux

LUMINOUS FLUX "F" (the unit is the "Lumen") is the measure of the illuminating power flowing across a given area per second. The lumen is defined as the Flux emitted into a unit solid angle by a uniform point source having a luminous intensity "I" of 1 candela

thus $F = I\omega$ ("ω" in steradians).............. A.8.1

If the luminous flux is incident at some angle "ϕ" to the normal of this surface,

$F = I\omega \cos\phi$ A.8.1a

Objectively speaking (ie. independently of any observer) luminous flux or power can be defined as

$F = \frac{dQ}{dt}$ watts A.8.2

where "Q" is the luminous energy in ergs say.

Quantity	Symbol	Defining Equation	Unit	Abbreviation
Quantity of light (luminous energy)	Q	$Q = \int_{380}^{760} K_\lambda U_\lambda d\lambda$	lumen-hour ‡*talbot* (lumen-second)	lm-hr lm-sec
Luminous density	q	$q = dQ/dV$	lumen-hour per cubic centimeter	lm-hr cm^{-3}
Luminous flux	F	$F = dQ/dt$	lumen	lm
Luminous flux density at a surface Luminous emittance Illumination ‡*Illuminance*	L E	$L = dF/dA$ $E = dF/dA$	lumen per square foot footcandle (lumen per square foot) lux phot	lm ft^{-2} fc lx ph
Luminous intensity (candlepower)	I	$I = dF/d\omega$ (ω = solid angle through which flux from point source is radiated)	candle (lumen per steradian)	c
Photometric brightness (luminance)	B	$B = d^2F/d\omega(dA \cos\theta)$ $= dI/dA \cos\theta$ (θ = angle between line of sight and normal to surface considered)	candle per unit area stilb footlambert lambert apostilb	c $in.^{-2}$, sb fL L asb

Table A.8.T.1 Photometric Units.

	Directional Sources			Lambert Sources		
$B \downarrow$	cd/cm² sb	cd/m²	cd/ft²	asb	L	ftL
1 candela/cm² 1 stilb (cd/cm² = sb)	1	10^4	$\frac{10^4}{a^2} = 929$	$10^4\pi = 31415.9$ (1 stilb = $10^4\pi$ apostilb)	$\pi = 3.14159$	$\frac{10^4}{a^2}\pi = 2919$
1 candela/m² = (cd/m²) NIT	10^{-4}	1	$\frac{1}{a^2} = \frac{9.290}{10^2}$	$\pi = 3.14159$	$\frac{\pi}{10^4} = \frac{3.14159}{10^4}$	$\frac{\pi}{a^2} = 0.2919$
1candela/sq.ft = (cd/ft²)	$\frac{a^2}{10^4} = \frac{1.076}{10^3}$	$a^2 = 10.76$	1	$a^2\pi = 33.82$	$\frac{a^2\pi}{10^4} = \frac{3.382}{10^3}$	$\pi = 3.14159$
1 apostilb (asb) = luminance of ideal diffuse reflector with 1 lux illumination	$\frac{1}{10^4\pi} = \frac{3.183}{10^5}$	$\frac{1}{\pi} = 0.3183$	$\frac{1}{a^2\pi} = 0.02957$	1	10^{-4}	$\frac{1}{a^2} = 0.09290$
1 lambert (L) = luminance of 1 lumen/cm² on "white"	$\frac{1}{\pi} = 0.3183$	$\frac{10^4}{\pi} = 3183$	$\frac{10^4}{a^2\pi} = 295.7$	10^4	1	$\frac{10^4}{a^2} = 929.0$
1 foot lambert (ftL) = equivalent footcandle = apparent footcandle = luminance of 1 lumen/ft² on "white"	$\frac{a^2}{10^4\pi} = \frac{3.426}{10^4}$	$\frac{a^2}{\pi} = 3.426$	$\frac{1}{\pi} = 0.3183$	$a^2 = 10.76$	$\frac{a^2}{10^4} = \frac{1.076}{10^3}$	1
black body luminance at point of solidification of platinum	$c = 60.00$ (by definition)	$10^4c = 60 \times 10^4$	$\frac{10^4c}{a^2} = 5.574 \times 10^4$	$10^4c\pi = 1.885 \times 10^6$	$c\pi = 188.5$	$\frac{10^4\pi c}{a^2} = 1751 \times 10^4$

Note: 1 metre = a feet = b inches $a = 3.2808$ $b = 39.370$ c = luminance of black body at point of solidification of platinum, equal to 60.

Table A.8.T.2 Luminance Units. (from [244])

This is the definition which need be used in determining the response of a photoconductor, as say a Vidicon, to the incident luminous flux. Expression A.8.1 and A.8.2 can be related together by the fact that "I" is related to the sensitivity of the human eye, and thus a subjective quantity (even though a "standard" eye may be defined). The sensitivity of the eye differs for different wavelengths, and thus incident radiant energy is weighed at each wavelength by the sensitivity or "luminosity" factor "K_λ". $K_\lambda = 1$ at $\lambda = 5550\mu$ and is less than 1 at all other wavelengths. "K_λ" is shown in Figure A.8.1.(b). Similar "photo-sensitivity" curves, or "spectral response" curves are available for each photo sensitive device. For Vidicon Photoconductors, these are shown in Figs. 66 (a) (c) and 66 (b) (c).

To obtain the luminous flux "F" in lumens, the radiant energy need be found. "P" in watts /sq.centimetre/ steradian, within a small wavelength interval "$d\lambda$" is say "$P_\lambda . d\lambda$" where "P_λ" is the "energy - spectrum" distribution.

The total amount of luminous flux emitted into 1 sq. cm into 1 sterradian is

$$F = 680 \int_{3800}^{7600} K_\lambda . P_\lambda . d\lambda \qquad \text{............ A.8.3}$$

where the factor "680" is used by standard convention as for a radiant power of 1 watt / cm^2/ steradian, F = 680 lumens at = 5550 μ.
The limits 3800 μ and 7600 μ are the wavelengths limits for visible light.
The above expression A.8.3 is used in the later section to determine Vidicon sensitivity under "W" phosphor illumination.

A.8.1.3. Luminous Intensity

LUMINOUS INTENSITY ,"I" (the unit is the "candela")
(or 1 lumen / 1 steradian)
is the measure of the illuminating power of a luminous source in some given direction under consideration.

(i) For a point source,"I", in a particular direction, is equal to "F" in lumens emitted in a unit solid angle into that given direction

$$ie \quad I = \frac{F}{\omega} \qquad \ldots\ldots\ldots\ldots\ldots \quad A.8.4$$

(ii) For a surface ,"I" is equal to the normal incident illumination "E_n" (in foot-candles) on the source surface, multiplied by the square of the normal distance between that point of measurement and the surface

$$F = E_n \cdot d^2 \qquad \ldots\ldots\ldots\ldots\ldots \quad A.8.4 \text{ (a)}$$

A.8.1.4. Luminance, Brightness

LUMINANCE, or "Physical" BRIGHTNESS ,"B" (unit is the "nit" or the "foot-Lambert" (ft-L) is the measure of luminous flux emitted by a projected unit area of a surface, or the intensity per projected area of a radiating surface in a given direction

$$B = \frac{dI}{d(\text{Area})} \quad \text{or} \quad \frac{I}{\text{Area}} \qquad \ldots\ldots\ldots\ldots\ldots \quad A.8.5$$

(1) For a point source, a source of intensity of 1 candela radiates 4π lumens.

(2) For a surface radiating in a preferred direction, a source of intensity of 1 candela radiates per unit area 1 nit (1 lumen / square centimetre)

(3) For a ideally diffusing surface (a Lambert radiator), an intensity of 1 candela normal to the surface radiates a

flux of π lumens (this is proven in most books on optics and photometry)

(4) For an ideally diffusing surface emitting a flux of 1 lumen / square foot, the luminance is 1 ft-Lambert

Thus for a diffusing surface

1 candela / unit area = π lumens / unit area

Note the difference between the two sets of units.

(i) For flat surfaces radiating in specified directions the units are "nits".

(ii) For perfectly diffusing surfaces, (such as CRT phosphor cathodoluminence) the units are ft-Lamberts.

A.8.1.5. Luminosity

LUMINOSITY, "L" is the total flux "F" radiated by a unit area of surface into a solid angle of 2π steradians; it is measured in "foot-candles" (ie 1 lumen / sq ft.)

A.8.1.6. Illumination

ILLUMINATION, "E" ("foot candles" or lumens / sq foot) is a measure of the luminous flux received by a unit area of a surface

$$E = \frac{F}{d^2} \qquad \text{.............} \qquad A.8.6$$

For light being incident at an angle "θ" to the normal,

$$E = \frac{L \cos \theta}{d^2} \qquad \text{.............} \qquad A.8.6\ (a)$$

For an ideal reflecting surface (whether a specular surface or diffusing surface), the total incident flux equals the reflected flux. Thus an illumation "E" of 1 ft-C (1 lumen / sq. ft) results in a brightness of 1 lumen / sq. ft. or 1 ft -Lambert

ie for a perfect diffusing surface

$$E\ (ft\text{-}C) \;\approx\; B\ (ft\text{-}L) \qquad \text{........} \qquad A.8.7$$

Normally a fraction "T_{rs}" is reflected only and thus

$$B \text{ (ft-L)} = E \text{ (ft-C)} \cdot T_{rs} \quad \ldots\ldots\ldots\ldots \quad \text{A.8.7 (a)}$$

If luminance "B" is due to an incident illumination of intensity in candelas, then

$$B \text{ (ft-L)} = \frac{E \text{ (candelas)}}{\pi} T_{rs} \quad \ldots\ldots \quad \text{A.8.7 (b)}$$

These Units and their defining relationships are tabluated in Table A.8.T.1

A.8.2 VIDICON SENSITIVITY

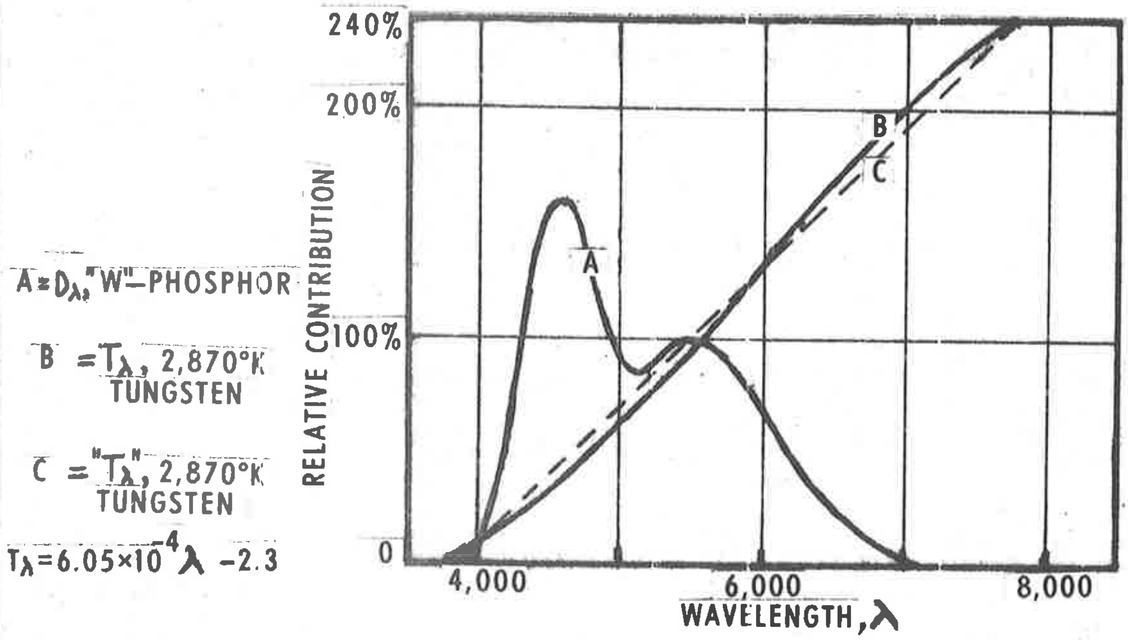
A.8.2.1 Introduction

The above definitions provide the basis for the illumination relationship linking the CRT emitted luminance with the incident illumination on the Vidicon. They also provide the basis for determining the sensitivity of the Vidicon photoconductor of incident illumination of a spectral distribution other than the 2870°K Tungsten filament illumination used to obtain the characteristics provided by the manufacturers (fig.66) Finally they allow the scaling factor in the expression for the Vidicon photo-conductor time constant to be found (expressions 4.15 and 4.16)

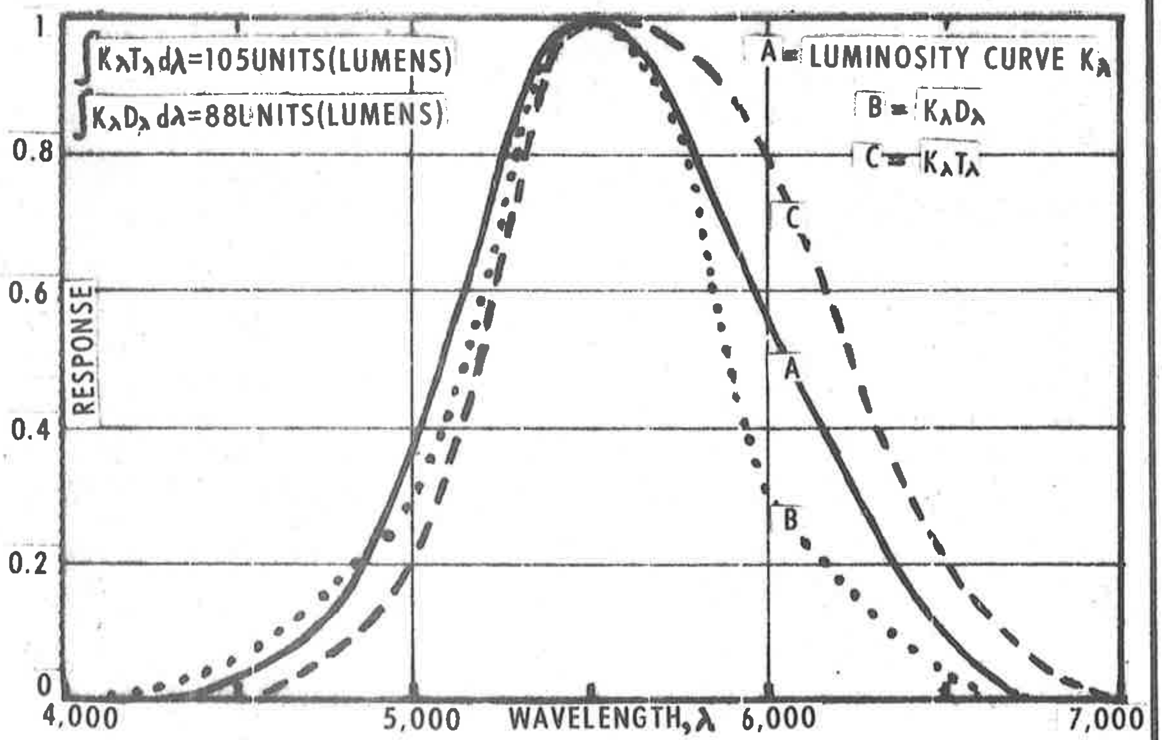
A.8.2.2. Sensitivity of Vidicon photoconductors to "W" phospher spectral illumination.

A.8.2.2.1. Introduction

From Figure 66 (b), the "current Vs illumination" ("I Vs E") are given for a standard illuminating source of "2870°K Tungsten Illumination" in foot-candles. "Colour temperature" illuminating sources radiate a specific radiant energy output of a given spectrum. For 2870 K° Tungsten this spectrum is shown in Figure



(a) Spectrum distribution of 'W' phosphor and of the standard Tungsten illuminant and its approximation.



(b) Conversion of radiant energy to lumens.

Figure A.8.1 Spectra of Illuminants and Luminosity Curves.

A.8 1(a) Where the output energy is given in watts, an "objective quantity". Conversion is thus required to the "subjective" illumination units.

The spectral energy distribution of the required "W" phosphor is shown in Figure 64 (b). Also available for the phosphor is its luminance (which again is a subjective quantity) at certain given levels of beam current etc. The radiant energies can be calculated from these two curves, and it is by a comparison of the objective luminous or radiant energies, that a valid comparison may be made with respect to their effect on a Vidicon photoconductor.

A.8.2.2.1. Equivalence Relationship between Radiant energy in Watts and Illumination in Foot-candles.

If "T gives the relationship of energy spectrum of the Tungsten illumination (shown in Fig.A.8.1.a) then according to expression A.8.3 the total equivalent amount of luminous flux is

$$F_T = 680 \int_{3800}^{7600} K_\lambda . T_\lambda . d\lambda \text{ lumens} \quad ..... \quad A.8.8$$

if T = 1 watt at λ = 5500 μ.

As the total radiant energy in this interval is

$$\int_{3800}^{7600} T_\lambda . d\lambda \quad \text{watts} \quad ...... \quad A.8.9$$

then 1 watt of "T_λ" illumination is equivalent to

$$F = 680 . \frac{\int_{3800}^{7600} K_\lambda . T_\lambda . d\lambda}{\int_{3800}^{7600} T_\lambda . d\lambda} \text{ Lumens} \quad .. \quad A.8.10$$

As 1 ft-Candle = 1 lumen / sq.foot, then 1 watt of "T_λ" illuminant incident on 1 cm² is equivalent to

$$929.680 \cdot \frac{\int_{3800}^{7600} K_\lambda . T_\lambda . d\lambda}{\int_{3800}^{7600} T_\lambda . d\lambda} \quad \text{ft Candles} \ldots \quad \text{A.8.10 (a)}$$

as 1 sq. ft = 929 sq cms.

Similarly if "D_λ" is the "W" phosphor spectrum energy distribution as per Figure A.8.1 (a), 1 watt of "D_λ" illumination is equivalent to, on a 1 square cm.,

$$929.680 \cdot \frac{\int_{3800}^{7600} K_\lambda . D_\lambda . d\lambda}{\int_{3800}^{7600} D_\lambda . d\lambda} \quad \text{ft Candles} \ldots \quad \text{A.8.10 (b)}$$

A.8.2.2.2. <u>Current output of Vidicon due to Illumination</u>

From the "spectral response curves" of the Vidicon photoconductor Class I and II, shown in Figure 66 (a) (c) and 66 (a) (b), the output signal current in μA per square cm. of photoconductor target area is given when the Vidicon is illuminated by a 'μW' (microwatt) of radiant energy of a given wavelength "λ". Ideally to correspond to visual response, this spectral response curve should have the identical shape as the "luminosity curve" in Figure A.8.1 (b). It can be seen that the maximum sensitivity is nearly identical for both at $\lambda = 5550\mu$. Calling the response curve for Vidicon Class I Photoconductor "$S_{\lambda I}$"

(and "$S_{\lambda 2}$" for class II Photoconductor), the output current from the Vidicon for each square cm. of target area when illuminated by "P_λ" watts in a frequency band "$d\lambda$" is

$$10^6 \cdot P_\lambda \cdot S_{\lambda I} \cdot d\lambda \qquad \mu A \qquad \text{.......... A.8.11}$$

Over the visible spectrum then, the output current due to

$$\int_{3800}^{7600} P_\lambda \cdot d\lambda \text{ watts is}$$

$$10^6 \int_{3800}^{7600} P_\lambda \cdot S_{\lambda 1} \cdot d\lambda \qquad \mu A/cm^2 \quad \text{.... A.8.12}$$

Thus 1 watt of this illuminating source results in

$$10^6 \cdot \frac{\int_{3800}^{7600} P_\lambda \cdot S_{\lambda 1} \cdot d\lambda}{\int_{3800}^{7600} P_\lambda \cdot d\lambda} \qquad \mu A/cms^2 \quad \text{... A.8.13}$$

Using the radiant energy and luminous flux equivalence relations of expressions A.8.10 (b) then 1 ft-candle of "T " illumination results in

$$\frac{10^6}{680.929} \cdot \frac{\int_{3800}^{7600} T_\lambda \cdot S_{\lambda 1} \cdot d\lambda}{\int_{3800}^{7600} K_\lambda \cdot T_\lambda \cdot d\lambda} \qquad \mu A/cm^2 \quad \text{.... A.8.14}$$

While for 1 ft-candle of "D_λ" illumination, the similar expression holds

$$\frac{10^6}{680.929} \cdot \frac{\int_{3800}^{7600} D_\lambda . S_{\lambda 1} . d\lambda}{\int_{3800}^{7600} K_\lambda . D_\lambda . d\lambda} \quad \mu A/cm^2 . \qquad \text{...... A.8.14 (a)}$$

A Vidicon target 9.6 x 12.8 mms has an area of 1.23 sq. cm^2; hence the above expression need be multiplied by "1.23" to give the total Vidicon output current.

From the known or supplied relationship "$S_{\lambda 1}$", "$S_{\lambda 2}$", "D_λ", "T_λ", and "K_λ", all of the above can be found. "$T_\lambda . S_{\lambda 1}$", "$D_\lambda . S_{\lambda 1}$", and "$T_\lambda . S_{\lambda 2}$" and "$D_\lambda . S_{\lambda 2}$" are shown in Figure A.8 2 (a)(b) reppectively. "$K_\lambda . T_\lambda$" and "$K_\lambda . D_\lambda$" are shown in Figure A.8 1 (b) respectively. Due to the distinctly and narrowly defined, "K_λ" the luminosity curve, both "$K_\lambda . T_\lambda$" and "$K_\lambda . D_\lambda$" are nearly identical.

The integrals are found by linear approximation of the corresponding relationship and found by graphical intergration. Evaluating the constant (and multiplying by 1.23 (sq.cm^2) results in the constant being "$1 \cdot 95 \times 10^{-6}$"

For "T_λ" it is found that

$$\int K_\lambda . T_\lambda . d\lambda = 1 \cdot 05.10^3 \text{ lumens}$$

$$\int T_\lambda . S_\lambda . d\lambda = 84 \ \mu A / \mu W / cm^2$$

Hence for class I Photoconductor, 1 foot-candle of Tungsten illumination generates 0.155 μ A at the Vidicon phototarget. However the spectral sensitivity curve is specified at some given "dark current" (or signal plate Voltage "V_s") and a given output current (o.o2 μA) (since the output current is not a linear function of illumination and depends on the Signal Plate Voltage.) For the RCA 8572A Photoconductor class I Vidicon, the output current of 0.02 μ A, corresponds from the " I_s vs E " curves to 0.17 ft-C (see Figure 66 (a) (b)). Thus 0.17 ft-C corresponds to an output current (0.17 x 0.155) = 0.026 μ A. This corresponds within the 50% tolerance limits to the 0.02μA output current obtained from the manufacturers' curves.

Similarly for a class II Photoconductor, from the curves the following are found:

the constant is $1{\cdot}95.10^{-6}$

$$\int K_\lambda . T_\lambda . d\lambda = 1{\cdot}05.10^3 \text{ lumens}$$

and

$$\int T_\lambda . S_\lambda . d\lambda = 244 \ \mu A/\mu W/\ cm^2$$

resulting in a current of 0.454μA for an equivalent of 1 ft-C of illumination.

Again for a class II Photoconductor Vidicon RCA 7262A for a dark current of 0.02μA and a signal current of 0.02 μA , then from the characteristic curves this occurs at an illumination E of 0.03 ft-C (see Figure 66 (1) (a)). The current thus evaluated is (0.454 x .03) = 0.014 μA which corresponds within 50% to the data supplied of 0.02μA.

A.8.2.2.3. <u>Vidicon Sensitivity with Arbitrary Spectral Distribution Illuminations.</u>

The object of the above exercise is to compare the effectiveness of the photoconductor illumination having the spectrum of the "W"-phosphor with that standard Tungsten illumination for which the manufacturers' data is provided. The above evaluation of the expressions and comparison with actual manufacturers' figures shows the validity of the derived expressions.

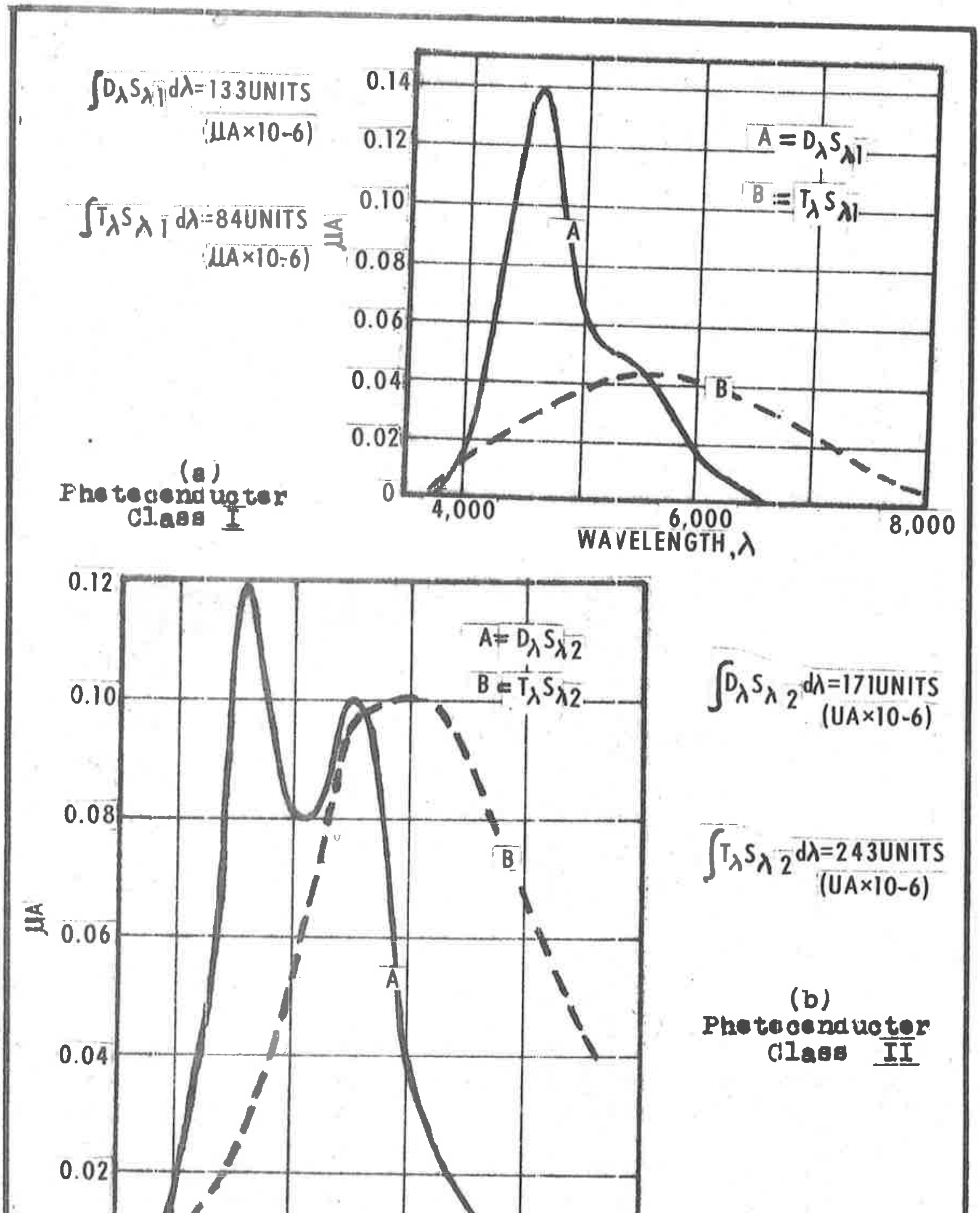


Figure A.8.2. Determinatien ef Vidicen Sensitivity.

Absolute values of output current cannot be calculated for the W-phosphor as the various "light vs current" relationships are not linear. What can be determined however, is whether an equivalent W-phosphor foot-candle of illumination is greater or less than the output current due to a foot-candle of Tungsten source illumination.

Hence comparing expression A.8.14 with expression A.8.14 (a) then

$$R_{D\lambda - T\lambda} = \frac{\left.1\text{ft-c}\right]_{D\lambda}}{\left.1\text{ft-c}\right]_{T\lambda}} = \frac{\int_{3800}^{7600} D_\lambda . S_{\lambda 1} . d\lambda \; . \int_{3800}^{7600} K_\lambda . T_\lambda . d\lambda}{\int_{3800}^{7600} T_\lambda . S_{\lambda 1} . d\lambda \; . \int_{3800}^{7600} K_\lambda . D_\lambda . d\lambda} \quad .. \text{ A.8.15}$$

If $R_{D\lambda - T\lambda} > 1$, then the "W"-phosphor illuminant is more effective than the Tungsten illuminant.

If $R_{D\lambda - T\lambda} < 1$, then the Tungsten illuminant is more effective than the W-phosphor illuminant

These various relationships "$D_\lambda . S_{\lambda 1}$", "$D_\lambda . S_{\lambda 2}$", "$K_\lambda . D_\lambda$" are shown plotted in Figure A.8.2. As ratios are only required, only the ratios of the areas under the curve need be compared. It is found that :

$R_{D\lambda - T\lambda}$ = 1·5 for class I Photoconductor.

and $R_{D\lambda - T\lambda}$ = 0·86 for class II Photoconductor.

To a first degree of approximation, the "W"-phosphor illumination can be considered to be as effective as the standard Tungsten illuminant. Hence the manufacturers' performance curves are valid for "W"-phosphor illumination.

A.8.2.3. <u>Number of photons per foot-candle of Illumination</u>

A photon has an energy "E" where

$$E = h\gamma \text{ ergs} \quad \text{............. A.8.16}$$

where "h" = $6{\cdot}63.10^{-27}$ erg.sec. (Planck's constant)

"γ" = frequency

Now 1 electon-volt = $1{\cdot}6.10^{-12}$ ergs

.... A.8.17

E = $1240/\lambda$ electron-volts, if "λ" is in $M\mu$.

For visible light "λ" is from 380 - 760 $m\mu$

Thus for "blue", $E_B = \frac{1240}{380}$ eV = $3{\cdot}26$ eV

for "red", $E_R = \frac{1240}{760}$ eV = $1{\cdot}63$ eV

It can thus be seen that photoconductors for visible light need have bandgaps "E_B" such that

$1{\cdot}63 < E_B < 3{\cdot}26$ (electron-volts) A.8.18

(Typically Sb S_3 has a bandgap of some 1·8 V (253))

Now 1 Amp = $6{\cdot}24.10^{18}$ electrons;

thus if the "quantum efficiency" (see section 5.7) is 1 (ie for each incident photon 1 electron or carrier is generated). then

"1 Watt of blue light" is equivalent to $\frac{6{\cdot}24.10^{18}}{3{\cdot}26} = 1{\cdot}92.10^{18}$ photons

and "1 watt of red light" is equivalent to $\frac{6{\cdot}24.10^{18}}{1{\cdot}63} = 3{\cdot}38.10^{18}$ photons

For each watt of visible light, somewhere between 1.9 and $3.8{\cdot}10^{18}$ photons can be expected.

Generally "N," the number of photons / watt of energy at some wavelength "λ", is

$$N = \frac{6{\cdot}24.10^{18}}{1{\cdot}24.10^{4}} . \lambda \quad \text{(photons / watt of energy at "}\lambda\text{")}$$

$$= 4{\cdot}95.10^{14} . \lambda$$ A.8.19

Thus for a illuminant of "T_λ" spectral distribution

$$\int_{3800}^{7600} T_\lambda . 4{\cdot}95.10^{14} . \lambda . d\lambda$$ photons are provided

by $\int_{3800}^{7600} T_\lambda . d\lambda$ watts of energy

Hence

$$1 \text{ watt}\Big]_{T_\lambda} = 4.95.10^{14} \frac{\int_{3800}^{7600} T_\lambda.\lambda d\lambda}{\int_{3800}^{7600} T_\lambda . d\lambda} \text{ photons} \quad \text{........ A.8.20}$$

As expression A.8.10(a) gives the equivalence between one watt of "T_λ" illumination and its equivalent illumination in foot-candles then

1 foot Candle of "T_λ" illumination is equivalent to

$$\frac{4 \cdot 95 . 10^{14} \int_{3800}^{7600} T_\lambda.\lambda d\lambda}{680.929 \int_{3800}^{7600} K_\lambda . T_\lambda . d\lambda} \text{ photons incident on 1 square cm.} \quad \text{.... A.8.21}$$

For a Vidicon of area 1.23 sq. cms, the above constant

$$\frac{4 \cdot 95.10^{14}}{680.929} . 1.23 = 9 \cdot 65.10^{8}$$

while $\int_{3800}^{7600} K_\lambda . T_\lambda . d\lambda = 105$ lumens

"T_λ" can be approximated from the graph in Figure A.8.1.(a)

as $T_\lambda = 6 \cdot 05.10^{-4}. \lambda - 2.3$ A.8.22

and thus $T_\lambda.\lambda = 6 \cdot 05.10^{-4}. \lambda^2 - 2.3\lambda$ A.8.23

Hence $\int_{3800}^{7600} T_\lambda.\lambda.d\lambda = 27 \cdot 7.10^{6}$ photons A.8.24

giving 1 foot-candle of "T λ" Tungsten illumination is equivalent to $2{\cdot}54.10^{13}$ photons / second

Redington (252) states that incident on a Vidicon Phototarget are "about 2.10^{16} photon / second / square foot for that colour temperature illumination (ie 2870°K Tungsten)". This is equivalent to $2{\cdot}6.10^{13}$ photons on 1·23 sq.cm, which checks very well with our $2{\cdot}54.10^{13}$ photons.

The above expression however, can be used to obtain the number of photons for any arbitrary illumination spectrum; it thus can be used to evaluate the photocurrent time constant for any arbitrary illumination.

A.8.3 OPTICS RELATIONSHIPS.

A.8.3.1. On-Axis Illumination from Sources.

For a Lambert diffussing radiator, an element of area "dA_1" radiating at a luminance of "B" ft-L/ft^2 (or $\frac{B}{\pi}$ candelas) has a resulting luminous intensity "I" on an element of Area "dA_2" at an angle "θ" to the normal line joining the two areas, of length " r " units, of

$$I = \frac{B}{\pi}.\cos\theta_1.dA_1 \quad \text{candelas} \quad \ldots\ldots \quad A.8.25$$

As $I = \frac{F}{\omega}$, where "F" is the luminous flux in lumens and "ω" is the solid angle subtended by "dA_1" at surface "dA_2" then

$$\omega = \frac{dA_2 \cos\theta_2}{r^2} \quad \ldots\ldots\ldots\ldots\ldots \quad A.8.26$$

Where "θ_2" is the angle of inclication of "dA_1" to the line of "r " units.

Hence the luminous flux received by "dA_2" due to "dA_1" is

$$F = \frac{B}{\pi} \cdot \frac{\cos\theta_1 . \cos\theta_2}{r^2} . dA_1 . dA_2 \quad \text{lumens} \quad .. \quad A.8.27$$

Assume a circular Lambert radiating source of Luminance "B" ft-L. It is required to find the illumination on the small elemental area "dA_2" on the axis linking the centre of the source and "dA_2". Dividing the source "S" into their annular rings of area

$$dA_1 = x \cdot d\phi_1 \cdot dx \qquad \ldots\ldots\ldots\ldots\ldots \quad A.8.28$$

and noting that $\theta_1 = \theta_2$, then the illumination at "dA_2" due to "dA_1" is from the definitions A.8.5 and A.8.7,

$$\text{Thus} \quad \frac{\text{Luminous flux at } dA_1}{\text{Area of } dA_2}$$

$$\frac{B}{\pi} \cdot \frac{\cos\theta_1 \cdot \cos\theta_2 \cdot x \cdot d\phi \cdot dx}{r^2} \qquad \ldots\ldots\ldots \quad A.8.29$$

$$= \frac{B}{\pi} \cdot \cos^2\theta \cdot \frac{x \cdot d\phi \cdot dx}{r^2} \qquad \text{foot candles}$$

$$\ldots\ldots\ldots \quad A.8.29\ (a)$$

As $x = h . \tan\theta$ then $dx = h \sec^2\theta . d\theta$

and as $\frac{x}{h} = \tan\theta$ then $\left(1 + \frac{x^2}{h^2}\right) = \sec^2\theta$

Thus the illumination "E" due to the elemental annulus is

$$dE = \frac{B}{\pi} . \cos\theta . \frac{h^2 . \tan\phi . \sec^2\theta . d\theta . d\phi}{h^2 \qquad \sec^2\theta} \qquad \ldots\ldots \quad A.8.30$$

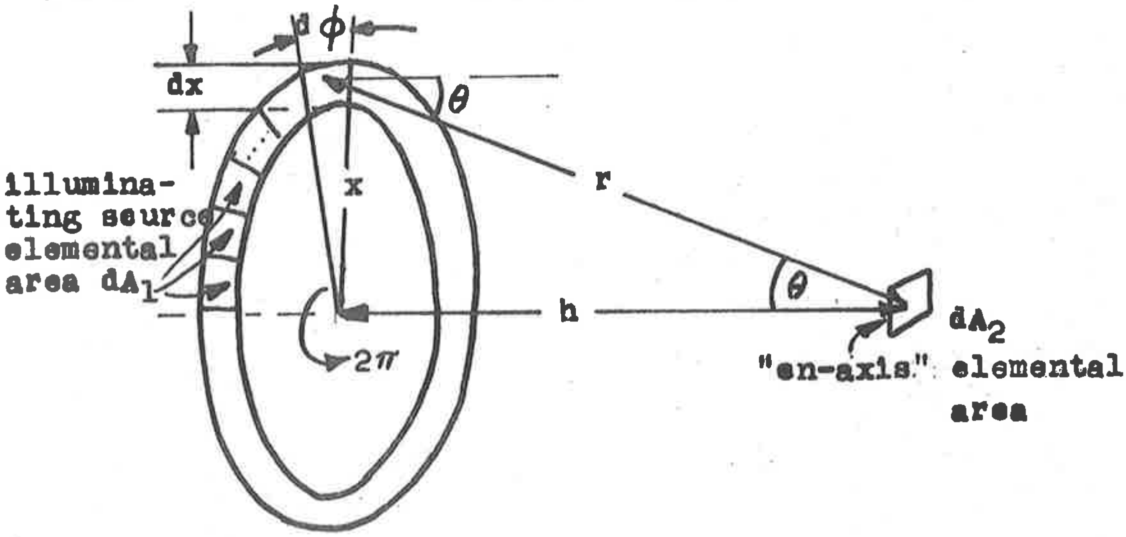
$$= \frac{B}{\pi} . \cos\theta . \sin\theta . d\theta . d\phi \qquad \ldots\ldots \quad A.8.30\ (a)$$

Integrating for "ϕ" from 0 to 2π as the annulus is circular, and for "θ" from 0 to "θ" results in the illumation "E" on the axial point "L" units distant as

$$E = 2\pi . \frac{1}{2} . \frac{B}{\pi} \sin^2\theta \qquad \ldots\ldots \quad A.8.31$$

$$= B . \sin^2\theta \qquad (\text{ft-C's}) \qquad \ldots\ldots \quad A.8.31\ (a)$$

This then is the illumination or incident flux at the entry pupil of the lens or an optic fibre (expression 6.20)



For the Schmidt Optical System, with a central blanked out area the integration is from "θ_1" to "θ_2", resulting in $E = B\ (\sin^2\theta_2 - \sin^2\theta_1)$ (expression 6.8).

A.8.3.2. Off-axis Illumination from Sources.

For an elemental area off the central axis, at an angle "α" to the axis then

(i) the effective radiating elemental annulus is $dA_1 \cdot \cos\alpha$

(ii) the effective receiving elemental area is $dA_2 \cdot \cos\alpha$

(iii) the effective distance is "$r \cdot \sec\alpha$" or "$\frac{r}{\cos\alpha}$"

modifying the above expression of illumination for an off axis receiving location to

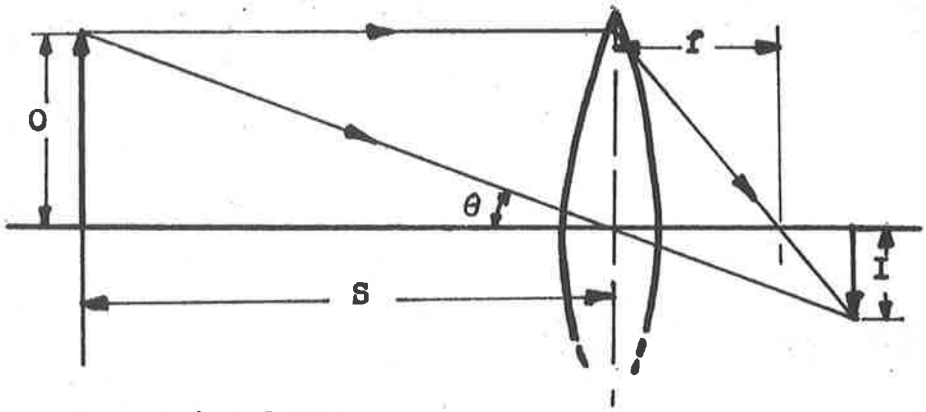
$$E = B \sin^2\theta \cos^4\alpha \qquad \text{.......... A.8.32}$$

Hence for extended receiving surfaces, the worst case illumination should be calculated using this expression.

In the case of an optic fibre, $\alpha \approx 1^{\circ}$ making "$\cos^4 \alpha$" equal to 1.

A.8.3.3. Lens Image Illumination.

If the expression represents the illumination at the entry pupil of a lens, then the resultant illumination at the image plane is found as follows:



For a converging lens:

$$\sin\theta = \frac{O}{\sqrt{O^2 + S^2}} \approx \frac{O}{S} \quad \text{as} \quad S \gg O$$

here "O" is the object height

"S" is its distance from the lens

The magnification "M" $= \dfrac{\text{Image Height}}{\text{Object Height}} = \dfrac{I}{O} = \dfrac{f\,(1+M)}{S}$

(by similar triangles)

Hence $\quad \sin\theta = \dfrac{I}{M} \cdot \dfrac{1}{f\,(1+M)}$ A.8.33

Allowing for transmission lens losses, a fraction "T" of the incident luminous flux on the lens contributes to the image illuminations,

then since $\dfrac{\text{image area}}{\text{object area}} = \dfrac{I^2}{O^2}$

and if the entry pupil illuminations is "E_O" and the image illumination is "E_I", then

$$E_I \cdot I^2 = T \cdot E_O \cdot O^2 \qquad \text{A.8.34}$$

or

$$E_I = T \cdot E_O \cdot \left(\frac{O}{I}\right)^2$$

$$= T \cdot \frac{E_O}{M^2} \qquad \text{A.8.34 (a)}$$

Substituting for $E = B \sin^2\theta \cos^4\alpha$ from expression A.8.32 and for $\sin\theta = \frac{I}{M.f.(1+M)}$, then

$$E_I = T \cdot \frac{B_O}{M^2} \left(\frac{I}{f\ (1+M)}\right)^2 \qquad \text{A.8.35}$$

If the "F-number" is defined as $\frac{f}{2.0}$ the ratio of the effective entry pupil diameter "O" and "f" is the focal length then

$$E_I = T \cdot B_O \cdot \frac{\cos^4\alpha}{4F^2\ (1 + M)^2} \quad \text{(ft-c)} \qquad \text{A.8.35 (a)}$$

giving the image illumination due to a lens of magnification "M" from an object of luminance "B" ft-L (expression 4.9).

A.8.3.4. Vidicon Lens Depth of Field.

In systems containing lenses, three-dimensional objects form "3-dimensional" images. For any given image plane (such as the photoconductor target surface of the Vidicon camera) only the points on a particular object plane are in sharp focus. Objects remote from this, ie not on the plane, are defocussed to a greater or lesser degree (and hence as the effective image size of these defocussed locations changes, so does the effective illumination). Points near the object plane are reproduced on the image plane as "circles of confusion", with the diameter increasing as the distance of these objects increases from the focussed object plane.

If the resulting "circle of confusion" for any image point is smaller than the smallest resolvable element in the image plane, then such a defocussed object will still appear in focus.

Thus for the eye, if the "circle of confusion" subtends less than 1 minute or so (see section 2.2.2.1), the object still appears focussed.

In TV systems, the smallest detail that can be resolved is the display element size; hence the "circle of confusion" which can be tolerated in less than or equal to the element size on the photoconductor target. For a 0.375" x 0.5" Vidicon (9.6 x 12.8m) with about 600 active lines, the element size is some

$$\frac{3}{8} \times \frac{1}{600} \approx 0.00062" \quad \text{(Call this "}d_c\text{")}$$

The total "depth" of an object along the axis of a lens producing an image where such "circles of confusion" is acceptable (it may exceed the above limits) is called the tolerable "depth of field".

It is shown in most books on optics (and photography (281)) , that if the distance of the in-focus object plane from the lens is "S" and if the diameter of the circle of confusion is to be $\leq d_c$, "M" is the lens magnification and "D" is lens aperture (diameter of the entry pupil of the lens) then the total depth of field "ΔS" is given by

$$\Delta S = \frac{d_c \cdot S}{M.D - d_c} + \frac{d_c \cdot S}{M.D + d_c} \quad \ldots\ldots \quad A.8.36$$

$$\doteq \frac{2d_c\ S.\ M.D.}{M^2D^2 - d_c^2}$$

The Vidicon lens is a 25 mm lens (1")
"M" for an Projection CRT with an effective display area of 3" x 4" is

$$M = \frac{\frac{3}{8}}{3} = \frac{1}{8}$$

"S" can be found from the geometry to be

$$S = \frac{f(1+M)}{M} = f.\ 8(1.125) = f.9''$$

and the "F-number" $= \frac{f}{D}$

$$\therefore\ D = \frac{f}{F}$$

Hence substiting into "ΔS" results

$$\Delta S = \frac{2.\ d_c . f . \frac{(1+M)}{M} . M . \frac{f}{F}}{M^2 \left(\frac{f}{F}\right)^2 - d_c^2} \quad \ldots\ldots\ldots \text{A.8.36 (a)}$$

Substituting the values as above and $d_c = 0.00062''$

$$\Delta S = \frac{2 . 0{\cdot}00062 .\ 1\ (1.125) . 1}{(0{\cdot}125)^2 . \frac{1}{F} - (0{\cdot}00062)^2 . F}$$

and as $(0.00062)^2 F \ll (0.125)^2 . \frac{1}{F}$

then $\Delta S = 2(0{\cdot}00062) . \frac{1.125}{(0{\cdot}125)^2} . F = 0{\cdot}089F \quad \ldots\ldots$ A.8.36(b)

With F = 1·9   S = 0·169"
F = 2·8   S = 0·25"

The CRT display screen is curved, thus giving rise to a non-flat object plane and hence circles of confusion. Table I gives the screen deviations from flatness, $\Delta S'$.

For the RCA projection screen

$$\Delta S > \Delta S' \quad , \quad \text{for} \quad F = 1.9$$

While for the Phillip's MW 13-38

$$\Delta S > \Delta S' \quad , \quad \text{for} \quad F = 2.8$$

The relationship between the tolerable screen deviation "$\Delta S'$" from flatness, and the F-Number is important as the illumination reaching the Vidicon target is $\propto \frac{1}{F^2}$

Hence for a larger "F-number" the illumination is smaller but greater screen curvature can be tolerated as the depth of focus is increased.

APPENDIX 9

PHOSPHORS AND CATHODOLUMINESCENCE

APPENDIX 9

PHOSPHORS AND CATHODOLUMINESCENCE

A.9.1 PHOSPHORS AND CATHODOLUMINESCENCE

A.9.1.1 Introduction

The CRT screen luminance at required display locations results from the impact of high velocity electrons of the electron beam onto the phosphor screen. The luminance is due to the exchange of the kinetic energy of the electrons into luminous energy by "cathodoluminescence".

In chapter 4, the requirements of the CRT luminance to realize the feasability of VIDIOGRAPHIC are given. These include:

(i) the available luminous output of the phosphor

(ii) the form of the persistance of the luminance after the beam excitation has ceased

(iii) the spectrum of the phosphor luminance (this has been treated in Appendix A.8.2.2).

(iv) "buildup" effects of luminance when a screen location, previously unexcited, is excited.

Within this Appendix the results used in chapter 4 are derived, along with a very brief explanation of cathodoluminescence.

A.9.1.2 Phosphors

Phosphors for CRTs are usually inorganic crystalline compounds containing additives ("activators") in varying quantities (from 0.01 % to several %). The resulting phosphors are, electrically, semi-conductors.

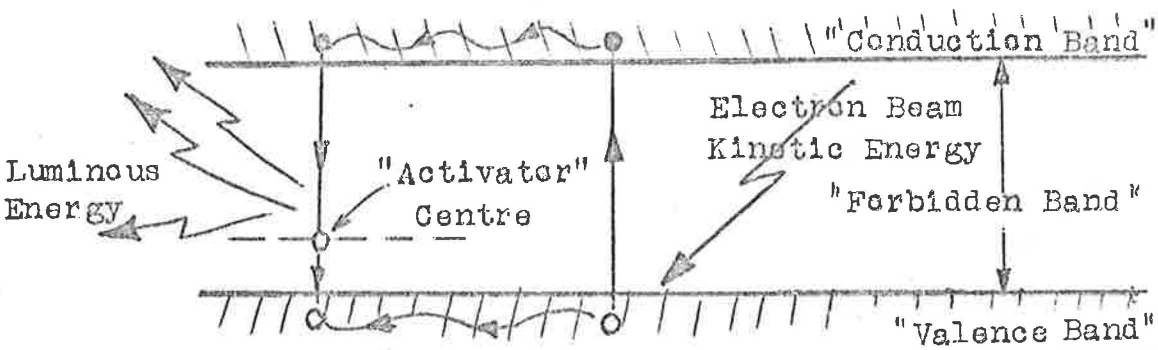
Most commonly the phosphors are sulphides of Zn, and silicates of Mg, Ca-Mg, Zn-Be, and phosphates, tungstates, fluorides and oxides of their elements. The "activators", usually Cu, Ag, Mn and Zn, enhance the luminous output and alter the persistance characteristics. Their action is explained below.

The phosphor is deposited on the inside of the glass envelope of the screen. The electron beam incident on the screen emits light equally well towards the viewer (through the glass screen) and back into the tube. Thus the light output is inefficiently utilized. Contrast also suffers, as the backward emitted luminance can, by internal tube reflections, be output again at different screen locations. In addition, the screen, being semi-conductive/insulating, will tend to accumulate charge (even though secondary emission is present), thus altering the screen potential and hence beam landing characteristics. To overcome these effects, a thin aluminium screen (some 500 A thick) is placed to sandwich the phosphor between it and the glass. This precludes luminance in the backward direction, reflecting such luminance back in the viewer-direction (and hence effectively doubling the viewer luminance); simultaneously contrast is enhanced. In addition, the screen potential is kept constant, the aluminium screen being conducting. However for the beam electrons to reach the phosphor through the aluminium film, the accelerating anode potential must be increased. The tradeoff between increased luminance and increasing the accelerating potential favours the presence of the aluminium screen if the accelerating potential is above some 5KV.

A.9.1.3 Phosphorescence

The basic process of the generation of luminance within the phosphor is difficult to describe with any degree of accuracy, due to the complexity of obtaining a valid model of the phosphor structure, particularly modelling the electron energy levels within the structure; on these depend the energy exchanges leading to the luminance output. The complexity of the problem is sufficiently described in (248).

The possible electron energy levels in an insulator are shown in the diagram. The electron energy increases vertically.



In the "filled" or the "valence band", all the electron energy levels are filled; hence electrons cannot be energized to move to higher levels within this band. The next allowable energy region is the "conduction band", separated by an energy difference, equal to the width of the "forbidden band", within which, for pure materials, no energy states can occur. Imperfections or "activators" introduce intermediate energy levels within the "forbidden band".

During electron beam impact, electrons in the filled band gain enough kinetic energy to jump into the conduction band, leaving a "hole" in the filled band. Both the hole and the conduction band electron have mobility and hence move. If due to this motion, the hole approaches an "activator" center, with its energy level in the "forbidden band", an electron in this level jumps to the lower energy state of the hole, destroying it. The "conduction band" electron, similarly during its motion, may become adjacent to the empty hole in the activator energy level, and jump into it; during this, it releases energy in the form of luminous energy.

From the expression A.8.1.7 in Appendix 8, it is seen that the wavelength of the emitted luminance depends on the electron energy release (in electron-volts). Hence a choice of activators, with different "forbidden band" energy levels result in different energy release and hence in different wavelengths. The type of activator thus determines the phosphor luminance spectrum.

The density of activator locations and thus the speed with which electron-hole recombinations occur, determines the intensity and persistance of the output luminance. Temperature, affecting the mobility of electrons and holes, also affects the persistance and luminance peak output.

In practice, the above simple model is more complex, with several energy levels in the forbidden band, and hence the type of interactions and their probability of occurance tend to complicate the picture. However the above elementary model is adequate for a simple explanation.

A.9.1.4 Phosphorescence Decay

The fact that the output luminance depends on the motion of electron/holes prior to recombination in the forbidden band, means that luminance is generated some time after the initial excitation producing these free electrons/holes, leading to the "persistance" of luminance after the initial excitation. The output luminance consists of two distinct phases:

(i) fluorescence, which is the luminance generated during the instant of excitation and within 10^{-8} secs. after the termination of excitation

(ii) phosphorescence, is the luminance generated after 10^{-8} secs after the excitation has been cut off.

The form of the luminance decay is usually classified as

(i) exponential decay

(ii) some form of inverse power law.

Generally, exponential decay has its decay time - constant independent of the initial level of excitation, while in the inverse power law decay, the decay and persistance is dependent on the initial level of excitation (248). Sulphide phosphors are predominatly power law decaying, while silicates are predominantly exponential decaying.

However with appropriate curve fitting, either exponential or power law decay (and buildup) curves can be fitted to experimental curves, leaving the predictability of decay with varying and high level excitation difficult to forecast (see the various curves fitted to the experimental results in (248)).

A.9.1.5 CRT Phosphors

The "W" phosphor, or its U.S. equivalent, "P-4" phosphor, used in TV Projection CRTs, have as their main requirements:

(i) the spectrum needs to be "white", or approximate the eye's spectral response (see Fig. A.8.1(b)). In practice, the phosphor has two distinct peaks in the spectrum (Fig. A.8.1(a)), one in the blue region, one in the yellow.

(ii) the luminance is to decay to a negligible value one frame time (40ms in the 50 Hz, 625-line system; 33ms in the 60Hz, 525-line system) after excitation. For the "P-4" phosphor this requirement is defined as "decreasing to no more than 7% of its initial output after 33ms".

Within these requirements there is ample scope for selecting the phosphor type and type and amount of activators. For example, in the literature (249), three forms of P-4 phosphors are usually reported-a sulphide, a sulphide-silicate, and a silicate. The point of all this is that each manufacturer has his own P-4 or monochrome phosphor. Their characteristics are not readily available nor readily supplied due to proprieatory rights. The data for W-phosphor being available and being typical of TV-Projection CRT phosphors, the following expressions derived are based on its data.

(iii) The conversion efficiency defined by

$$\frac{\text{"Output luminous energy"}}{\text{Input electric energy}}$$

is in the range of 1-20%, typically it is about 5-10%.

A.9.2 PRACTICAL CRT LUMINANCE RELATIONSHIPS

A.9.2.1 CRT Output Luminance

When an electron beam is incident on a phosphor screen there are four parameters which can be varied to control the resultant screen luminance:

(i) the beam current "i_b" which controls the number of phosphor-exciting electrons incident on the screen.

(ii) the accelerating final anode potential "V_A", which controls the final energy of the electron prior to screen impact

(iii) the area of excitation by the beam, at any instant, or the beam cross-sectional area at screen impact "A", controlling the current density incident on the screen.

(iv) the duration of excitation of the screen, or the "beam dwell time", "T_d", which determines the total number of electron incident on any location of the screen.

The expression relating the above factors to the resultant screen luminance B(t) is found to be (247):

$$B(t) = \frac{K}{A} \cdot (i_b \cdot \gamma) \cdot T_d \cdot (V_A - V_o)^n \qquad \text{A.9.1}$$

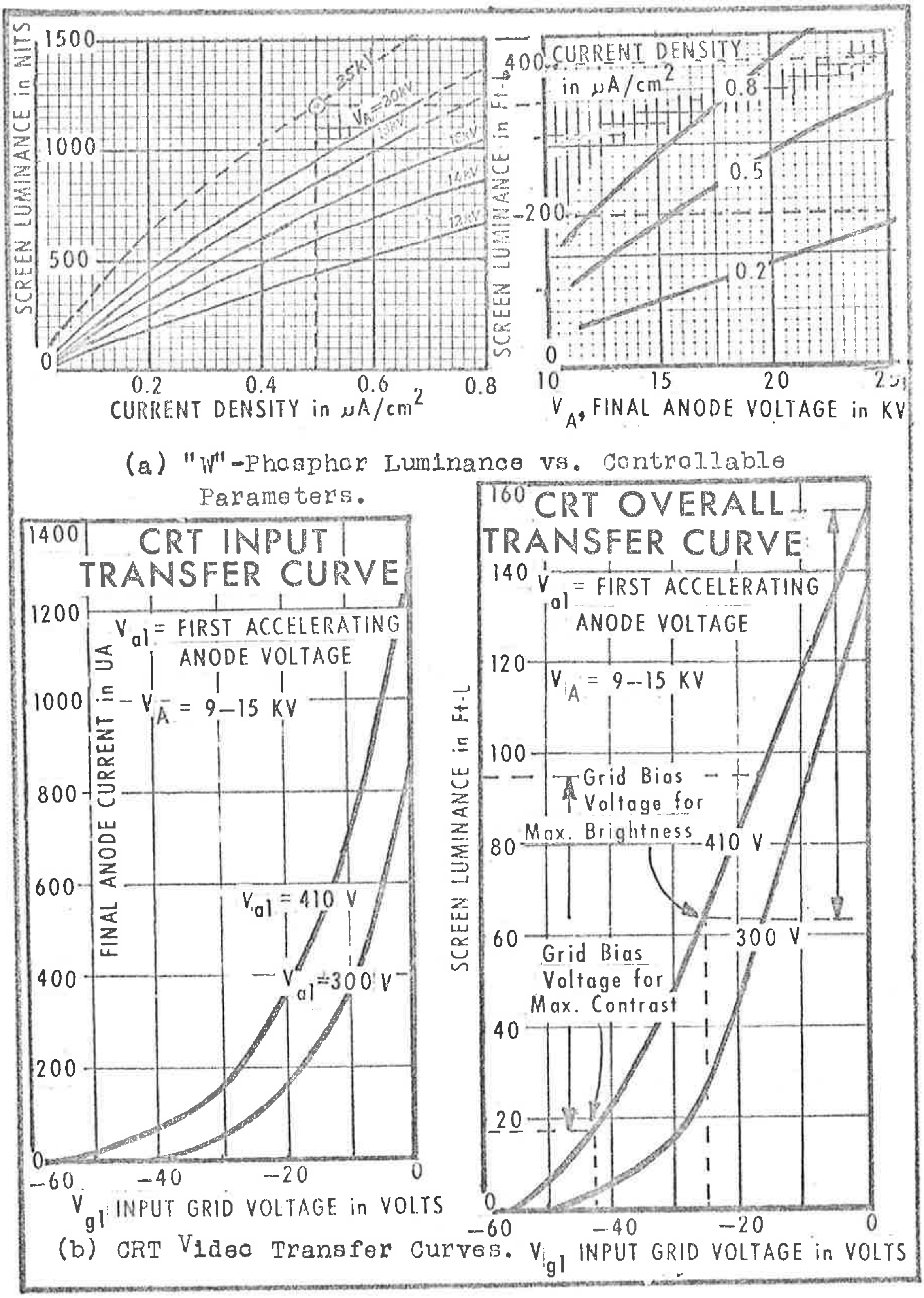
where "K" is a constant

"V_o" is a constant, and typically of the order of 100V (depending on the "surface purity" of the phosphor). With an aluminium-backed screen it is of the order of several thousand volts.

"γ" is a "constant" at low values of current (for "low" - see below) equal to 1, and at higher values of "i_b", it decreases. It indicates saturation of output luminance at higher values of "i_b".

"n" the "gamma" is of the order of 1·5→2, tending toward n=1 at "V_A" approaching 25KV.

Due to the limited range where the above expression holds (i.e. where saturation effects are not present), manufacturers occassionally supply phosphor luminance data in the form of graphical data where the output luminance is given as a function of the screen current density "loading" in $\mu A/cm^2$ for several values of accelerating voltage "V_A".



(a) "W"-Phosphor Luminance vs. Controllable Parameters.

(b) CRT Video Transfer Curves.

Figure A.9.1 "W"-Phosphor Luminance Output and Available CRT (AW 36-48) Input/Output Transfer Curves.

Such data is shown in Fig. A.9.1(a). Also given in the figure is "Output luminance vs Accelerating Voltage, V_A" for several values of screen current density. For the beam current of "i_b" and a beam area of A = "d x d", the screen loading

$$i_u = \frac{i_b}{A} \quad \mu A/cm^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{A.9.2}$$

This current acts during the dwell time "T_d"; hence on the average (per frame time "T_F") the actual screen loading/sec is

$$i_u = \frac{i_b}{A} \cdot \frac{T_d}{T_F} \quad \mu A/cm^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{A.9.2(a)}$$

"T_F" can be expressed in terms of "T_d"; due to the "T_H", the "Horizontal Blanking time," and "T_V", the "Vertical Blanking time," it follows simply that

$$T_F = T_d \cdot H.V. \left(1 + \frac{T_H}{64}\right) \cdot \left(1 + \frac{T_V}{20}\right) \quad \ldots\ldots\ldots\ldots \text{A.9.3}$$

where "H" = the display screen width in cms

"V" = the display screen height in cms

"64" = is the horizontal line duration in μs

"20" = is the vertical display duration in ms

"T_H" $\approx$ 10 us

"T_V" $\approx$ 2 ms

hence $T_F = 1{\cdot}3 . T_D . H.V$

Thus

$$i_\mu = \frac{i_b}{1{\cdot}3 .\ H.V} \quad \mu A/cm^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{A.9.2(b)}$$

For a 4" x 3" CRT display area this results in

$$i_u = \frac{i_b}{100} \quad \mu A/cm^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{A.9.2(c)}$$

From the curves shown in Fig A.9.1(a), for V_A = 25KV, the luminance corresponding to 350 ft-L (that given in section 6.3.5 as satisfying the requirement for adequate viewer screen brightness and Vidicon illumination) occurs at a current density loading of 0.5 u/cm^2.

Hence a beam current of 50 uA is required, which is quite an average value for beam currents (247).

A.9.2.2 CRT Input/Output Transfer Curves

For any given CRT, two more curves are necessary from which the resultant screen luminance "B(t)" can be obtained for any given input video voltage at the control grid. These are either supplied by the manufacturers or may be derived from other curves.

One curve, the "Input Transfer Curve" relates the input control grid voltage "V_{g1}" (obtained from the amplified CRT video input) with the final anodes current; this is usually supplied by the manufacturers, for several values of "first accelerating anode" Voltage "V_{a1}". For the 14" PYE TV monitor AW-36-48, the curve is shown in Fig A.9.1(b).

The second set of curves, relate the resultant screen luminance with the final anodes current in the previous set of curves. This is the CRT "Output Transfer Curve"; this is also supplied by the manufacturers (see Fig 64(c)).

These two curves are combined together to give the "Overall Transfer Curve", relating the input video control grid voltage "V_{g1}" to the CRT screen luminance. These are shown plotted in Fig A.9.1(b), for the AW-36-48 CRT in the available TV monitor.

A particular CRT is operated at some constant accelerating voltage. This is 14KV for the available CRT, and 25KV for a typical

Projection CRT ; and at some fixed value of "first anode accelerating anode, V_{a1}" ; in the AW-36-48, this is +410V.

The selection of the operating point at the control grid (i.e. the static control grid voltage, giving even background CRT luminance with no input signal voltage) is provided by the "brightness control", which provides a variable positive DC voltage at the cathode (in the case of the AW-36-48 this cathode voltage can be varied from +7V to +43V.). The control grid is kept at 0V, and thus relative to the cathode at -7 to -43V. The tube operating point can thus be varied between these limits.

The video control grid signal, positive going, can, if sufficiently high (+43V p. to p.), vary the screen luminance from about 20ft-L to 160 ft-L, if the grid is biased at "lowest" brightness (i.e. at +43V).

For the 14" PYE Monitor, the video amplifier stages, have a gain of 22dB max (i.e. a voltage gain of 12.5), with a 3dB bandwidth of 7·5Mhz. The maximum video input is limited from 0.5V to a max of 2V, resulting in a maximum control grid signal voltage, for a 2V input, of 25V. Thus from the curve it is seen that at the "best contrast" (low brightness), the output varies from a "no signal" luminance of 18ft-L to 94ft-L at maximum input signal (i.e. a contrast ratio of 5.2:1). For a maximum luminance output of 156ft-L, the biasing must be at -25V giving a no signal output of 70ft-L; the contrast is thus of just over 2:1.

In either case, the available CRT, AW-36-48 is wholly unsuitable for our purposes, even in the simple requirement of providing adequate CRT luminance.

This is not unusual as the PYE TV monitor is for studio monitor work, to be observed at close range by the operator; high luminances are thus not required.

Thus only Projection CRTs provide this adequate luminance required for VIDIOGRAPHIC.

A.9.3 CRT LUMINANCE DECAY

A.9.3.1 General Form of Luminance Decay

The CRT display location luminance as quoted in the above and given on the characteristic curves, are the average luminance per frame, i.e. the luminance averaged out over the interval between the instants of screen excitation, that is 40ms, due to the screen excitation in the beam dwell time, "$T_d \approx 0.1$us". As mentioned in section A.9.1.5, available luminance persistance/decay curves are not readily available. The data on the RCA Projection tube phosphor "P-4" is unavailable, and thus the "W" phosphor produced by Philips is utilized (247). This persistance curve is shown in Fig A.9.2(in) and redrawn on a log-log scale in Fig. A.9.2. This is for a "0.1 uA/cm^2 screen loading" whereas from above, for the 350 ft-L, a screen loading of 0.5 uA/cm^2 is required. Thus for this case, either the same decay will result (if W-phosphor is an exponential decay) or else for an inverse power decay, the decay will be faster. In either case this is adequate for the Vidicon photo current buildup when illuminated by this luminance - the faster the luminance decay, the faster is current buildup in the Vidicon and the smaller the remanent photo current at the end of the frame time. Thus the above W-phosphor characteristic can be used as a "worst-case" example of W-phosphor decay, for our purpose.

The decay curve has been divided into 4 segments, to simplify curve fitting. The numerical values of the expressions do not necessarily have any significance as far as determining the structure or behaviour of the phosphor. They have merely been derived for best-fit for integration purposes.

(i) Within the interval $10^{-5} < t < 10^{-3}$ sec,

$$B(t) = B_{max}\left(\frac{1}{9 \cdot 5 \cdot 10^4 t}\right)^{0.47} \qquad \text{A.9.4}$$

(ii) Within the interval $10^{-3} < t < 16.10^{-3}$ sec,

$$B(t) = B_{max}\left(\frac{1}{1+5 \cdot 3.10^3 t}\right)^{1.1} \qquad \text{A.9.5}$$

(iii) Within the interval $16.10^{-3} < t < 40.10^{-3}$ sec,

$$B(t) = B_{max}\left(\frac{1}{1+1 \cdot 95.10^3 t}\right)^{1.45} \qquad \text{A.9.6}$$

(iv) For $t > 40.10^{-3}$ sec,

$$B(t) = B_{max}\left(\frac{1}{0 \cdot 35.10^3 t}\right)^{2.4} \qquad \text{A.9.7}$$

These above analytically-fitted curves are shown plotted to indicate their variation (negligible) with the actual decay curve. Thus the average luminance within a frame time is given by:

$$B_{AV} = \frac{1}{T_F}\int_0^{T_F} B(t).dt \qquad \text{A.9.8}$$

$$= \frac{1}{40.10^{-3}}\int_0^{40.10^{-3}} B(t).dt \qquad \text{A.9.8(a)}$$

as "T_F", the frame time, is equal to 40ms.

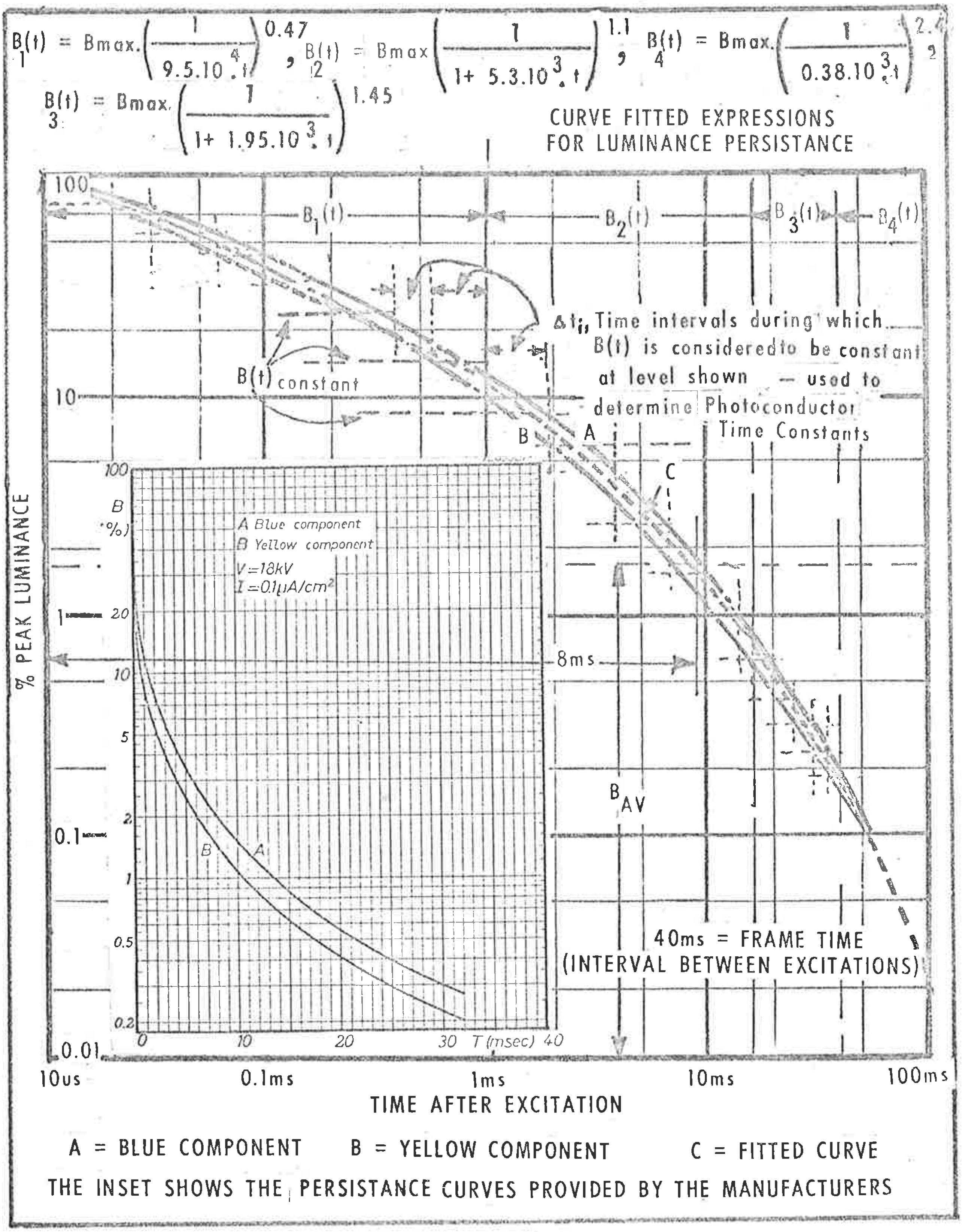


Figure A9.2 "W"-Phosphor Luminance Decay and Associated Parameters.

Hence

$$B_{AV} = \frac{1}{40.10^{-3}} \cdot \left(\int_{10^{-5}}^{-10^{-3}} \left(\frac{1}{9\cdot5.10^{4}.t}\right)^{0\cdot47} . dt + \int_{10^{-3}}^{16.10^{-3}} \left(\frac{1}{1+5\cdot3.10^{3}.t}\right)^{1\cdot1} . dt + \int_{16.10^{-3}}^{40.10^{-3}} \left(\frac{1}{1+1\cdot95.10^{3}.t}\right)^{1\cdot45} . dt \right) \quad \ldots A.9.8(b)$$

The integration poses no problems. Evaluating gives

$$B_{AV} = 0.0158\ B_{max} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots A.9.8(c)$$

$$\text{or } B_{max} = 63\ B_{AV} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots A.9.8(d)$$

Thus knowing "B_{AV}" for any given beam current values such as in Fig. A.9.1(a), "B_{max}" is deduced, and the persistance curve Fig. A.9.2 gives the luminance at any instant within the frame time.

From this curve the illuminance at the Vidicon photo-conductor target, modified by the scaling factor given by expression 4.11, can be obtained. Hence the photo current buildup and decay time constants, and illumination can be obtained as shown in Appendix A.10.5. The time intervals over which "B(t)" is approximated as being constant are also shown; these are used in the above mentioned Appendix A.10.5.

A.9.3.2 Light Pen Illuminating Peak Luminance

From the same curves, the average luminance available to the CRT Light-pen Illumination Source described in section 6.6.7.3 can now be derived. The 390-odd Optic Fibres have contributing luminance from a CRT area which is scanned within 0.735 ms. Within 0.735ms the decay from "B_{max}" is quite appreciable. Consequently the luminance output at the pen must be calculated using the average luminance within this 0.735ms interval.

Thus

$$B_{L\text{-}PEN\ AV} = \frac{1}{0{\cdot}735.10^{-3}} \int_0^{0{\cdot}735.10^{-3}} B(t).dt \qquad \ldots\ldots \text{A.9.9}$$

$$= \frac{B_{max}}{0{\cdot}735.10^{-3}} \int_0^{0{\cdot}735.10^{-3}} \left(\frac{1}{9{\cdot}5.10^4.t}\right)^{0{\cdot}47} .dt$$

$$\ldots\ldots \text{A.9.9(a)}$$

which, on evaluating, gives

$$B_{L\text{-}PEN\ AV} = 0{\cdot}23 . B_{max} \qquad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{A.9.9(b)}$$

Consequently the number of Optic Fibres had to be increased, as the CRT Illuminating Source, being extended, effectively has a decreased output luminance over the illuminating source area.

The increase in Optic Fibres compensated for this decreased light source luminance.

A.9.4 PHOSPHOR LUMINANCE BUILDUP

As mentioned in section 4.1.3 , and from the persistance curve of Fig A.9.2, remanent luminance is still present at times greater than 40ms, the frame time interval, contributing to the luminance in subsequent frametimes. Thus for any CRT display location, at any frame time, some fraction of its luminance is due to contributed luminance from previous frames. This fraction is of interest, as at instants of new display information being inserted, either via the User Light-pen or the CPU I/O

Interface, no luminance contribution from previous frame times contributes to the output luminance. Of interest is this reduced luminance, to see its effect on Vidicon photo·current buildup.

In the persistance curve of Fig A.9.2, the decay for "$t > 40ms$" is approximated by

$$B(t) = B_{max}\left(\frac{1}{0\cdot35.10^3.t}\right)^{2.4} \qquad \text{A.9.7}$$

Now, assuming that the Principle of Superposition holds for luminance (see expression A.9.1) i.e. saturation effects are negligible, then if two separate values of luminance result for two separate effects, then if those two effects are simultaneously applied, the resultant luminance is the sum of the previous two values. For the relatively low values of current loading of the CRT screen, it is seen from the curves of Fig A.9.1 that we are still operating in the linear region where Superposition still holds.

At any one frame time interval, the luminance is due to the luminance generated by the excitation at the beginning of the frame time plus the luminance remanent from the prior frame times, as follows:

$$B(t)_{\text{PRESENT CYCLE}} = \frac{1}{T_F}\int_0^{T_F} B(t).dt \qquad \text{(present cycle)}$$

$$+ \frac{1}{T_F}\int_{T_F}^{2T_F} B\ t).dt \qquad \text{(1st prior cycle)}$$

$$+ \quad \cdots\cdots$$

$$+ \frac{1}{T_F}\int_{2T_F}^{3T_F} B(t).dt \qquad \text{(2nd prior cycle)}$$

$$\vdots$$

$$+ \frac{1}{T_F}\int_{n.T_F}^{(n+1).T_F} B(t).dt \qquad \text{(n-th prior cycle)}$$

$$= \frac{1}{T_F}\int_{0}^{T_F} B(t).dt + \frac{1}{T_F}\int_{T_F}^{\infty} B(t).dt \quad \ldots\ldots \text{A.9.10}$$

The luminance contribution due to the prior cycle is thus given by:

$$B(t)_{CONT} = \frac{1}{40.10^{-3}}\int_{40.10^{-3}}^{\infty} B(t).dt$$

$$= \frac{1}{40.10^{-3}}\int_{40.10^{-3}}^{\infty} \left(\frac{1}{0\cdot 35.10^{3}.t}\right)^{2\cdot 4} .dt \quad .. \text{A.9.11}$$

where T_F = 40.10^{-3} secs, the frame time.

Integrating then

$$B_{CONT}(t) = 0.00128\ B_{max} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{A.9.11(b)}$$

As $\quad B_{AV\ TOTAL} = 0.0158\ B_{max}$

then $\quad \frac{B_{CONT}}{B_{AV\ TOTAL}} = \frac{0.00128}{0.0158} \approx 0.081 \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{A.9.12}$

Thus about 8% of the total steady state luminance is contributed by the luminance generated prior to that particular frame time, or the luminance output from newly inserted display data is some 92% of the steady state display location luminance.

To allow for approximations etc., a reduction of 10% (rather than 8%) of the steady state luminance was used to calculate the photo-currents and Vidicon output signals due to newly inserted data.

. . . .

APPENDIX 10

PHOTOCONDUCTIVITY

APPENDIX 10

PHOTOCONDUCTIVITY

A.10.1 PHOTOCONDUCTORS AND PHOTOCONDUCTIVITY

A.10.1.1 Introduction

When the CRT screen luminance or user light pen luminance is imaged onto the Vidicon phototarget, a charge is built up at phototarget locations dependent on the illumination. The resultant charge pattern on the phototarget, when discharged by the scanning electron beam, results in a video time signal representative of the incident imaged CRT display or light pen input. The process whereby charge is built up due to incident illumination is termed "photoconductivity"; and the material which enables this is a "photoconductor" which, by definition, is a material whose electrical conductivity is increased by the absorption of incident light or other suitable radiation.

Of interest here are the properties of photoconductors specifically applied to Vidicons, the parameters on which their performance depends and by the knowledge of which Vidicon performance can be optimized, and particularly suitable expressions for photocurrent rise times and decays to ensure that within the scanning interval or frame-time of 40 ms duration, new levels of photoconductor illumination (due to newly inserted graphics data) can be detected by the system, and existing display locations can be erased as required. The expressions for photocurrent decay and photocurrent rise are derived here as no such expressions have been found in the literature.

For good reasons has photoconductivity been described as "one of the most complex phenomena of solid state physics" (331). Even within the restricted field of Vidicon-application photoconductors great discrepancies exist between experimental and theoretical results, the "discrepancy" often being several orders of magnitude! Even where experimental results apparently have agreed with theoretical or predicted results, doubts have been expressed as to their validity, or the theory, or conclusions. For example "There is no basis for the conclusions derived in the work of Redington and a number of other works..." (332) and "The assumptions used for the derivation of.. [the "fundamental" expression for photoconductors]..are not of general validity" (333) and so on. (As indicated below the main problem is the great difficulty in deriving a valid model of the photoconductor energy band structure, particularly in the presence and distribution of "trapping centres").

It is not our province, nor is it presumed, that within this short Appendix, the above mentioned discrepancies can or will be resolved. Rather than go to first principles and hopefully derive expressions which may or may not be correct, the expressions derived herein are based on existing manufacturers' curves (as in Figs.66); the real test of course is to see whether the expressions derived check for performance figures as supplied by the manufacturs. In that respect the derived expressions have been fairly succesful.

The emphasis within this short review is photoconductivity and photoconductors as applied to Vidicon requirements.

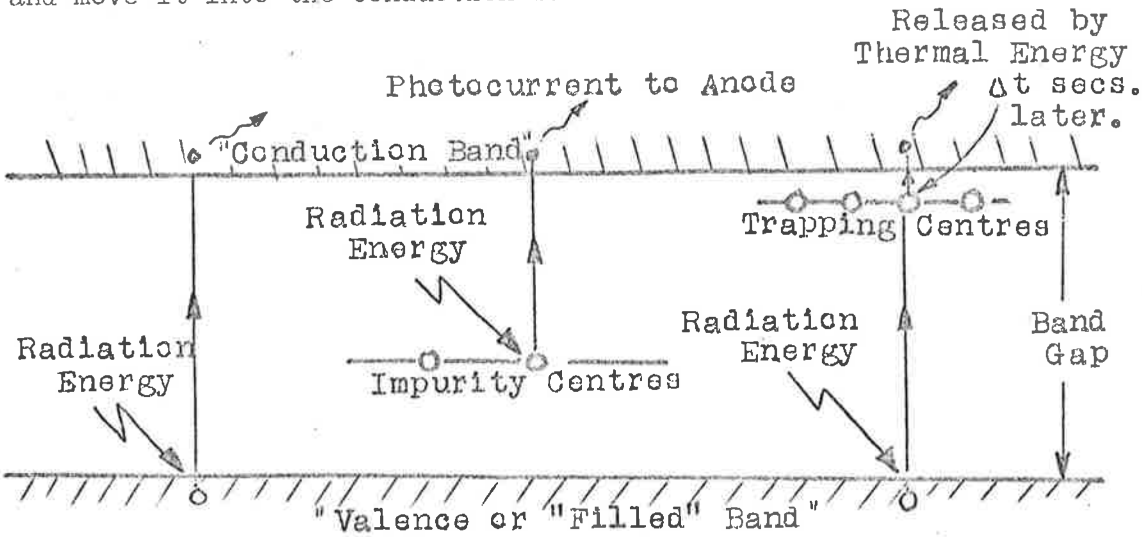
A.10.1.2 Photoconductivity

Photoconductivity occurs in every insulator or semi-conductor, and even in certain organic materials. From the view-point of suitability for illumination by "white light" and matching the spectral response of a viewer's eye, and from the viewpoint of response and conductivity, suitable photoconductors are rather limited. For Vidicons, "red" antimony trisulphide, Sb_2S_3, is almost exclusively used; other materials such as amorphous selenium (Se), PbO (in"Plumbicons"), ZnSe,have been used with greater or lesser success.

The intrinsic resistivity of Sb_2S_3 is in the range of $10^{11} - 10^{13}$ Ω cm, which characterise it electrically as an insulator. Due to light incident on it, photoconductivity reduces the effective resistivity 100 - 1000 fold.

For insulators or semiconductors exhibiting current conducting properties, recourse must be made to the energy diagram of electrons, to see how external sources of energy (light energy in this case) gives rise to electrical currents. This is somewhat analogous to the process of cathodo-luminescence (Appendix 9) where the external energy (the kinetic energy of impacting beam electrons) gave rise to luminance after a similar process of generation of charge carriers. Indeed the similarity between cathode-luminescence and photoconductivity (or rather between phosphors and photoconductors) has been recognized and the corresponding phenomena have been explained in similar energy/electron generation terms (334).

Basically the process of photoconductivity occurs as follows. Within an insulator the highest energy band containing electrons is the "Valence band"; this is completely filled. The next electron energy states which are possible lie in the conduction band, separated from the valence band by the "band gap" or "forbidden gap", which is characteristic of the material. The band gap is that energy required to free an electron from the valence band and move it into the conduction band.



When light energy is incident on the photoconductor, the constituent photons have associated energy dependent on the wavelengths of the illuminating light. This associated energy, in electron-volts, has been given in Appendix A.8.2.3, where it was found that a photon corresponding to

(1) "red light" (7000 A) has an associated energy of about 1.8 eV

(2) "blue light" (3800 A) has an associated energy of about 3.2 eV.

When such light is incident, the photon energy is made available on collision to the electrons in the valence band. For injection into the conduction band, the photon energy exchanged must be greater than the band gap. For visible light energy this means that the band gap must be no greater than 1.8 eV. If it is higher, then the "red" response of the photoconductor is very poor, and if it is lower the photoconductor would respond to infra-red (this has some uses but notin TV applications).

Sb_2S_3 has a band gap of 1.8 eV (333), hence its almost exclusive use in Vidicons.

Normally there are certain impurities and defects in the photoconductor, which result in energy levels in the forbidden band. This lowers the energy necessary to excite an electron into the conduction band and is also responsible for the "dark current" (see below, section 10.1.3.1).

In addition, "traps" modify the behaviour of photoconductors even more drastically (see below section 10.1.3.3).

Thus incident illumination generates conduction electrons whose motion due to some external electric field gives rise to photo-currents. If the photoconductor has a very high transverse resist-ance (by making the thickness of the photoconductor much smaller than the dimensions of an illuminated area) and an electric field exists across the photoconductor (due to "V_s", the signal plate voltage), the conduction band electrons result in an electric current in the localized illuminated area, depleting charge from that location and giving rise to a charge corresponding to the illuminated location. The effectiveness of this depends on the photoconductor conductivity, transverse resistance, etc., explained in more detail below.

A.10.1.3 Photoconductor Properties

The performance and characteristics of a photoconductor depend on four parameters (335), which are:

(i) Dark Current, or the electrical activity under no illumination.

(ii) Spectral Response, the relative sensitivity to radiation of different wavelengths.

(iii) Speed of Response to changes in illumination intensity.

(iv) Photosensitivity or the steady-state change in photo-current for a steady-state change in illumination intensity.

A.10.1.3.1 Dark Current

Normally, with no incident illumination (or rather with the existing stray or ambient background illumination), thermal excitation is present which generates carriers; hence some photo-currents flow and charge leakage at the photoconductor surface exists. Its magnitude depends on the band gap, the amount and type of impurities present and on the electric field across the photo-conductor (i.e. on "V_s", the signal plate voltage). The magnitude of the dark current is usually kept small with respect to the required signal, say 10-20 times less. For a Vidicon the relationship between dark current and "V_s" is provided by the manufacturers (see Fig.66 (a)).

A.10.1.3.2 Spectral Response

The band gap for a pure material corresponds to the longest wavelength (smallest energy) of the incident illumination which results in a free or conduction-band electron being generated. When impurities or material imperfections are present resulting in energy levels in the forbidden band, less energy (i.e. longer wavelengths) is required to inject electrons in the conduction band. Hence photocurrents are generated for longer wavelengths (in the red region) than otherwise predicted.

In practice a peaked photocurrent/wavelength response occurs at wavelengths of energy very close to the band gap. For wavelengths whose energy is near or less than the band gap, the decrease in photocurrent is obvious. For higher energy wavelengths the lowered response occurs for 3 reasons:

(i) there are less photons for the same energy in shorter wavelengths.

(ii) complete absorption of short wavelengths occur in the surface regions of the photo-conductor in comparison to the bulk of the photoconductor (due to different structure).

(iii) the glass faceplate in Vidicon tubes absorbs more energy from shorter wavelengths than longer wavelengths.

A.10.1.3.3 Speed of Response

Certain expressions have been derived some years ago when the study of photoconductivity was intensively pursued (336) which indicated large photocurrents and fast responses for given illuminations. When compared with experimental results, discrepancies of several orders of magnitude were found to exist. These discrepancies were explained by "traps" or "trapping centres" which are energy states introduced by impurities, which temporarily remove electrons from the free or conduction-band state, and then return them to the free state at somelater time when sufficient thermal energy has been supplied from the material lattice (335).

The total amount of free electrons due to some illumination may be that predicted from the theory, but the time interval during which their presence is detected (by the resultant photocurrent or charge being depleted) is lengthened. This appears as a delay or buildup in photocurrents, or when illumination is cut off, as a "long" photocurrent decay. Similarly the amplitude of the photocurrent is affected, for if the total number of freed electrons "N" is as predicted by the theory, then, for the associated increased time delay "Δt", the detected photocurrent $i = \frac{N.e}{\Delta t + t}$, a smaller value than if "t" was acting only.

During initial illumination, the photocurrent delay occurs because trapping centres are filled by the removal of photo excited electrons, until some steady state value is reached when the rate of trapping equals the rate of release from the centres.

When excitation illumination has stopped, the trapped electrons are released slowly, resulting in a slow photocurrent decay time. (This decay is not to be confused with the apparent signal current decay in Vidicons due to incomplete beam discharge at low levels mentioned in Chapter 5).

The presence, distribution, density and nature of these traps vary from material to material (and even from sample to sample) and depend on the process used in making the photoconductor, as the structure of the material depends on this. The exact nature of a photoconductor sample, as far as the determination of the above quantities go, is difficult to assess and numerical figures are even harder to determine. This is one reason for the difficulty in correlating theoretical results with experimental results.

In a trap-free material, the speed of response (i.e. the photocurrent rise or decay time) would be identical to the life-time of a free electron. With trapping centres, however, if, as in mostly the case, the density of trapped electron exceeds the density of free electrons, the rise and decay of the photocurrent is the time required to free an electron from a trapping centre.

In between these two extreme cases, a multitude of proportions of trapped/free electron ratios exist so that rise and decay times, as well as the interrelated magnitude of photocurrents, are difficult to predict with any sort of accuracy.

However because of this interrelation between rise and fall times and magnitude of photocurrents due to the total amount of free electron generated at any given illumination being constant, valid expressions for the rise and fall times can be obtained from the "output current vs. illumination" curves supplied by the manufacturers and shown in Figs.66, 67.

This is the basis for deriving the expressions for rise and decay times of photoconductors below.

A.10.1.3.4 Photosensitivity

Photosensitivity is a measure of the efficiency of conversion of light energy to the output electric current by the photoconductor.

There are several ways of expressing this, but for Vidicon applications, "Photoconductor Gain, G," is used:

$$\text{Gain, } G = \frac{\text{Number of electrons passing between the electrodes}}{\text{Number of photons absorbed}}$$

This can be recognized as being equal to "Q", the "Quantum Efficiency" (see expression 5.20). It has been found that (336),

$$G = \frac{\text{Lifetime of free electron}}{\text{Transit time of free electron between electrodes}}$$

The value of "G" has been the main cause of controversy (332, 333, 337); it has been assumed for a long time that for Vidicon photoconductors at least, its value must be such that $G \leqslant 1$. Lately, after doubt has been expressed (332, 333) the value of $G < 10$ has been accepted as possible (337).

The "transit time" depends on the mobility of the electron i.e. on the motion of the electron (or holes) in an electric field; it depends on the field presentand on the material.

The "electron lifetime" depends on the density of the recombination centres, their "capture cross-section" and so on, again quantities characteristic of the material.

The gain "G", however, depends also on the electric field across the photoconductor, which means that either "V_s", the voltage producing that field, or the thickness of the photoconductor can be varied (see below).

Photosensitivity depends on the number of photons giving rise to a given photocurrent. The illuminating source must thus be standard for two photoconductors if a valid comparison is to be made, as then the number of photons for the same illumination is the same for both photoconductors at some given level of illumination. The standard source used in the $2870^{o}K$ Tungsten Filament source. For non-standard illuminants, the required conversion factors can be found as in Appendix A.8.2.

This brief introduction to photoconductivity will now be extended specifically to Vidicon photoconductors to see how their requirements are met and whether their performance can be improved as regards our specific application.

A.10.2 VIDICON REQUIREMENTS FOR PHOTOCONDUCTORS

A.10.2.1 Introduction

The parameters of the photoconductor target in a Vidicon are not solely determined by the photoconductive properties themselves but also by the external sub-systems with which the photoconductor interacts, namely:

(i) the input stage of the video amplifier

(ii) the signal plate voltage "V_s" setting up an electric field across the photoconductor

(iii) the scanning electron beam discharging the stored charge on the photoconductor surface

(iv) the glass faceplate

(v) the electric contact between the signal plate and the photoconductor.

Examining the effect of these one by one, the apparently restrictive parameters of a Vidicon concerning output current signals, tube dimensions etc., will become apparent.

A.10.2.2 Signal Current Consideration

The main requirement in the Vidicon Camera is the ability to detect the output video signal at the signal plate output; video amplifiers are thus required. For a typical video bandwidth of

5 MHz, well designed amplifiers inject an RMS noise current of some 2.10^{-3} µA (262).

As a S/N ratio of some 100:1 is required for an acceptable noise free image when viewed on a CRT (see section 5.7.3), an output signal current "i_s" of some 0·2 µA is required (in fact a maximum current of some 0·5-0·6 µA is allowed for).

A.10.2.3 Phototarget Capacitance

The built up surface charges, "Q", on the beam side of the photoconductor due to photocurrent flow, have an associated surface potential "ΔV" associated with "Q", given by "$\Delta V = Q/c_m$", where "c_m" is the capacitance associated with the location where "Q" is stored. As mentioned in section 5.5.2, surface potentials "ΔV" cause beam bending and thus uneven output signals, and due to added deflection, cause loss of resolution. For acceptable output signals "ΔV" should be kept below 10-12V.

The stored change "Q" is restored by the scanning electron beam in a beam dwell time "T_d", which for the 625-line system is of the order of 0.1 us.

The output signal "i_s" is given from

$$i = c.\frac{dV}{dt}$$

Thus

$$i_s = c_m.\frac{dV}{dt} = c_m.\frac{\Delta V}{T_d}$$

Whence, assuming total charge discharge,

$$c_m = \frac{i_s \cdot T_d}{\Delta V} \quad \ldots\ldots\ldots \quad A.10.1$$

Substituting $\Delta V_{max} = 12$ V, $i_s = 0{\cdot}2$ µA, $T_d = 0{\cdot}1$ µs

gives $$c_{m\ min} = 1{\cdot}7.10^{-15}F \quad \ldots\ldots\ldots \quad A.10.2$$

As there are some $4{\cdot}5.10^5$ locations on a photoconductor scanned target in a 625-line system "C_{min}", the maximum phototarget capacitance is

$$C_{min} = 4{\cdot}5.10^5 . c_{m\ min} = 760 \text{ pF} \quad \ldots \quad A.10.3$$

A.10.2.4 Phototarget Thickness

During scanning, the electron beam is deflected, implying non-orthogonal beam landing (unless corrective action is taken (section 9.3.6)); this leads to uneven output signals. These can be tolerated if the deflection angle is kept to below 6-7°. Thus either the Vidicon scanned area is reduced or else the length of the deflecting field is increased. Scanned areas cannot be lowered below 1 cm diameter (333), as difficulty is experienced in obtaining a small enough beam diameter for acceptable resolution. Nor is the tube size increased to lengthen deflecting fields, due to tube and camera size considerations (these could be relaxed in our requirements). Consequently the photoconductor target is usually made 1" in diameter, the central $\frac{3}{8}$" x $\frac{1}{2}$" area (0·96 x 1·28 cm) being the scanned area. Its area is 1·23 cm^2.

For the above minimum scanned-area capacitance, the maximum photoconductor thickness can be deduced. The capacitance of a plane area is given by

$$C = \frac{\epsilon . \epsilon_o . A}{d} \text{ Farads}$$

where "ϵ_o" is the free space permittivity and equal to $8 \cdot 85.10^{-14}$ F/cm.

"ϵ" is the relative permittivity of the photoconductor

"A" is the area, equal to $1 \cdot 23$ cm^2 for the Standard Vidicon

"d" is the photoconductor thickness.

Hence substituting for C_{min} = 760 pF, results in

$$d_{max} = \frac{8 \cdot 85.10^{-14} . 1 \cdot 23 . \epsilon}{760 . 10^{-12}} \quad \ldots\ldots \text{A.10.4}$$

For Sb_3S_2, the usual Vidicon photoconductor, "ϵ" is typically 12, making

$$d_{max} = 17.10^{-4} \text{ cms} \quad \ldots\ldots\ldots \text{A.10.4(a)}$$

Recently, however, new methods of desspositing Sb_3S_2 have been implemented, in the presence of some gaseous atmosphere, which result in a "spongy" or porous mass (338) of better photoconductive properties. The layer being partly free space, partly Sb_3S_2, has a relative permittivity "$1 < \epsilon < 12$". A typical value of $\epsilon = 5$, gives then,

$$d_{max} \approx 7.10^{-4} \text{ cm} \quad \ldots\ldots\ldots \text{A.10.4(b)}$$

The Vidicon under consideration was quoted by the manufacturers to have a thickness between "$3 \rightarrow 6.10^{-4}$ cm".

A.10.2.5 Maximum Capacitance of Phototarget

The maximum capacitance (and hence minimum thickness) is determined by the fact that the scanning electron beam has a finite impedance "R_b" at low surface potentials "ΔV" (see Appendix A.7.2) which is of the order of 10^6-$10^7 \Omega$. Low surface potentials of some "$\Delta V \approx$ 2-3 V" are typical in normal TV applications. This impedance in conjunction with the element capacitance "c_m" form a delay time constant which must be much smaller than the beam dwell time "T_d", otherwise this "capacitive lag" would cause incomplete charge erasure and thus "smearing" of the display. The requirement is thus

$$c_m R_b < T_d \qquad \text{A.10.5}$$

whence on substitution for "R_b" and "T_d"

$$c_m < \frac{0{\cdot}1 \times 10^{-6}}{10^7}$$

$$< 10.10^{-15} \text{ F} \qquad \text{A.10.5(a)}$$

This sets an upper limit to the target capacitance of 4500 pF and a minimum thickness of some 10^{-4} cm for the photoconductor.

Typically Vidicons have a scanned area capacitance of 1500-2000 pF, and thicknesses typically of the order of $3 \rightarrow 6.10^{-4}$ cm. Considering that only two or three materials, Sb_3S_2, Se and PbO, are suitable as photoconductors in the visible spectrum, and that for efficient light utilization all the incident illumination should be absorbed, it can be seen that with the restricted range of thicknesses, high efficiencies are difficult to obtain.

A.10.2.6 Phototarget Resistance

Yet another requirement is the ability to "store" incident illumination to increase its sensitivity. The incident illumination need be integrated during the interval between scanning intervals to most efficiently utilize the incident illumination. "Integration" immediately brings to mind long time constants. The requirement is that the resistance, "r_m" associated with each scanned area location together with the capacitance of that location "c_m" resulting in a time constant, be much greater than the frame or scanning interval "T_F", which is equal to 40 ms.

Thus $$r_m . c_m > T_F \quad \ldots\ldots\ldots\ldots \quad A.10.6$$

and if the number of scanned area locations is "m" (equal to $4.5.10^5$), then as the total target capacitance

$$C = m\,c_m$$

and the total target resistance

$$R = \frac{r_m}{m}$$

then $$r_m \cdot c_m = RC$$

Hence it is required that

$$RC > 40.10^{-3} \quad \ldots\ldots\ldots \quad A.10.6(a)$$

or $$R_{AV} > \frac{40 \,.\, 10^{-3}}{C_{AV}}$$

$$> 3.10^7\,\Omega \text{ , for } C_{AV} \approx 1500 \text{ pF} \quad . \quad A.10.6(b)$$

This can be expressed as a resistivity "ρ". For any sample

$$R\,C = \rho \cdot \epsilon \cdot \epsilon_o$$

hence $$\rho_{AV} > 10^{11}\ \Omega\,\text{cm} \qquad \ldots\ldots\ \text{A.10.6(c)}$$

From solid state physics considerations (333) it can be shown that materials with band gaps of 1.8 eV have a resistivity of $10^{13}\ \Omega$ cm. The porous Sb_3S_2 deposit thus satisfies the requirement of resistivity admirably.

A.10.2.7 Phototarget Contacts

It is however found that the contacts with the photoconductor have a determining influence on the nett photoconductor resistance (333), lowering the effective resistance. The contacts in this case are the electron beam, which injects electrons into the layer, effectively lowering the resistance. The other contact is the signal plate providing "V_s", the signal plate voltage. Both contacts, being "injecting contacts", effectively lower the resistance and give rise to a space change limited current (339, 340).

In both cases a contact potential exists which must be overcome for electrons to be injected into the photoconductor. For the free electrons in the beam, this is equal to the band gap of 1.8-2 eV which has been accounted for and used in the beam discharge and potential stabilizing values in Appendix A.7.1-3 and sections 5.2-5.3. For the signal plate contact, depending on the conducting material (usually having a negligible "band gap") the contact potential is also of the same order.

It is the presence of these injecting contacts which has given rise to the controversy in determining the properties of photoconductors(see for example (333, 339, 340)), mainly because of the exact nature of the resulting "space charges" adjacent to the boundary layers; these alter the value of the gain "G".

A.10.3 EXPRESSIONS FOR PHOTOCURRENT, DARK CURRENT ETC.

A.10.3.1 Dark Current

Assume a voltage exists across the photoconductor (say due to "V_s"), giving rise to carriers which move towards the anode (thus depleting the beam side of the photoconductor by "Q" and giving rise to a surface potential "ΔV").

The current due to these carriers is thus

$$i = \frac{Q}{T_R} \qquad \text{. . . . . A.10.7}$$

where "T_R" is the transit time of the carriers across the photoconductor

"Q" is the charge leaked away from the beam side photoconductor surface (given by $Q = C.V_s$).

"T_R", the transit time is defined in (341) as,

$$T_R = \int_0^d \frac{dx}{\mu E}$$

where "d" is the photoconductor thickness

"E" is the field within the photoconductor due to "V_s"

"μ" is the carrier mobility (see below).

For a constant field (a homogenous photoconductor of constant thickness)

$$E = \frac{V_s}{d}$$

giving

$$T_R = \frac{d}{\mu E} = \frac{d^2}{\mu V_s} \qquad \text{A.10.8}$$

Substituting above in expression A.10.7 for "Q" and "T_R" gives

$$i = \frac{V_s^2 \cdot C \cdot \mu}{d^2} = K \cdot V_s^2 \qquad \text{A.10.9}$$

which is the "space charge" form dependence of the dark current. (The "space charge" analogy is because of its similarity to the space charge current between anodes and cathodes in a vacuum.)

A.10.3.2 Electron Mobility

The mobility is defined in (336) as,

$$\mu = \frac{1}{\rho \cdot n \cdot e} \qquad \text{A.10.10}$$

where "ρ" is the resistivity of the material

"e" is the electron charge

"n" is the number of free carrier/unit volume (which may or may not be trapped carriers).

Initially at low "V_s", the mobility is very small due to "trapping" - i.e. the number of free carriers giving rise to the current is small. As "V_s" increases more current is available, i.e. more carriers can fill the traps and be released from the traps due to thermal excitation and thus more are available for the current. Consequently the mobility is increased. The resultant dark current thus increases much faster than the above square-law form. This increase apparently makes "$i \propto V_s^{3-4}$" rather than "$i \propto V^2$" as predicted above.

A.10.3.3 Photoconductor Gain, "G"

A most important quantity defined previously, is the "gain of a photoconductor, G". This was defined as

$$G = \frac{\text{number of electrons passing through the photoconductor}}{\text{number of absorbed photons}}$$

This is clearly equivalent (337) to,

$$G = \frac{\text{lifetime of a free electron}}{\text{transit time of electron through photoconductor}} = \frac{\tau}{T_R}$$

as an electron under the action of a field travels across the photoconductor unidirectionally. The lifetime of a free electron does not end when it reaches the anode since a new electron takes its place at the cathode. Clearly "τ" can be longer than "T_R", resulting in $G > 1$.

From expression A.10.8 and from A.10.10

$$T_R = \frac{d^2}{V_s \mu} = \frac{d^2}{V_s} \cdot \rho n e \quad \ldots\ldots\ldots \text{A.10.10(a)}$$

For a photoconductor of thickness "d" and Area "A", the total number of free carriers is "$n.A.d \equiv N$", or "$n = \frac{N}{A.d}$".

Now "τ" is given from an expression obtained in (342) (which will not be derived here) as,

$$\tau_o = \tau \cdot \frac{N_T}{N} \quad \ldots\ldots\ldots\ldots \text{A.10.11}$$

where "τ_o" is the rise or decay time of photocurrent build up

"N_T" is a "quantity determined by the dynamical relationship between free and trapped electrons"

"N" is the total number of free carriers.

Substituting into "G" above, gives

$$G = \tau_o \cdot \frac{N}{N_T} \cdot \frac{V_s}{d^2 \cdot \rho\, ne} \quad \ldots\ldots\ldots \text{A.10.12}$$

$$= \tau_o \cdot \frac{N}{N_T} \cdot \frac{V_s}{d^2} \cdot \frac{1}{\rho \cdot e \cdot \frac{N}{A.d}}$$

$$= \frac{\tau_o}{N_T} \cdot \frac{A}{d.\rho} \cdot \frac{V_s}{e} \quad \ldots\ldots\ldots \text{A.10.12(a)}$$

Multiplying top and bottom by "C" the phototarget capacitance, and noting that "$\frac{d\rho}{A}$" is equal to the resistance of the phototarget "R", gives

$$G = \frac{\tau_o}{N_T} \cdot \frac{1}{RC} \cdot \frac{CV}{e}$$

$$= \frac{\tau_o}{RC} \cdot \left(\frac{Q}{e}\right) \cdot \frac{1}{N_T} \quad \ldots\ldots\ldots \text{A.10.12(b)}$$

as CV = Q, the charge present at the anode. Hence "$\frac{Q}{e}$" is the number of charges at the anode; this is called "N_A".

Hence

$$G = \frac{\tau_o}{RC} \cdot \frac{N_A}{N_T}$$

$$= \frac{\tau_o}{RC} \cdot M \quad \ldots\ldots\ldots \text{A.10.12(c)}$$

Depending on how "M", or rather its constituents are interpreted, is the basis of the problem of photoconductivity, its speed of response, and its gain in Vidicon phototargets.

"N_T" has been stated to be (337), "the total number of electron charges, free plus trapped, in thermal contact with the conduction band" (i.e. capable of being released from the traps into the conduction band due to random, thermal lattice agitation).

Hence "G" depends on the density and distribution of trapping centres (and "R", "C" etc.) which in turn depends on the structure and composition of the photoconductor.

As mentioned before, until recently it was thought that G" and particularly "M" could not exceed "1"; however values of "G" and "M" of the order of 5 have been reported (332). A theoretical study of trap distributions (337, 342) have indicated that values of "M" of the order of 10-20 or more are possible with certain trap distributions and certain backplate voltages "V_s". (In some CdS photoconductor samples, "M" as high as 10^2-10^3 has been found (335)).

Two classes of Vidicon photoconductors have been distinguished, called Class I and Class II photoconductors; they have different spectral response, different photosensitivity and different buildup and decay response (see Figs.66(a) and (b)).

There is good evidence to suggest that photoconductors of Class I (of which early Vidicons were manufactured) have M=1 (341). Indeed, because for several years this was the only type of photoconductor available, small wonder that "theory" showed $M \leqslant 1$ for "all" Vidicon photoconductors.

For photoconductors Class II, "M" has been taken to be of the form

$$M = a + M'b$$

where

$$a + b = 1$$

$$1 < M' < 10 ,$$

"M" being dependent on "V_s", the signal plate voltage. This has been taken from a study of the theoretically allowable values of "M" (342) and to explain lower values of response times "τ_o" than otherwise predicted if "M" had been kept to the usually accepted value of "1".

A study of the diagrams of the variation of "M" against "V_s" for various trap distributions (342) reveals that for low "V_s" increasing, "M" increases from "1" up to about "10-20". Abruptly at some $V_s > V_s'$, "M" drops back to a value of "1", thereafter remaining constant. This fact explains some of the values of the time constants of photocurrent buildup and decay, where the photoconductor trap distribution is such that $M > 1$ (i.e. Class II photoconductors).

A.10.4 DERIVATION OF PHOTOCURRENT BUILDUP AND DECAY TIME CONSTANTS

Above the Gain "G" was defined as

$$G = \frac{\text{number of electrons passing through the photoconductor}}{\text{number of absorbed photons}}$$

If the electron count is made within a unit time, then clearly

$$G = \frac{\text{photoconductor current}}{\text{"photon" current associated with incident illumination}}$$

$$= \frac{i_p}{K.E} \quad \ldots\ldots\ldots\ldots\ldots \text{A.10.13}$$

where "i_p" is the photoconductor current

"E" is the illumination on the photoconductor in foot-candles

"K" is a conversion factor giving the number of photons in a ft-C of illumination (see Appendix A.8.2.3).

In Chapter 4, expression 4.20 gave the equivalence between "i_p" and "i_s", the output Vidicon signal current as

$$i_s = \frac{1}{T_F}\int_o^{T_F} m.i_p(t)\,dt \quad . . \text{(expression 4.20)}$$

where "T_F", the frame time, is the interval between scans

"m", is the number of scanned area locations.

For a constant "$i_p(t)$",

$$i_s = m \cdot i_p$$

$$\therefore \quad G = \frac{m\, i_p}{m.K.E} = \frac{i_s}{m.K.E} \quad \ldots\ldots \text{A.10.13(a)}$$

where "m.K.E" is the number of photons falling on the total scanned area associated with illumination "E" ft-C in unit time.

"i_s" is the steady state output current at an illumination of "E" ft-C.

From the manufacturer's supplied data curves of "i_s vs. E", (see Fig.66(a)(b), 66(b)(a)), the relationship of "i_s" for any dark current or signal plate voltage "V_s", is clearly seen to be

$$i_s = I_o (E)^{\gamma} \qquad \text{. . . . . . . A.10.14}$$

where "I_o" is the signal current "i_s" at 1 ft-C of illumination.

"E" is the illumination in ft-Cs.

"γ" , the "gamma" is the slope of the curves, and is typically 0·6-0·7, but for high values of "E" (E > 100 ft-C) may drop to 0·4-0·5.

Hence substituting, $G = \frac{I_o (E)^{\gamma}}{mK \cdot E} = \frac{\tau_o}{RC} \cdot M$ . . . . A.10.13(b)

whence

$$\tau_o = \frac{R \cdot C \cdot I_o \cdot (E)^{\gamma - 1}}{mK \cdot M} \qquad \text{. . . . . . . A.10.15}$$

where "τ_o" is the photocurrent rise or decay time constant at an illumination of "E" ft-Cs

"R" is the dark current resistance at the signal plate voltage "V_s" at which the Vidicon is operated at and can be found from manufacturers curve of "dark current vs. V_s" (Fig.66(a)(a)) by measuring the slope of the curve. A somewhat lower value than this is used to allow for the effective lowering of the nett voltage across the photo-conductor due to charge buildup, and to take into account the dynamic resistance rather than the static dark resistance as changing current flows.

"C" is the scanned area capacitance, typically equal to 1500 pF.

"I_o" is the signal current at 1 ft-C illumination for the dark current corresponding to "V_s" at which "R" is measured.

"E" is the illumination in ft-Candles.

"γ" is typically 0·6-0·7, the slope of the "i_s vs. E" for the given dark current at which the Vidicon is operated and must be evaluated at the appropriate illumination "E".

"mK" is the number of photons in 1 ft-C of illumination. For 2870°K tungsten illumination it evaluates at $2{\cdot}6.10^{13}$ photons/sec.

"M" is a quantity characteristic of the photoconductor and the density and nature of traps therein. For photoconductors Class I, "$M \approx 1$". For photoconductors Class 2, "$1 < M < 10\text{-}20$" depending on "V_s" the signal plate voltage. It is given as $M = (a + M'b)$

where $a + b = 1$

$1 < M' < 10$, a function of "V_s".

It increases from "1" to "10-20" (theoretically) at low "V_s", and then for $V_s > V_s'$ decreases to "1" and remains constant.

Since 1 ft-Candle of 2870°K Tungsten Illumination is equivalent to $2{\cdot}6.10^{13}$ photons/sec (Appendix A.8.2.3) and as 1 Ampere = $6{\cdot}24.10^{18}$ electrons/sec, then 1 ft-Candle is equivalent to a "current of photons" of 4·16 uA $\approx 4{\cdot}2.10^{-6}$ A.

Hence the photocurrent rise and decay time constants are given by:

(i) For Photoconductor Class I

$$_{o} = \frac{R \cdot C \cdot I_o \, (E)^{\gamma - 1}}{4 \cdot 2 \, . \, 10^{-6}} \quad \text{seconds (expression 4.15)}$$

(ii) For Photoconductor Class II

$$_{o} = \frac{R \cdot C \cdot I_o \, (E)^{\gamma - 1}}{4 \cdot 2 \, . \, 10^{-6} \, (a + b M')} \quad \text{seconds (expression 4.16)}$$

All the quantities therein are directly obtainable from the manufacturers' curves, except for Class II photoconductors, where $(a + bM')$ lies between 1 and 10.

Table A.10.1 shows the calculated values for the above expressions for several Vidicons for which time constant data exists, namely for photocurrent decay. When comparing calculated time constants with those read off manufacturers' curves, only the initial decay is taken under consideration from the manufacturers' curves, since at low output signal, "capactive" or "beam discharge" decay become significant.

As seen from Table A.10.1, for photoconductors Class I, the agreement between calculated and experimental results in very close, agreement being within 10-15%! This is because $M \approx 1$. For Class II photoconductors as "M" is not known precisely, and more-over is varying with "V_s" (i.e. with the dark current "I_d") nothing definite can be said whether agreement is close or otherwise. However it is seen that for low "V_s", $M > 1$ while at high "V_s", $M \approx 1$ as expected (see Section A.10.3.3 above).

TABLE A.10.1

COMPARISON BETWEEN CALCULATED AND SUPPLIED TIME CONSTANTS FOR VARIOUS VIDICONS

	PHOTOCONDUCTOR CLASS I						PHOTOCONDUCTOR CLASS II				
Vidicon Type	RCA 8572A (1")			RCA 8051 (1 1/2")			8134	RCA 7262A		RCA 8507A	
I_d, in uA (1)	0.004	0.02	0.1	0.005	0.02	0.05	0.035	0.02	0.2	0.02	0.1
V_s, in Volts	18	36	62	18	32	45	35	29	51	24	42
R, in $10^8\Omega$ (2)	45	18	6.2	35	9	4.5	6	10	1.7	7.7	2
C, in pF.	Nominal value of 1500 pF. used for all Vidicons.										
I_o, in uA.	0.018	0.07	0.19	0.042	0.12	0.22	0.27	0.20	0.50	0.20	0.40
E, in ft-Cs	96	9.8	2.8	12	2.5	0.88	1.2	1.8	0.6	2.0	0.6
γ, "gamma"	Average value of 0.65 used for all Vidicons.										
M, "gain"	Equal to 1 for Class I Vidicons						See below				
τ_o, (ms.) (calc.	5.3	19.4	28	22	28	37	52	58	37	43	35
τ_o, (ms.) (measd.	9	21	30	19	31	36	24-43	15	26	15	25
T_o, (calc.)	29	45	41	53	38	35	55	72	30	55	29
T_o, (measd.)	49	49	44	45	42	35	25-45	18	22	19	21
% variation	-40%	-8%	-7%	18%	8%	0	$M\approx2$	$M\approx3-4$	$M\approx1$	$M\approx3$	$M\approx1$
Form used	$\tau_o = 40.\ (E)^{-0.35}$ ms.						$\tau_o \approx 25.\ (E)^{-0.35}$ ms.				

(1) Id, the "dark" current.

(2) The value of "R" is taken less than its static value to allow for the reduction in the "effective" V_s due to charge buildup on the photoconductor and to introduce some dynamic resistance effects as the current is changing.

"Measured" Time Constants τ_o, (from data sheets) are for a signal current $\approx$ 0.3uA.

From these results the time constant for photoconductor Class I is taken as,

$$\tau_o = 40 \cdot (E)^{-0.35} \text{ ms.} \quad \text{(expression 4.15(a)}$$

$$= T_o \, (E)^{\gamma - 1}$$

while for photoconductor Class II it is taken as (mainly from manufacturers data),

$$\tau_o = 25 \cdot (E)^{-0.35} \text{ ms.} \quad \text{(expression 4.16(a)}$$

$$= T_o \, (E)^{\gamma - 1}$$

As a further check, Redington (252) has given a photo-current decay curve for a photoconductor Class I, illuminated by a mercury flash lamp. From manufacturers data (279), the luminance of such a lamp is of the order of $10^5 - 10^6$ ft-L giving an illumination "E", of the same order. The experimental time constant graphed indicates a $\tau_o \approx 0.45$ ms. Evaluating, using the above formula 4.15(a), gives a $\tau_o \approx 0.23$-0.57 ms. Hence again agreement is very close, considering that no exact figures are specified.

A.10.5 EVALUATION OF PHOTOCURRENT BUILDING AND DECAY

The general iterative expression for photocurrent buildup (and decay) at any given illumination "E", is given in chapter 4 by expression 4.14(a) as

$$i_s(t + \Delta t) = i(t) + \left(i_{ss}\left(t + \frac{\Delta t}{2}\right) - i(t)\right) \cdot \frac{\Delta t}{\tau\left(t + \frac{\Delta t}{2}\right)}$$

. . . . . .(expression 4.14(a))

where "Δt" is the small iteration time-interval during which the illumination "E" is considered to be constant. This is read off the W-phosphor persistance curve, scaled by the appropriate peak illuminance operating at the CRT screen. Fig. A.9.2 shows the % of peak CRT luminance and the "Δt" iteration time increments used in these calculations.

"$i_s(t+\Delta t)$" is the required output signal at time "$t+\Delta t$".

"$i(t)$" is the calculated current evaluated at the prior time "t", i.e. "Δt" secs. prior.

"$i_{ss}(t+\frac{\Delta t}{2})$" is the "steady state" current for the given illumination "E" acting at time "$t+\frac{\Delta t}{2}$" and is read off the manufacturers "i_s vs. E" curves (see Fig.66(a)(b), 66(b)(a)).

"$\tau(t+\frac{\Delta t}{2})$" is the time constant evaluated at the illumination "E" operative at time "$t+\frac{\Delta t}{2}$". The expressions 4.15(a), 4.16(a) are used for this.

The above expression is used to obtain the curves in Fig.67(a), (b), for the various peak values of "E" shown.

For the initial current to obtain the signal output without prior illuminance at that given location, the current at t = 0 is the dark current corresponding to the operating "V_s", the signal plate voltage.

A sample calculation for the first few iterations is given on Table A.10.2. This is for a photoconductor Class II

TABLE A.10.2

SAMPLE EVALUATION OF PHOTOCURRENT BUILDUP DUE TO W-PHOSPHOR ILLUMINATION—NO PRIOR ILLUMINATION

Photoconductor Class II RCA 7262A E peak=600 ft-candles I dark=0.01 ua $\tau_o=25(E)^{-0.35}$ ms.

t (ms.)	Δt (ms.)	$E(t+\frac{\Delta t}{2})$ (ft-cs)	$\tau(t+\frac{\Delta t}{2})$ (ms.)	$\frac{\Delta t}{\tau}$	$i_{ss}(t+\frac{\Delta t}{2})$ (uA)	i(t) (uA.)	$i_{ss}-i(t)$	$(i_{ss}-i(t)) \times \Delta t/\tau$	$i(t+\frac{\Delta t}{2})$ (uA)
0	0.03	480	2.23	0.0129	5.08	0.01	5.07	0.0655	0.0755
0.03	0.03	336	2.55	0.0118	4.20	0.0755	4.12	0.0485	0.124
0.06	0.04	264	2.90	0.0138	3.64	0.124	3.52	0.049	0.173
0.1	0.1	192	3.40	0.0294	2.97	0.173	2.80	0.082	0.255
0.2	0.2	144	3.82	0.0524	2.54	0.255	2.28	0.120	0.375
0.4	0.3	96	4.4	0.0683	2.04	0.375	1.665	0.113	0.488
0.7	0.3	78	5.0	0.060	1.84	0.488	1.35	0.081	0.569
1	1	49.5	6.0	0.167	1.40	0.569	0.83	0.139	0.708
2	2	27	7.1	0.282	0.97	0.708	0.26	0.073	0.781
4	3	15.6	9.0	0.33	0.70	0.781	-0.08	-0.026	0.755
7	3	9.6	10.8	0.313	0.51	0.755	-0.245	-0.077	0.678
10	5	5.7	13.0	0.384	0.365	0.678	-0.313	-0.120	0.558
15	5	3.5	15.6	0.32	0.262	0.558	-0.294	-0.094	0.464
20	5	2.3	18.4	0.271	0.20	0.464	-0.264	-0.0715	0.393
25	5	1.8	20.0	0.25	0.175	0.393	-0.218	-0.055	0.338
30	5	1.32	22.7	0.22	0.144	0.338	-0.194	-0.043	0.295
35	5	1.08	24.4	0.205	0.126	0.295	-0.169	-0.035	0.260
40	0.03	480	2.23	0.0129	5.08	0.260	4.82	0.062	0.322
40.03	0.03	336	2.55	0.0118	4.20	0.322	3.88	0.038	0.360
40.06	0.04	264	etc.	etc.	.....	.....	.....	....	..
etc.	etc.	....	...	Next frame illumination cycle				...	....

Facing p. A. 159

Vidicon (RCA 7262A), with a dark current at 0.01 uA ($V_s \approx +15 \rightarrow +32$ V) and a peak incident illuminance of 600 ft-C. It is for the first frame time illumination for curve "A" shown plotted in Fig.67(a).

. . . .

APPENDIX 11

MEASUREMENT OF DISPLAY DISTORTION AND VIDICON SCANNED AREA DISTORTION : MOIRE FRINGES

APPENDIX 11

MEASUREMENT OF DISPLAY DISTORTION AND VIDICON SCANNED AREA DISTORTION : MOIRE FRINGES

A.11.1 INTRODUCTION

The maintenance of a steady display in "VIDIOGRAPHIC" requires that at each frame-time, the display locations are relocated to result in a geometrically linear or distortion-free display. The amount of this relocation is derived from the stored Graphical Distortion Correction Information. This Correction Information in turn, must be specified to a higher degree of precision that the correction requiring to be implemented. The measurement of display distortion from which this correction information is derived must itself be yet more precise. As the correction to be implemented is to a precision where the remanent distortion after correction is to be less than 0.25% of the display height (or width) (expression 7.13), the precision of measuring distortion is to be of the order of 0.1% of the display height (or width).

As current methods of measuring distortion are cumbersome, tedious, and not over accurate (unless precise but very expensive electronic and optical equipment is used), techniques of measuring display and TV camera distortion using Moire Patterns had to be developed and implemented. ("Had to be" because without accurate knowledge of the distortions present, corrections cannot be implemented and no steady display would result in the VIDIOGRAPHIC system; while lack of funds precluded the other alternative method). As it turned out, the resultant two distinct methods of obtaining display distortion using Moire Patterns generated distortion information which were in almost "correct" form for encoding graphically within the Graphic Distortion Correction Information storage.

It would be pointless to present again within this Appendix the theory, derivation of expressions, method of implementation and results of the Moire Pattern Methods of obtaining distortion, when a fairly comprehensive and self-contained paper has been published on the subject. Moreover the calculations given therein and the illustrations pertain to the Vidicon and the PYE TV Monitor under test, which have been mentioned throughout this thesis. Hence a reprint of the paper is included in the back-cover pocket of the thesis; it must be taken as a direct section of the thesis, as its contents are not reproduced elsewhere within the thesis. Except for some general remarks about the universality of the Moire Fringe Methods in measuring distortion of all classes of displays, the paper specifically refers to the measurement of distortion of a TV-raster scanned CRT and a TV-raster scanned Vidicon camera.

A very short resume of the two methods developed for measuring distortion is given as well as an explanation how the distortion results obtained by those two methods give the necessary information to be used in the Graphic Distortion Correction Information Storage described in Sections 7.3-7.5.

A.11.2 DISTORTION CONCEPTS

Two "distortion" concepts must be distinguished when measurements to obtain distortion are performed. If the display area under measurement has a horizontal dimension "H", and a vertical dimension "V", and if it has "N_H" discrete addressable locations in the Horizontal direction and "N_V" discrete addressable locations in the Vertical coordinate, then ideally

(i) each horizontal coordinate is spaced "$\frac{H}{N_H}$" units apart

(ii) each vertical coordinate is spaced "$\frac{V}{N_V}$" units apart.

Now it can be imagined that on the display an imaginary grid can be drawn, the vertical grid line spacing being "$\frac{H}{N_H}$" units and the horizontal grid line spacing being "$\frac{V}{N_V}$" units.

If in a real display, all the addressable display locations are displayed at intersections of these imaginary grid lines, the display area can be considered to be distortion-free or the display has complete geometrical linearity.

A.11.2.1 "Location-to-Location" Distortion

In practice, due to distortion, displayed locations do not fall on these grid line intersections but in-between them.

Consequently the actual spacing between adjacent addressable locations is no longer "$\frac{V}{N_V}$" or "$\frac{H}{N_H}$" units but,

(i) for the norizontal coordinates, the adjacent location spacings are given by

$$\frac{H}{N_H}\cdot(1 + F_H\,(x,y)) \quad \text{.....................} \quad A.11.1$$

where "$F_H(x,y)$" is the percent deviation from ideal horizontal location spacing.

and

(ii) for the vertical coordinates, the adjacent location spacings are given by

$$\frac{V}{N_V}.(1 + F_V (x,y)) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad A.11.1(a)$$

where "$F_V(x,y)$" is the percent deviation from ideal vertical location spacing.

The form "$F(x,y)$" is used to indicate that at any constant vertical coordinate "y", say, "$F_V(x,y)$" may still be a function of "x", the horizontal coordinate. This is to say that in a raster-scanned display for example, adjacent raster lines are not parallel over their length. The same is valid for the horizontal distortion.

The distortion "$F_V(x,y)$" or "$F_H(x,y)$" may be called the "Location-to-Location" Distortion.

A.11.2.2 "Relative" or "Cumulative" Distortion

The second form of distortion is that commonly used in display distortion measurements and which has been defined in section 2.2.2.7. It is that distortion which has been used in determining the Display Correction in Chapters 7-8.

Referring to Fig. 17(a), if at some given horizontal coordinate, a display point ideally is to be displayed at "y=A" (measured from some origin "y=0"), but in fact due to Vertical Distortion, it is displayed at "y=A'", then Vertical distortion, "E_A" at point "A'" is

$$E_A = \frac{OA' - OA}{V} = \frac{\Delta A'}{V} \quad \text{.................... A.10.2}$$

where "V" is the display area height.

As within the Vertical interval "OA'" there are several vertical location spacings, each subject to its own vertical "Location-to-location" Distortion "$F_V(x,y)$", the distortion "E_A", called the "Vertical Relative Distortion" (or "Cumulative Distortion") is due to a cumulative effect of individual "Location-to-location" Distortions.

A.11.3 MOIRE PATTERN METHODS FOR MEASURING DISTORTION

Two separate methods for measuring Display Distortion using Moire Patterns are reported in the attached reprint of the published paper. Each has its advantages as far as encoding the distortion in a form suitable for Graphic Correction Information goes.

The two methods are:

(i) "The Hyperbolic Moire Fringe" (H.M.F) Method
(section IV in the reprint)

(ii) "The Linear Moire Fringe" (L.M.F) Method
(section V in the reprint)

The names have been coined due to the characteristic shape of the resultant Moire Patterns.

Briefly each of the two methods gives:

(i) The H.M.F. Method gives either "$F_V(x,y)$" directly, or "$F_H(x,y)$" directly. The two distortions must be measured and obtained separately.
A particular Moire fringe or "line" traces out "$F_V(x,y)$" or "$F_H(x,y)$" directly on a linearly calibrated field.

The "calibration" and "scale" on which the magnitude of this distortion is read-out depends on the line-spacing of the "reference line grid", which is the reference pattern which interacts with the distorted display pattern resulting in the Moire Pattern. The display area can itself be made this "calibrated field", and thus a photograph of the resultant Moire Pattern, can be used as direct record of the "Location-to-location" Distortion "F(x,y)" for one of the two co-ordinates. The centre or "zero-th" Moire fringe traces out this distortion. A separate photo-record must be obtained to show the distortion in the other coordinate.

In our case, as will be shown below, only the vertical distortion is of interest.

Fig. 6 (in the reprint) shows the Vertical Distortion trace for the AW 36-48 PYE Monitor, showing a "$F_V(x,y)$" of some 1.3% over 90% of the scanned area.

Fig. 15 (in the reprint) shows the Vertical Distortion trace of the Vidicon Camera, showing some 2.5% distortion over 90% of the scanned area. (These values are for CRTs and Vidicons with uncorrected barrel or pincushion distortion.

(ii) The L.M.F. Method gives a resultant Moire Pattern from which the "Relative Distortion" or "Cumulative Distortion" "E_A" can be derived simultaneously for both the horizontal and the vertical directions.

This is an advantage over the other method, as only one measurement and one photograph lead to this result; this however is at the cost of having to perform some elementary calculations (section V.C in reprint).
Fig. 10 (in the reprint) show typical CRT display records with this form of Moire Patterns. Similar photographs show Vidicon generated Moire Patterns.
The elementary calculations in reducing the Moire Pattern information to Distortion Information are done at selected locations on the display area. Joining the points where the distortion is identical leads to lines of equal distortion, or "distortion contours or lines", and hence, for the whole scanned area, to "Distortion Contour Maps", mentioned and shown in Figs. 101, 104 of the thesis. On such Distortion Contour Maps for the Horizontal direction is based the Horizontal Coordinate Correction for both the CRT and the Vidicon. Consequently, for our purposes, the simple calculations need only be done to obtain H-coordinate Distortion Maps. (Vertical-coordinate Distortion Maps would have much the same appearance, if they were required also). As mentioned previously, the L.M.F. Method gives the "Cumulative Distortion" from such calculations.

A.11.4 GRAPHIC DISTORTION CORRECTION INFORMATION FROM MOIRE PATTERN MEASUREMENTS

A.11.4.1 Horizontal Correction Information

In the H-coordinate correction, due to the relatively limited amount of distortion information capable of being stored and read-out, the required corrections to be implemented are considered to be bounded by 0.25% Distortion Contours; within the region bounded by two adjacent contours, the distortion is considered to increase or decrease linearly. For the distortion within each such region, a given distortion contour is considered to be a local "distortion origin" and the distortions within that region can thus be considered to be "cumulative distortions". Consequently the H-distortion maps, specifying "relative" or "cumulative distortions", are the ideal way of graphically encoding the necessary Distortion Correction Information.

A.11.4.2 Vertical Correction Information

For Vertical Distortion Correction however, each horizontal raster line (i.e. each vertical coordinate) is scanned in its entire length, including that region of the scanned area reserved for Distortion Correction Information Storage. Consequently Distortion Information may be made available every vertical coordinate and hence correction can be made on a "per every coordinate" basis i.e. correction can be made from Vertical "Location-to-location" Distortion and intuitively the "Location-to-location" Distortion would be expected to enter the specification of Distortion Correction Information. This is indeed the case.

In section 7.5.7, the requirement for the "Vertical Distortion Line", specifying the Distortion Correction, is given.

The requirement there however is based on the Vertical Distortion Line being derived from a Cumulative Vertical Distortion

Contour Map, analogous to the Horizontal Distortion Map shown in Figs. 101,104. Since it is assumed that pincushion or barrel distortion is made negligible or is eliminated prior to this, the resultant Vertical Distortion Map would essentially consist of horizontal distortion regions. For any two adjacent horizontal raster lines, a vertical correction would apply equally well along the whole length of the raster lines (see section 9.2.1). (A pincushion or barrel distortion-free Horizontal Distortion Map would essentially consist of vertical distortion regions, rather than the distortion regions as shown in Fig 101, in which pincushion distortion was not eliminated).

The point of all this is that a "cut" vertically across the display, would specify a Vertical Distortion Trace, which would be identical at all other vertical "cut" locations. Hence only one "Vertical Distortion Trace" running the whole vertical length of the display area, is required to specify the display Vertical Distortion.

In section 7.5.7 this Vertical Distortion Trace is called "fV(n)", where "n" is the raster line count and hence is indicative of the vertical coordinate where this Vertical Distortion is measured at. It is defined there as

$$fV(n) = \frac{\Delta Y(n)}{Y} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{A.11.3}$$

where "Y" is the display area height (identical to "V" in the previous expression A.11.1(a);

and "$\Delta Y(n)$" is the Cumulative Distortion at the "nth"-raster line.

Now for the actual implementation of generating a linear display, a current staircase waveform feeds the vertical scanning coils in both the CRT and the Vidicon. Each current step height "ΔI" corresponds to the beam deflection equal to the distance on the scanned area between two adjacent vertical coordinates, "Δy", such that

$$\Delta y = K.\Delta I \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(expression 7.14)}.$$

where "K" is a constant.

For a linear display this is to be true for all vertical coordinates.

Due to display distortion being present, the spacing "Δy" is no longer constant but given by "$\Delta y + \delta y$" where

$$\Delta y + \delta y = K.\Delta I \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{A.11.4}$$

Hence for constant "Δy", giving a linear display

$$\Delta y = K.\Delta I - \delta y \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{A.11.5}$$

$$= K(\Delta I - \delta I) \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{A.11.5(a)}$$

i.e. a small correction current "δI" need be applied at each staircase step to generate a linear display.

Expression 7.8, in section 7.5.7 states that this correction current at the "n-th" step, is given by

$$K.\delta I = Y.f'V(n) \quad \ldots\ldots\ldots\ldots\ldots \text{(expression 7.8)}$$

where "$f'V(n)$" is the derivative of the Cumulative Distortion trace at the "nth" raster line.

In the reprint of the paper, the Vertical Relative Distortion or Cumulative Distortion given by the Vertical Distortion Trace "$fV(n)$", is given in expression 32 by

$$f\,V(n) = \frac{\sum_i (D')^i_V \cdot F^i_V(x,y)}{H} \quad \ldots\ldots \text{(expression 32)}$$

where "H" is the display height (identical to "Y")

"$(D')_V^i$" is a vertical distance within which the average Vertical "Location-to-location" Distortion is given by "F_V^i (x,y).

This interval "$(D')_V^i$" is that specified by two consecutive Moire Fringes (see section $\overline{V}$.C in the reprint).

Now it makes no difference in the L.M.F. Method to the magnitude of the distortion "fV(n)", no matter how small or large the intervals "$(D')_V^i$" are made between the Moire Fringes. For practical reasons of ease and accuracy of measurement, they are made to be of the order of 1", but can equally well be made equal to "Δy" the raster line vertical spacing.

Hence, equally well,

$$fV(n) = \frac{\sum_{1}^{n} \Delta y . F_V(x,y)}{Y} \quad \text{.............. (expression 32(a))}$$

where "$F_V(x,y)$" is the Location-to-location Distortion.

Clearly then (taking derivatives),

$$f'V(n) = \frac{\Delta y . F_V(x,y)}{H} \quad \text{...................... A.11.6}$$

and substituting in expression 7.8 above

$$K.\delta I = \frac{Y}{H} . \Delta y . F_V(x,y) \quad \text{...................... A.11.7}$$

or $$\delta I = F_V(x,y).\Delta I \quad \text{......................... A.11.7(a)}$$

as $Y \equiv H$.

Substituting in expression A.11.5(a) this results in

$$\Delta y = K.\Delta I.(1 - F_V(x,y)) \quad \text{.................. A.11.8}$$

If the maximum "$F_V(x,y)$" is labelled "$F_V(x,y)max$", and if the basic current staircase step "$\Delta I'$" is made equal to

$$I' = \Delta I \left(1 - F_V(x,y)_{max}\right) \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{A.11.9}$$

then the correction current "$\Delta I(n)$" to be added to "$\Delta I'$", the "nth" raster line, is given by

$$\delta I(n) = \Delta I . \left(F_V(x,y)_{max} - F_V(x,y)\Big]_n \right) \quad \ldots\ldots \text{A.11.10}$$

It was indicated previously that the H.M.F. Method gave "$F_V(x,y)$" directly, the "zero-th Moire Fringe" being "$F_V(x,y)$".

Hence if a vertical line is drawn through "$F_V(x,y)_{max}$" then the horizontal interval between it and "$F_V(x,y)$" specifies the correction current required directly (except for a scaling factor).

Reproducing this "$F_V(x,y)$" trace directly from the H.M.F. Method within the CRT scanned area interval of "1 us" lengths, "T_{V-VID}" and "T_{V-CRT}" (see Fig 105), resulting thus in the "Vertical Correction Line" then, during scanning, the interval defined by one of the edges of this interval, "T_{V-VID}" or "T_{V-CRT}", and the Vertical Distortion Trace, specifies a time interval directly proportional to "$F_V(x,y)$" and hence can be used directly to generate the correction current "$\delta I(n)$".

Thus the H.M.F. Method is recommended to obtain the Vertical Distortion Correction Information as the "zero-th" Moire Fringe directly specifies the Vertical Correction Line for either the CRT or the Vidicon Camera.

A.11.4.3 Conclusions

The two distinct Moire Fringe Methods to obtain CRT Display Distortion and Vidicon Scanned-Area Distortion, have each, from our viewpoint of obtaining Distortion Correction Information, distinct advantages. These are:

(i) To obtain the Vertical Distortion Correction Line, the H.M.F. Method should be used. It gives this V-correction line directly.

(ii) To obtain Horizontal Coordinate Correction Data, the L.M.F. need be used. Simple calculations on various distances between Moire Fringes on the one photo-record of the Moire Fringes, result in Distortion Maps, which are the necessary form of Horizontal Distortion Correction Information.

A.11.5 REPRINT ON MOIRE FRINGES

See inside of back cover.

. . . .

APPENDIX 12

TYPICAL TV—CRT AND VIDICON CAMERA SYSTEMS

APPENDIX 12

TYPICAL TV—CRT AND VIDICON CAMERA SYSTEMS

A.12.1 INTRODUCTION AND REQUIREMENTS

The block diagram and circuit diagram of the available TV-CRT display on loan and the purchased Vidicon Camera are given here primarily to illustrate typical commercially available equipment and to indicate the suitability or otherwise for using them or parts of them in "VIDIOGRAPHIC".

Only the Video Amplifier subsystem and the H- and V-Deflection Subsystems are given; power supplies etc are omitted, although in "VIDIOGRAPHIC", power supplies need be well regulated, particularly that part supplying the H- and V- Deflection System. Current regulation of the order of 0·1% or better is required.

The requirements which the commercially available CRT and Vidicon camera satisfy are quite different to the requirements of "VIDIOGRAPHIC".

(i) CRT Requirements

(a) A video bandwidth from 50Hz to some 4·5-5MHz is set by broadcast requirements; the received video signal requiring amplification can contain signals with frequency components within this range. The amplifier is required to be linear as it is required to amplify continuously varying amplitude signals corresponding to the halftones of a live or film scene being displayed. To eliminate signal distortion, and perhaps display element shape distortion, a flat response (ie to within

2-3dB) of the above bandwidth is required, as well as a linear phase characteristic.

(b) the display distortion present may be of the order of 3-5% or more. However the nature of the scenes televised are such that distortion is rarely visibly noticeable by the viewer.

For typical CRTs of usual size eg 10"-15" linear display area dimension, and with 90°-110° deflection angle, peak to peak scanning currents of the order of 200 - 600mA are typical. For the relatively low deflection angle in Projection CRT tubes (of the order 50° max for the diagonal deflection angle) (259)), the scanning current amplitudes are somewhat less than the above figures.

(c) The timing in a commercial TV-CRT receiver is usually provided by synchronysing signals transmitted with the Video signals and separated and decoded by the "synchronising circuits" in the receiver. These, in a 625-line, 50Hz system, generate the field timing signal to provide vertical scanning waveforms at a repetition rate of 50Hz and horizontal line scanning-waveforms at a repetition rate of 15·625 KHz.

(ii) Vidicon Camera Requirements

For the Vidicon camera, in closed-circuit or broadcast applications, the requirements are very similar to the TV-reciver, since the Vidicon camera generates the signal for the TV reciver and hence both must be mutually compatable. Normally, though, it is required for broadcast applications that video signal and display linearity be higher in the camera system (the transmitter) than in the TV-CRT display (the receiver). Thus the video bandwidth is at least 5-6 Mhz while the geometrical linearity is some 1-2%. However in industrial application cheap TV cameras, such as the Vidicon camera, are completely adequate; the requirements, particularly as far as displcy linearity and timing go, can thus be very much relaxed. Thus linearity may be no better than 4-5% while timing may be even completely unsynchronized.

In Closed-circuit TV applications, system timing for the camera and the TV-CRT display may be one of the following methods:

(1) an external timing and synchronizing source controlling the TV-camera Scanning Subsystems, resulting in a composite video signal containing both the video information and synchronizing signals as in the broadcast TV case (except that coaxial cable is used for transmission).

(2) the external source controls both the TV reciever and camera directly by separate synchronizing signal feeds, with the camera output consisting of only the video information signal.

(3) in the cheapest solution, most often encountered in CCTV, the camera has a simple free-running oscillator, nominally at H-line frequency of 15·625KHz, while the mains frequency sets the Vertical or field frequency at 50 Hz. The free-running oscillator randomly, varying ("jittering") by about ± 0·5%

from its nominal frequency in the short term, generates "random interlacing" (ie exact 2:1 interlacing does not occur) but the resulting "interlaced" lines, of variable spacing, "jitter" at a frequency of 50Hz, which the viewer's eye is not noticeable; a stable display, when viewed on a TV receiver appears to result. For "synchronizing" purposes it cannot be used.

(iii) "VIDIOGRAPHIC" Requirements

(a) The resulting display is generated on a grid of points or addressable locations and is not "continuous". Each location has a time interval of $\Delta t \approx 0{\cdot}162$us associated with it, and hence the bandwidth for adequate response has to be

$$\Delta f \geqslant \frac{2}{\Delta t} \geqslant 12 \text{ MHz.}$$

Low frequency response is of no great importance (there being no "continuous" lines or large "flat" areas).

Secondly the signal to be amplified is a two level amplitude pulse signal. The requirements for linearity in amplitude amplification and for linear phase response play little part. Thus:

(1) For the CRT a video pulse amplifier of at least 12MHz bandwidth (or having adequate response at 12MHz) is required.

These are available commercially (even up to 20MHz bandwidth) (eg (290)).

(2) For the Vidicon the effect due to a finite diameter scanning beam (the "aperture effect" (312,313)), lowers the effective frequency response of the Vidicon tube (as distinct from effects due to the video amplifier limited bandwidth). A bandwidth of 12MHz corresponds to a Vidicon imaging resolution of some 1000 TV lines (the cheapest Vidicons have a resolution (without aperture compensation) of some 400 TV lines). This required 1000 line resolution is achieved in existing Vidicons particularly if aperture compensation is used (312).

However even if Vidicon resolution is lower than this, this can be tolerated, as the output of the Vidicon is fed into, and processed by, "variable shift registers". The lowering of resolution means that, say, "n" adjacent display locations are no longer resolved as "n" separate locations but as a continuous signal of time duration $\Delta t' = n \Delta t$.

The use of J-K Flip-flops for the first stage of each shift register overcomes this and the output from the shift registers is again a discrete sequence of "n" pulses, say, as required (see section 8.5.3.2).

(b) the requirements for the linearity of the CRT display and Vidicon scanning require linearity to be better than 0·25%, as given by expression 7.13 in Chapter 7. This is for Distortion Correction being implemented and pincushion or barrel distortion minimized or eliminated.

(c) The timing or synchronizing is required to be stable and identical for both the CRT display and Vidicon camera. The H-line and V-field time-indicators require to be an integral number of $\Delta t = 0{\cdot}162$us pulses so that if an external source such as a stable crystal-controlled oscillator is used, simple division can be used to obtain the H- and V- timing signals.

From a comparison of the above requirements for "VIDIOGRAPHIC", and the performance parameters of typically low-to-medium priced commercially available equipment, it is seen that these are not suitable to be directly incorporated in "VIDIOGRAPHIC".

However since some experimental work was performed on them (eg Moire pattern measurements) and also to indicate the **main** components of CRT-TV displays and Vidicon cameras, the block and circuit diagrams for these systems are shown in Figs. A.12.1-5 and the main features briefly described.

A.12.2 THE TV-CRT SUBSYSTEM

The TV-CRT Display Subsystem, on loan from W.R.E, Dept. of Supply, Salisbury, is a PYE TV-Monitor Type 2738CF of 1955 Vintage (325).
The specifications of the Monitor are as follows:

- 625-line system, 50 field, 25 frames of 40ms duration.
- Signal inputs are
  - (i) Composite Video of 0·5-2V p. to p.
  - or (ii) $0{\cdot}5 - 2V_{p.\ to\ p.}$ video information and $0{\cdot}5\text{-}4V_{p.to\ p.}$ synchronizing information.

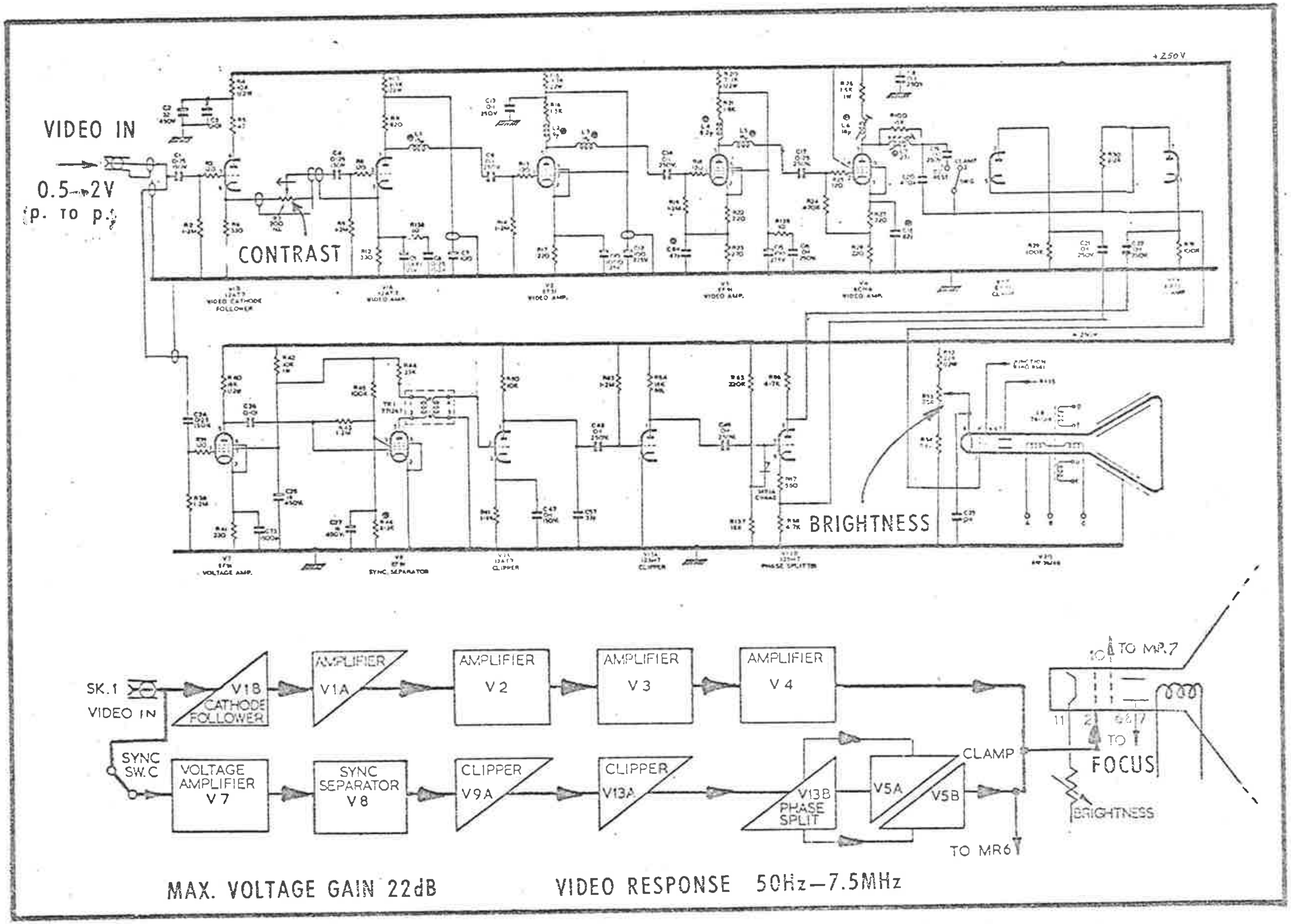


Figure A.12.1 Circuit and Block Diagram of High Quality Video Amplifier in Commercial TV Receiver — from the PYE TV Monitor Type 27[illegible]F Under Test.

- Video response of 50Hz - 7·5MHz (± 1dB)
- Display Linearity is better than 2%.
- Display size is 11·3" x 8·5" ± 20%
- The CRT Tube is a Mullard Aw -36 - 48 with W - Phosphor.

A.12.2.1 Video Amplifier Subsystem (Fig.A.12.1)

The video amplifier section is a multistage valve amplifier with better than 7·5MHz 3dB bandwidth.

DC-restoration is present whereby a DC or low frequency signal is inserted at the control grid to give a true representation of luminance of large display areas (ie low frequency signals) as the AC coupled amplifier stages may remove parts of these. This facility is not required in "VIDIOGRAPHIC".

Brightness or luminance is controlled by adjusting R_{33} which sets the cathode potential of the CRT and hence the DC voltage with respect to the control grid, ie it sets the tube operating point (see Fig.A.9.1(b)). This brightness control is required in "VIDIOGRAPHIC".

Contrast is determined by varying R_7, which adjusts the overall gain of the amplifier; the maximum gain is 22dB (or a voltage gain of 12·5). Contrast controls are not required in "VIDIOGRAPHIC" as only a two level signal is required.

A.12.2.2 Deflection Subsystem (Fig. A.12.2)

The Vertical Scanning Waveform is derived from a sawtooth waveform developed across the capacitors C_{35}-C_{71} by the blocking oscillator V_{10B}, which is synchronized by the trigger pulse applied at the anode of V_{9B}. The discharge

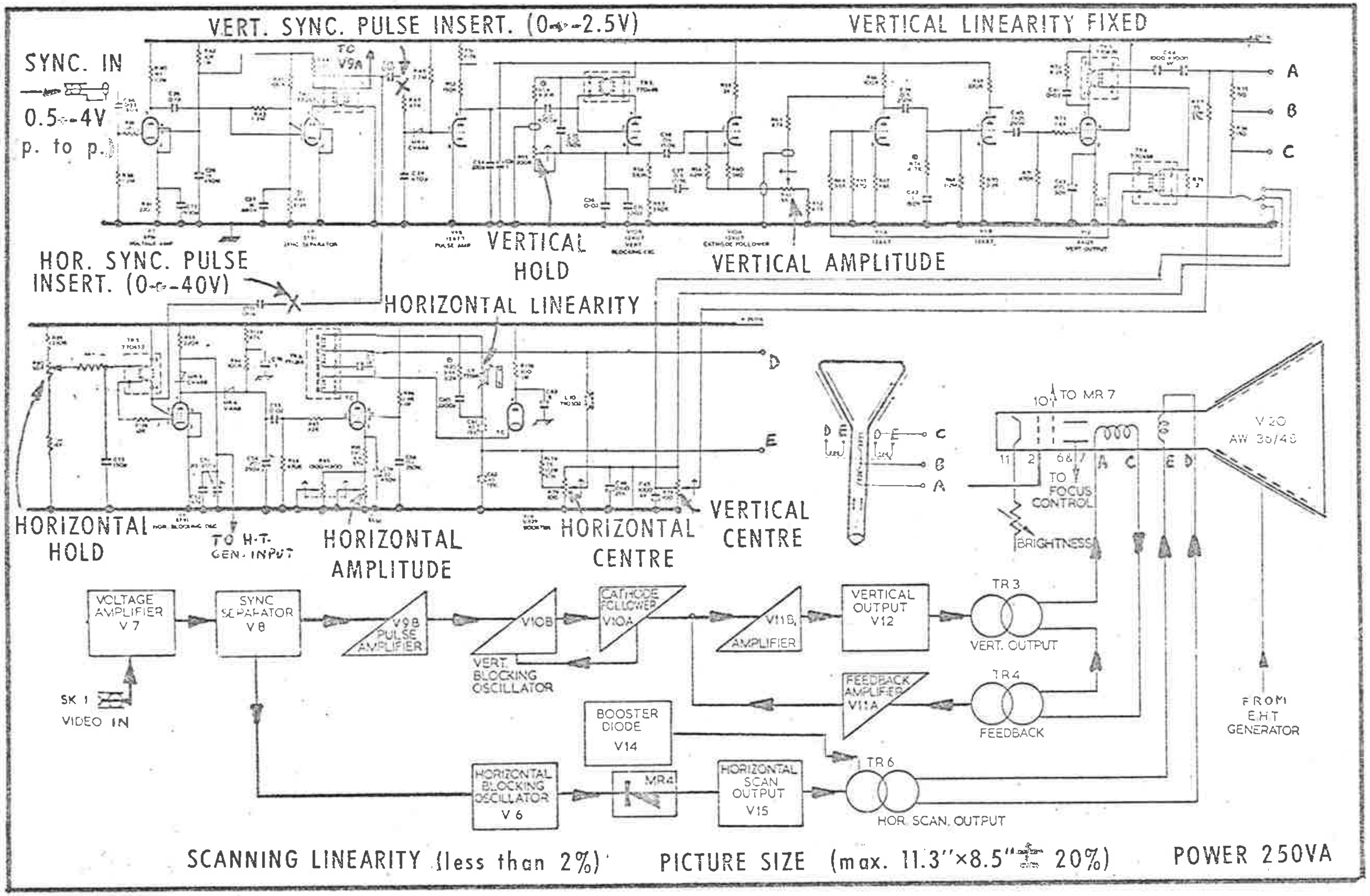


Figure A.12.2 Circuit and Block Diagrams of H- and V- Scanning Waveform Generators in PYE TV Reciever, Showing Points of Insertion of Separate H- and V- Sync. Timing Signals.

of these capacitors giving rise to the scanning waveform is via R_{56}-R_{57}, and the negative going sawtooth is applied to the cathode follower V_{10A}, where, by a bootstrap connection, the usual exponential discharge is linearized.

A fraction of the deflection coil waveform is fed back to give "good linearity" of the scanning waveform (better than 1%), giving overall display linearity of better than 2%.

<u>The Vertical Display Amplitude</u> (required in "VIDIOGRAPHIC") is by the variable resistor R_{61} which controls the input amplitude to V_{11B}.

<u>The Vertical Shift or Vertical Centre Control</u> R_{78}, controls a DC current, and its polarity, passing through the V-deflection coils. Again this is required in "VIDIOGRAPHIC".

<u>The Horizontal Scanning Waveform</u> is derived in much the same way as above.

<u>The Horizontal Amplitude</u> (R_{95}) controls the bias on the cathode of V_{15}, while the <u>Horizontal Centre Control</u> (R_{75}), as for the Vertical, controls the amplitude and polarity of a DC current through the Horizontal coils. Both are required in "VIDIOGRAPHIC".

<u>The Linearity</u> is adjusted by varying the core setting of a saturable reactor L_9 and a capacitor C_{54} in unison.

The <u>Timing Signals</u> normally are derived from a combined H- and V- Sync. signal train separated out at the Sync. Separator V_8.

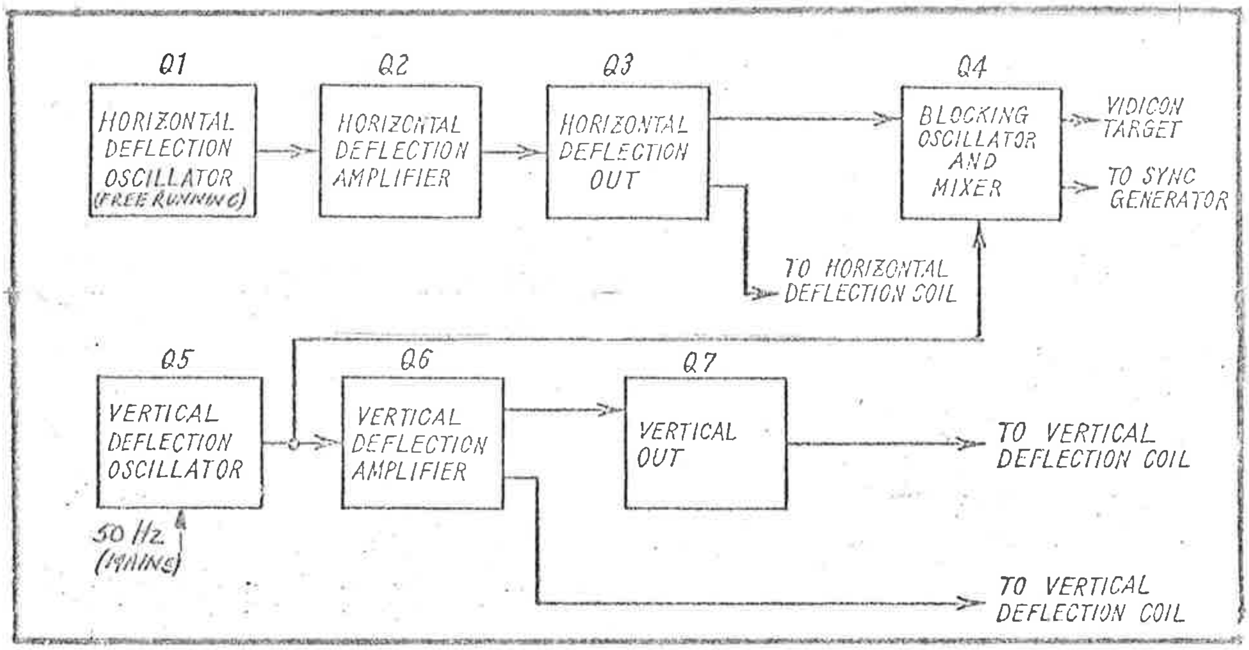


Figure A.12.3(a) Block Diagram of Scanning Waveform Generators in the Economic Vidicon Camera Sony CVC-100 under test.

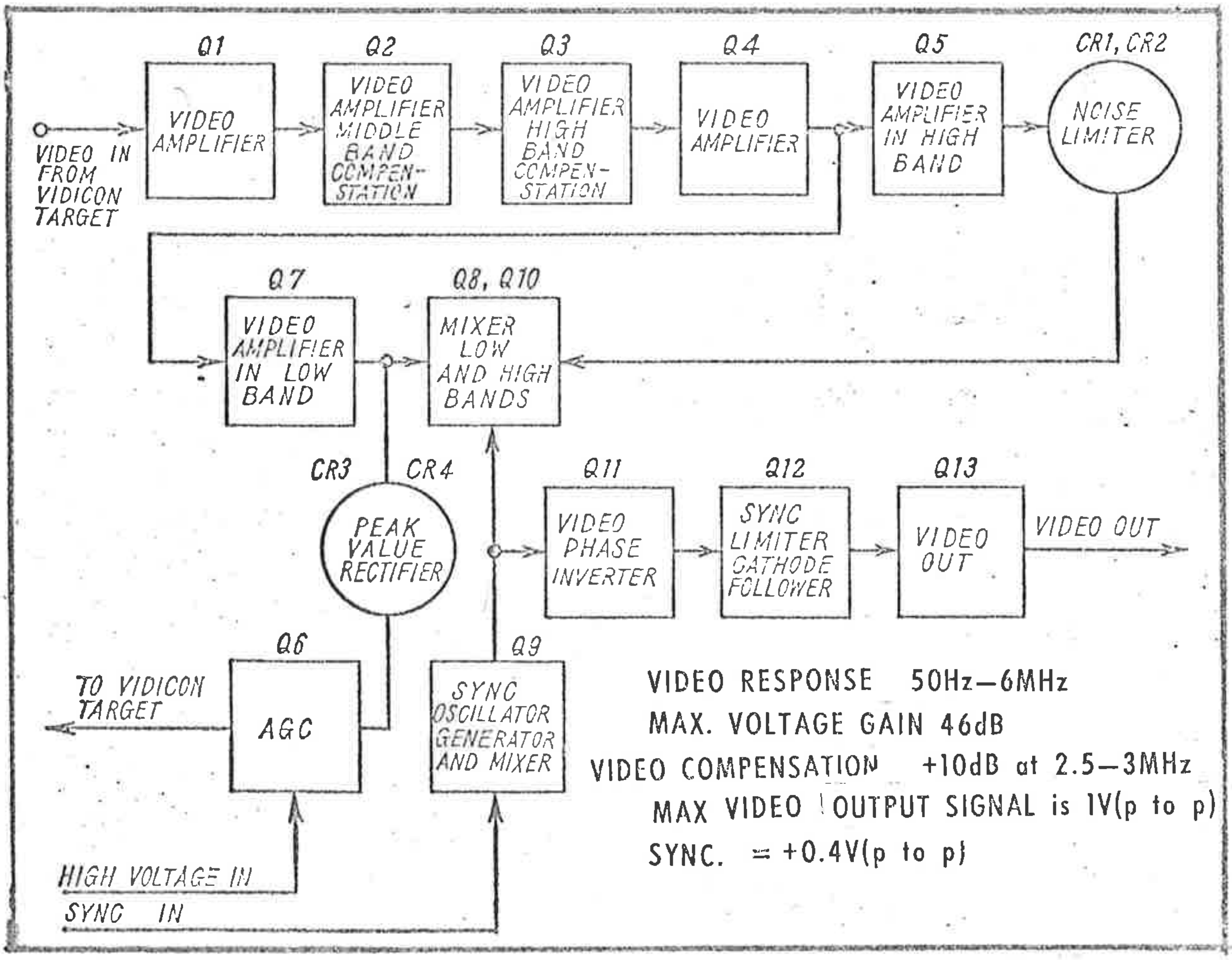


Figure A.12.3(b) Block Diagram of Video Amplifier in SONY CVC-100 Vidicon Camera.

In our case however, separate sync. signals were fed into the Vertical and Horizontal Scanning Subsystems by breaking the input lines as shown on the Fig.A.12.2 and inserting signals of the amplitude shown.

A.12.3 THE VIDICON CAMERA SYSTEM

The Vidicon Camera was purchased (cost $310), the choice being solely determined by cost; it was the cheapest on the market! It is a SONY TV Camera Model CVC-100 of 1964-1965 vintage (326).
The specification of the Camera are as follows:

- 625-line system, 50 field, 25-frames of 40ms duration, "random interlaced".
- Signal Output is a 1·4V $_{p.to\ p.}$ composite video signal with
  - (1) 1 $V_{p.to\ p.}$ video information
  - (2) 0·4 $_{p.to\ p.}$ H- and V- sync.
- Video response is 50 Hz-5MHz (better than 0·5dB) with a gain of some 46dB (a voltage gain of some 200).
- Video compensation of some +10dB peaking at 2·5 - 3MHz is provided to compensate for the aperture effect.
- the minimum luminance required is some 10ft-C.
- Linearity is of the order of 3-5%.
- Scanned area is 9·6mm x 12·8mm (0·375" x 0·5").
- the camera is provided with an F 1·9 Lens of 25mm focal length.
- the Vidicon is a SONY Vidicon M - 3016, of indeterminable class of photoconductor. Lenghty correspondence with the manufacturer was fruitless as they did not want to release spectral response (ie curves similar to

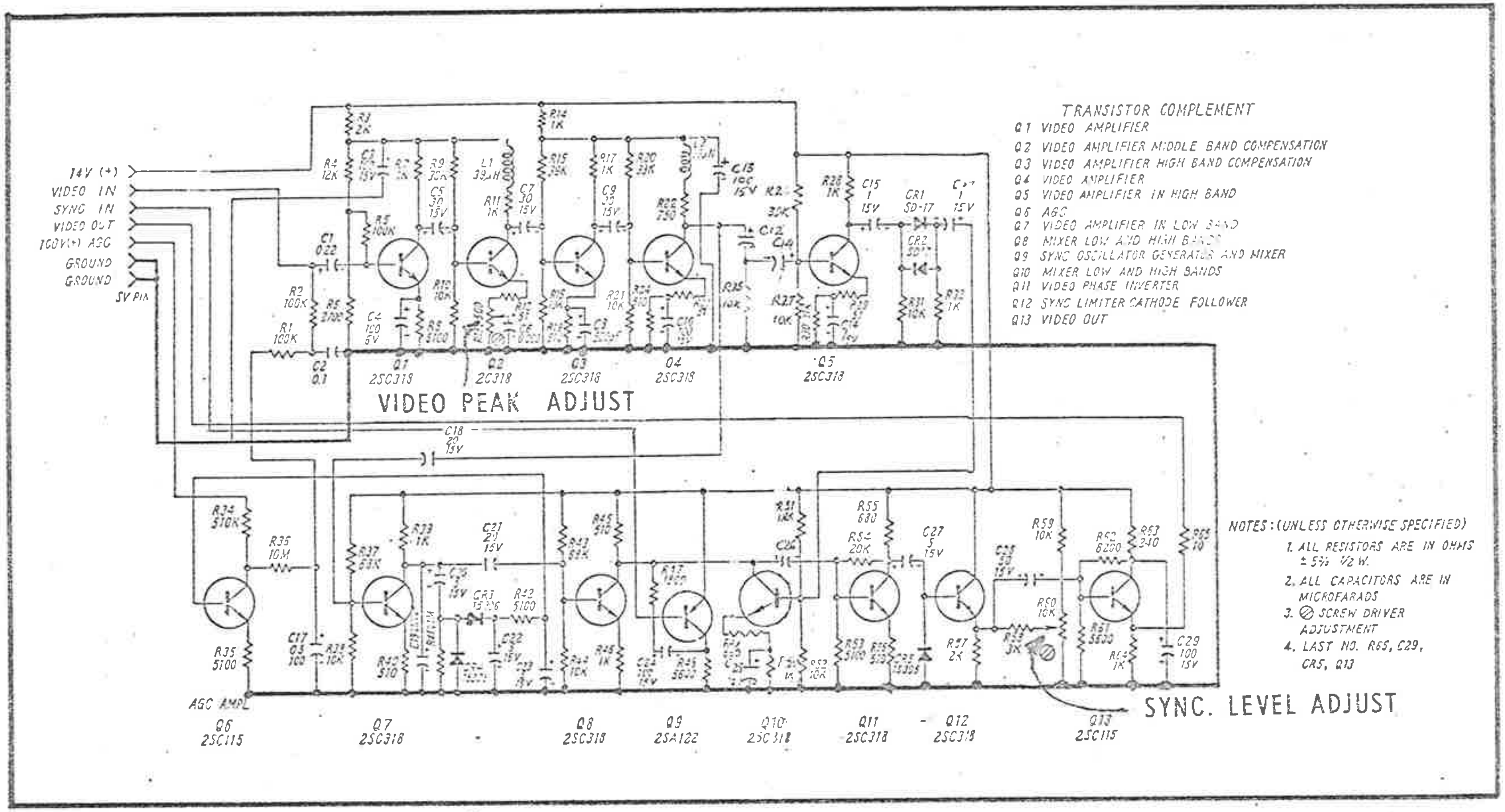


Figure A.12.4 Circuit Diagram of Video and Sync. Signal Mixer Generating the Composite Video Signal.

Facing p. A18

Figs. 66(a)(d), 66(b)(c)) nor decay curves.

A.12.3.1 The VIDEO Amplifier Subsystem (Fig.A.12.4)

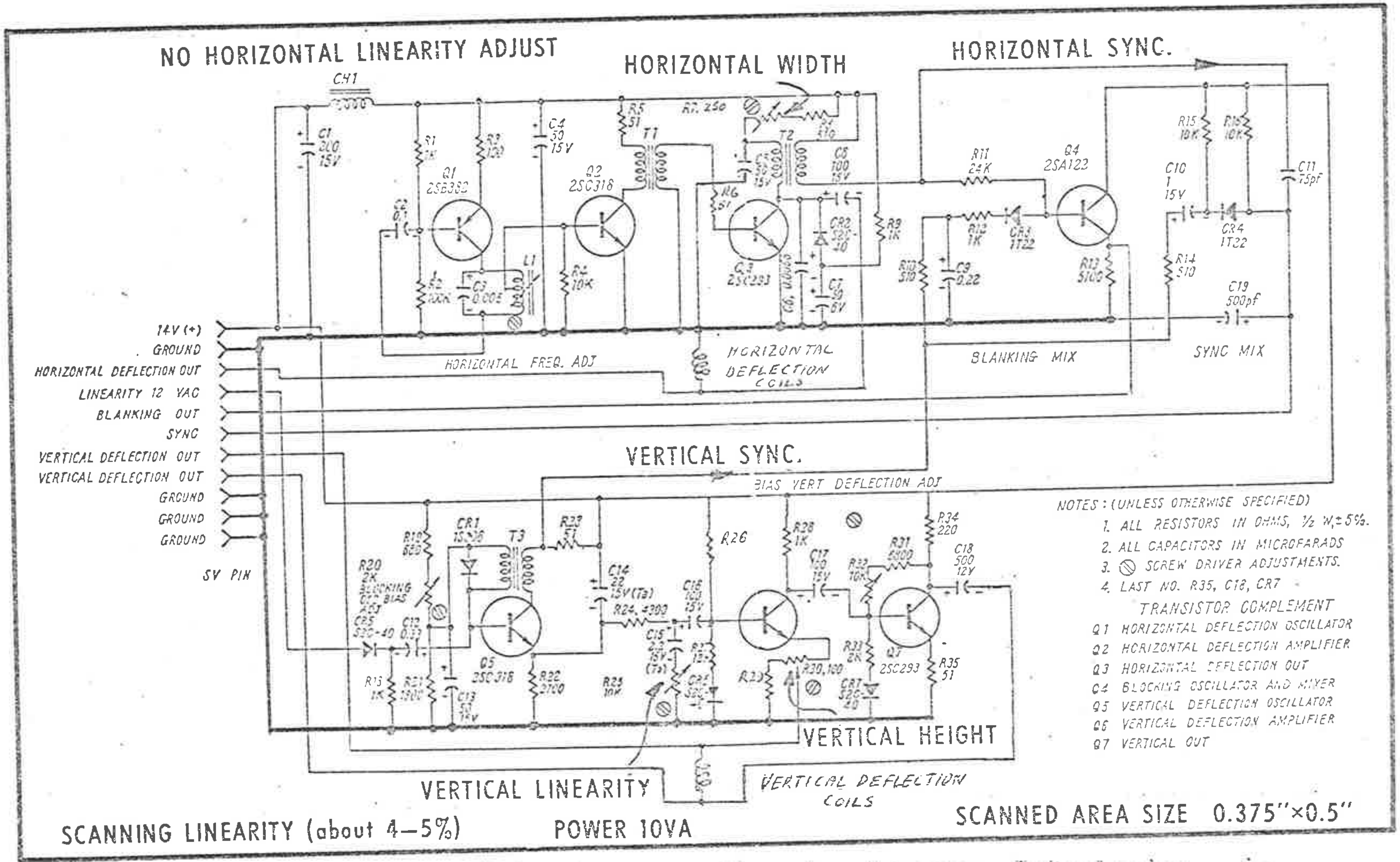
The Vidicon tube in conjunction with the Video amplifier gives an effective resolution at the centre of the Vidicon scanned area of some 450 TV lines, corresponding to a video bandwidth of some 5-6 MHz. To compensate for the aperture effect (which lowers the frequency response) the compensating networks at the emitters of Q_2, Q_3, Q5 give the extra gain of about +10dB, centered at around 2·5-3MHz.

Noise limiters are included as well as a sync. signal mixing circuit giving the output composite video signal.

The peak rectifier consisting of CR_3,CR_4, generate a rectified signal controlling the magnitude of the signal-plate voltage "V_S" via Q_4 and the resistors R_1,R_{34} and R_{36}. Thus extreme illuminations within the range of 10 - 1500 ft-C can be automatically compensated for by varying the sensitivity of the photoconductor by varying "V_S" (this response is such that this works only for changes in illumination over the whole scanned area and not merely localized highlights). This then is somewhat analogous to DC-restoration in CRT luminance control.

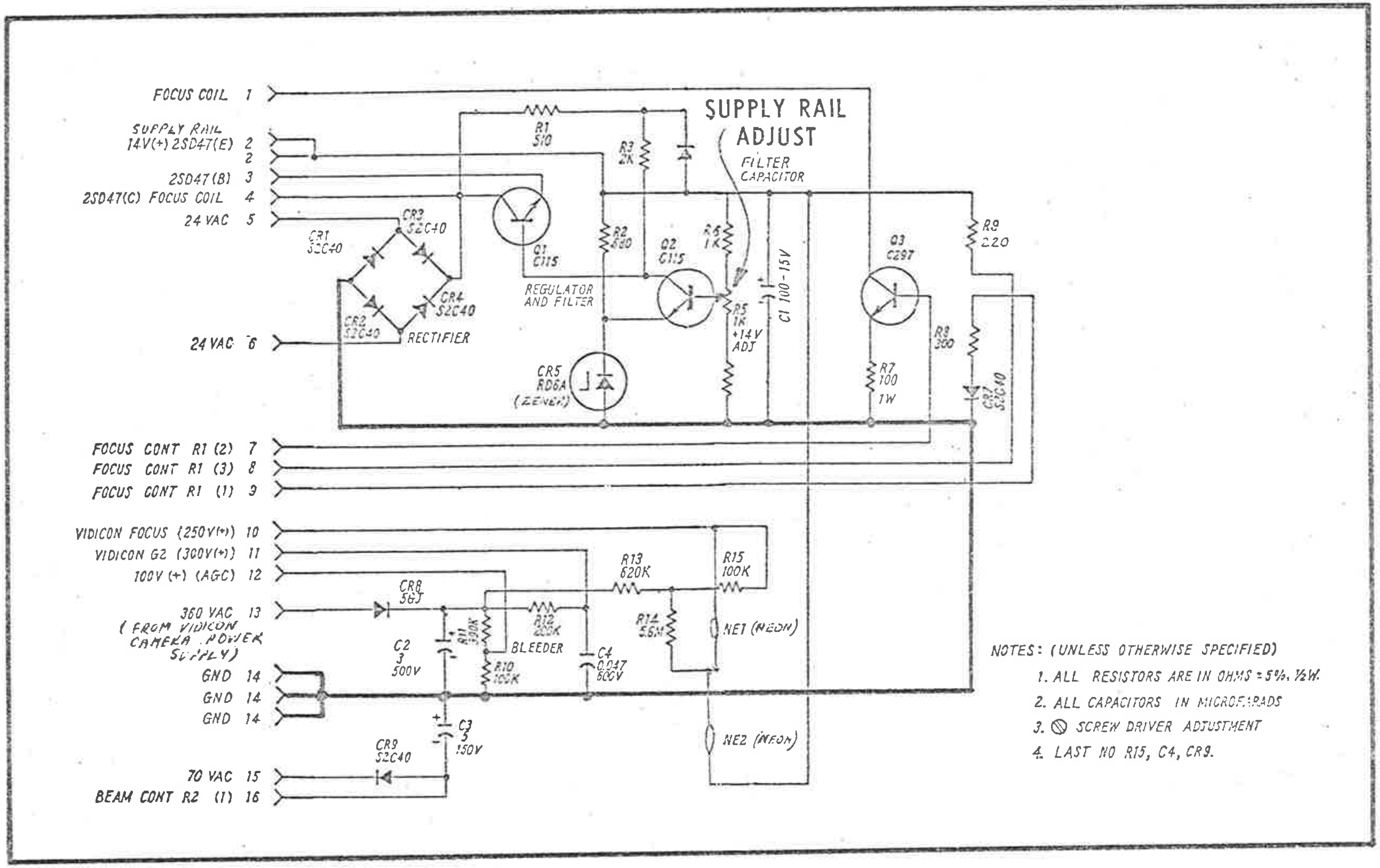
A.12.3.2 The Scanning Subsystem

The Vertical Scanning Waveform is initiated by the mains 50 Hz frequency triggering the blocking oscillator Q_5; when Q_5 is off (for most of the time interval between the V- timing pulses), the capacitor C_{14} discharges through R_{24} and R_{25}, giving rise to a linear exponential decay, the initial segment of which approximates the linear scanning waveform. Varying R_{25} varies the rate of discharge and hence the linearity.



The free running H-oscillator results in "random interlacing - ie completely unsynchronized timing.

Figure A.12.5 Circuit Diagram of H- and V- Scanning Waveform Generators in Sony CVC - 100 Vidicon Camera.



The high AC voltages are obtained from the mains voltages via Transformers not shown.

Figure A.12.6 Circuit of Vidicon Camera Rail Voltage Supply and Vidicon Grids and Anodes Voltage Supply.

The output is taken across the emitter resistor R_{29}- R_{30} of Q_6 and the collector of Q_7. Varying R_{30} varies the amplitude, and in conjunction with varying R_{32}, varies the nett DC emitter current in the V- coils and hence effectively Vertical centering is achieved.

The Horizontal Scanning Waveform is obtained by Q_3 acting as a switch. When it is "hard-on" conducting, during the active scanning period, the LC tank circuit components, due to scanning coils L and capacitor C_6, have a constant voltage impressed across them ($V_{CE\ sat}$), which results in a linear ramp current flowing in the coil.

Resistor R_7, controlling collector and hence emitter current, control "$V_{CE\ sat}$" and hence the scanning current amplitude.

The "tank" circuit being capacitively coupled has no DC current flowing and hence no centering control is required.

The V-blocking oscillator pulses as well as the H-recharge pulse in Q_3 are mixed in the network formed by R_{14}/C_{10} and C_{11}/C_{R4} to provide the sync. pulse for the composite video signal.

Timing is provided for the Vertical Scanning Waveform by the 50 Hz mains frequency, while the Horizontal Scanning Waveform is provided by the free-running oscillator Q_1.

COMMENT

Valves are high input and output impedance devices and thus are Voltage devices, while transistors are low input and output impedance devices and are thus current devices. Since at the scanning frequencies, and with the values of the scanning coil resistances and inductances (typically 1 - 100Ω and 1 - 50 mH) a relatively low impedance results, and thus these scanning coils also are current devices. Thus transistor circuits are inherently "natural" to drive deflection or scanning coils. Indeed transistor circuits can directly drive scanning coils or else by capacitive coupling, whereas for the valve circuits transformer coupling is necessary, and complex feedback is necessary to linearize the resulting current waveforms.

Transistor scanning coil driving circuits are thus very simple compared with the TV-CRT coil driving circuits. Cost, size and performance are improved, as can be seen in the relatively simple redesigned very linear H- and V-scanning circuits for the Vidicon Camera, described in Chapter 11. Transistor implemented scanning circuits for TV-CRTs are currently displacing valve circuits, again due to low cost, smaller size and improved performance (321).

. . . .

APPENDIX 13

REPRINTS OF PAPERS AND MISCELLANEOUS

1. REPRINT OF PAPER "COMPUTER DISPLAY DISTORTIONS"
2. REPRINT OF ITEM IN "COMPUTER GRAPHICS"
3. INVITATION TO BE PANEL MEMBER AT ACM CONFERENCE
4. COPIES OF ACCEPTED DESIGN NOTES
5. COPY OF ACCEPTED DESIGN NOTE
6. MISCELLANEOUS

APPENDIX 13

REPRINTS OF PAPERS AND MISCELLANEOUS

A.13.1 REPRINT OF PAPER "COMPUTER DISPLAY DISTORTIONS: CAUSES, MEASUREMENT, AND CORRECTION".

Invited paper presented at Computer Communication and Display Conference, Sydney, June 29 July 2, 1970.

Computer Communications & Display Conference

AMP THEATRE CIRCULAR QUAY, SYDNEY N.S.W. JUNE 29 - JULY 2, 1970

Conference Papers

Session 7—Papers 1,2,3
Session 8—Papers 1,2

Co-sponsored by the Association of Computing Machinery
and the U.S. Department of Commerce

A. 188

CONSULATE GENERAL OF THE
UNITED STATES OF AMERICA
Box R215, Royal Exchange
N.S.W., 2000

March 26, 1970

Mr. V.C. Sobolewski
Electrical Engineering Dept.
University of Adelaide
North Terrace
Adelaide, S.A. 5000

Dear Mr. Sobolewski:

The United States Trade Center is scheduled to open with its initial exhibition on June 29 with a Computer Communications and Display Equipment Exhibition. Approximately 25 U.S. manufacturers - including such large corporations as IBM, General Electric, Kodak, Burroughs and Control Data - will be displaying their latest developments in the computer communications and display fields.

Concurrently with the exhibition, the Trade Center is organizing a symposium which we plan to have co-sponsored by the Association for Computing Machinery in the United States. The symposium is scheduled for June 29, 30 and July 1 in the auditorium of the A.M.P.Building, two blocks from the Trade Center. The auditorium will accommodate 210 attendees. Attendance to the seminar will be by invitation only.

Papers will be given by both representatives from the companies exhibiting and other noted qualified people in the Computer Communications and Display field. Your name has been suggested to us as having an international reputation in the field of Data Graphics, and we would appreciate your consideration of presenting a paper during our symposium. Papers should not take longer than approximately 25 minutes to read. We plan to have a short question-and-answer period available after each paper is delivered. As publicity about the symposium will be printed and distributed along with the invitations and brochures describing the exhibition, we would require a short abstract of any paper planned to be given by April 21. At the same time, we would need to know the approximate length of time necessary to deliver the complete paper and also whether any special equipment such as a projector, blackboard, etc. would be needed.

/2......

We would be very willing to support your participation in the seminar by paying your round trip transportation and an honorarium of $25 to cover your expenses.

If you would like to have any further information on this seminar please do not hesitate to contact me, telephone 241-1031, extension 58.

Sincerely yours,

T.F. Welch, Jr.
U.S. Trade Center

THURSDAY JULY 2 1970

Session 7, 9 a.m. – 12 noon.

Chairman
CARL MACHOVER.

1. Application of High Speed Passive Graphics:
   WILLIAM D. SMYTHE,
   California Computer Products, Inc.
2. Electron Beam Recording-Microfilms, New Method of Talking with Computers:
   H.W. CLARKE,
   3 M Company.
3. Computer Driven Display Distortion Causes, Measurement and Remedies:
   V.C. SOBOLEWSKI,
   University of Adelaide.

COMPUTER DISPLAY DISTORTIONS

Causes, Measurement and Correction.

V.C. SOBOLEWSKI
Elec. Eng. Dept. University of Adelaide
and W.R.E. Salisbury, South Australia.

INTRODUCTION

In computer-driven displays, images are made up of discrete elements (e.g. the beam spot in Cathode Ray Tube displays (C.R.T.'s)), and assemblies of these discrete elements such as lines, boundaries etc. The finite size of elements leads to the concept of resolution, the measure of fineness of detail capable of being displayed, while the assemblies of points, specifying lines, shapes etc. leads to the concept of geometrical fidelity or linearity, and conversely the lack of it, which is termed geometrical distortion. Distortion of displayed images may be defined as deviations in any form from corresponding ideal images as specified by the analytic expressions describing them (if applicable), or by mismatches when a primary or ideal image is geometrically projected onto its corresponding displayed image. Such "general" distortion covers both geometrical distortion and element-shape distortion (the latter defect known in CRT's as "defocussing", leading to loss of resolution).

An ideal display in this respect is one which has constant-size elements (i.e. always correctly "focussed") and has no geometrical distortion over its display area.

Only CRT displays will be discussed here as they are by far the most common of computer driven computer-displays; electro-mechanical x-y plotters and the like are excluded. Moreover half-tone images and associated problems of display fidelity of half-tones will not be covered; only line graphics will be covered.

Focussed displays of high resolution and of high geometrical fidelity are not only desirable from a user-comfort point of view, but are necessary in precision engineering graphics (drafting, photo-mask layout etc.), photo-typesetting, and precise pattern-recognition such as the analysis of bubble-chamber pictures. In such applications, CRT displays may be required to have spot sizes (element size) of 0.005" diameter or less, and geometrical distortion to be less than 0.02% of display area height (the usual way of defining distortion), in a display area of about 15" x 15". On the other hand, a "good quality" commercial CRT display may have spot sizes of up to 0.025" diameter and geometric distortions of up to 3%.

To eliminate distortions from a given display, the distortions must first be measured to at least the same degree to which any remaining distortion may be tolerated. Knowledge of the causes or sources of distortion suggests means of correcting the distortions.

A brief listing of such distortions or "aberrations" is given together with a very simple but very precise method of measuring display distortion; finally mention is made of practical ways of eliminating them.

The magnetically-deflected CRT (M-CRT) will be dealt with, as the electrostatically-deflected CRT (E-CRT) not only has size and cost disadvantages, but also suffers from much poorer resolution due to considerable defocussing occurring with large deflections. From expressions derived in [1], giving the relative increase in spot diameter $\frac{\delta D}{D}$, or defocussing as a function of deflection angle γ, it can be easily shown that

$$\frac{\delta D_M / D}{\delta D_E / D} \approx \frac{1}{(\sec\gamma + 1).L/\ell} \qquad \text{..........(1)}$$

where $\frac{\delta D_M}{D}$ is the defocussing in M-CRT's

$\frac{\delta D_E}{D}$ is the defocussing in E-CRT's

ℓ = length of deflecting plates in E-CRT's
L = length of beam from center of plates to screen in E-CRT's.

As sec γ is positive and greater than 1, and $\frac{L}{\ell}$ is typically 10 or more, the defocussing and consequently loss of resolution is at least 10-20 more serious in comparable E-CRT's than M-CRT's. This fact allows larger beams (2-3 x) in M-CRT's with resulting increase in display brightness.

Speed advantages of E-CRT's are becoming less significant due to lower inductance deflecting coils being developed. Occasionally in the same CRT, E-deflection may be used to generate alphanumerics or small detail after the beam has been grossly positioned by M-deflection.

TYPES OF DISTORTION AND THEIR CAUSES

Distortions or aberrations may be classified as

(a) Mechanical or first-order aberrations : these are caused by physical misalignment of the axis of deflection coils, focus coils electron-gun assemblies and screen geometry defects, with the axis of symmetry of the whole CRT system (which in the ideal case is the undeflected electron beam). This causes lack of symmetry around the ideal centre of the display, leading to defocussing and distortion not only due to purely geometric effects (particularly noticeable in curved display screens) but also those due to secondary effects similar to deflection aberrations (see below), since an asymmetrically located beam with respect to the central axis may be considered as a slightly deflected beam.

Included here may be effects due to stray electric or magnetic fields (particularly the earth's magnetic field), which cause asymmetrical and/or transient deflections. (The whole display may be shifted unidirectionally by up to 1% of its own linear dimension, by the earth's field unless proper tube shielding is present).

Precision jigs enabling deflecting and focus coils to be positioned with great precision overcome the above aberrations.

(b) Chromatic aberration and aberration due to space-change both of which cause defocussing. Chromatic aberration is due to emitted electrons from the thermo-cathode having different initial velocities, which by the laws of electron optics are brought to different foci, hence the defocussing. Making final accelerating voltages so high that initial electron velocities are negligible with final electron velocities overcomes this defect. Similarly the electrons in an electron beam exert mutual repulsion resulting in a wider beam at the screen and thus defocussing.

(c) Deflection aberrations and 3rd-order aberrations. Deflection aberrations are those due to geometrical reasons, namely the variation of electron beam length as the display area is accessed by the beam, from maximum deflection (at the corners of the display) to zero deflection (at the center of the display).

Third-order aberrations are due to violations of the conditions of paraxiality. Paraxial rays are those which are at, or very near, the central axis of symmetry of the electro-optical system (i.e. the ideal undeflected beam). The formation of images on the display are due to electron beams whose trajectories are described by equations which contain the series expansion (about the central axis) of the deflecting and accelerating electrical and magnetic fields. Using only the first terms of the expansions results in first-order "Gaussian", linear or undistorted electro-optics; using the first two terms (the expansion is an odd-termed series) results in the so-called 3rd-orders aberrations being present giving expressions for the deviations from the ideal or linear case. Higher order aberrations are detected if more terms of the expansion are used.

There are eight 3rd-order aberration terms, each causing a distinct effect. Their expressions are fairly complex and are rarely used for numerical evaluation, not only for that reason, but mainly because they contain the series expansion of E-and M-fields consisting of the values of these fields at the axis and their higher order derivatives or gradients; these are very difficult to measure in practice and even more so to obtain accurate values.

Deflection distortion and 3rd-order aberrations can be conveniently lumped together by the effects they cause:

(1) defocussing aberrations - of which astigmatism, spherical aberration, coma etc. are examples and shown in fig. 1(a). The spot is not only enlarged with diffused boundaries but may be severely distorted. Certain conditions of M-and E-focussing and accelerating fields may minimise or eliminate these defects.

Deflection defocussing is due to the non-constancy of beam length under deflection. For a flat-faced screen, beam length, l, can be represented quite accurately by

$$l \approx L + \left(\frac{x^2 + y^2}{2L}\right) \qquad \text{..........(2)}$$

Where L is the undeflected beam length and x, y are the screen co-ordinates of the deflected beam.

A fixed magnetic field focus coil focusses a beam to some fixed focal length, say L; focussing to some variable length l, as above, requires "dynamic focussing".

(The geometry of magnetic deflection is shown in fig. 2).

(2) geometrical distortion aberrations which appear mainly as "pincushion distortion" and spiral distortion (fig. 1). Spiral distortion is negligible in M-CRT's, causing spot rotation and becomes only significant where long axial magnetic fields are present and low accelerating voltages such as in various imaging tubes (orthicons, vidicons). By far the most serious distortion, and more so in flat-screen CRT's than round screen CRT's, is pincushion distortion; for ideal linear sawtooth driving currents in the deflection coils, this distortion unless corrected can be up to 15% of full screen height, in 60^{0} tubes.

It has often been "shown" that the y-coordinate, say, in such a distorted display, can be given in terms of ideal (linear ramp) deflecting currents Ix and Iy, as

$$y = k_1 I_y \left[1 + L k_2 (I_y^2 + I_x^2) + \ldots \right] \qquad \text{..........(3)}$$

where k_1, k_2 are predictable constants
L is the distance of the centre of the deflection coils to the centre of the display screen
with an analogous expression for the x-deflection.

Such derivations are based on many assumptions, the main ones being:

(i) at any one instant the deflecting fields are constant within a sharply defined region and are zero outside this region. (the region of width ℓ in fig. 2).

(ii) the deflected beam is considered to originate from a "centre of deflection", C, which is constant in location and positioned midway along the deflecting coils (see fig. 2).

In fact this "centre of deflection" shifts with increasing deflections while magnetic deflecting fields are never sharply defined and indeed the actual intensity distribution along the length of the coils is often not exactly known.

To overcome any effect of these assumptions, a derivation from first principles of electro-optics is required which results in simultaneous differential equations, solutions of which give the required beam trajectories [2]. Arbitrarily shaped M-field distributions along the coil-axis can be substituted within the resulting equations (or even graphically integrated from the original differential equations); the concept of "centre of deflection" does not even enter the equations.

As a comparison, using equations from [2] with the assumption of sharply defined deflecting fields, the analogous expression to (3) for the y-deflection is

$$y = k_1 I_y . [1 + (L - \ell/4).k_2.(I_y^2 + I_x^2) + \dots\dots\dots] \qquad \dots .(3a)$$

where k_1, k_2, Ix, Iy are as before, but the term $(L - \frac{\ell}{4})$, decreasing the distortion, is due to the shifting "centre of deflection". "ℓ" is the length of the deflection coils.

Further terms, of the form $(I_x^2 + I_y^2)^{2n}$, n = 1, 2, ... are also present; their coefficiants also differ from those in (3).

For a non-flat curved screen, of a radius of curvature R, the equivalent expression is from [2],

$$y = k_1 I_y . [1 + L (1 - L/(R - \Delta S/2) - \ell/4L)(I_y^2 + I_x^2) + ..] \quad ..(4)$$

where k_1, k_2 L, ℓ are as before and ΔS, is the difference between the curved screen and the flat screen geometry (see fig. 2).

It can be seen that if R = L, i.e. the centre of deflection corresponds to the centre of curvature of the screen, pincushion distortion is still present albeit negligible. Usually, text-books state no distortion is present.

MEASUREMENT OF DISTORTIONS

It has been mentioned before that most expressions for the various distortions are rarely used to determine the magnitude of distortions. Distortions must be measured accurately under the condition of linear sweep currents feeding into the deflection coils; from such measurements the appropriately shaped sweep currents can be derived to produce distortion-free displays.

Current methods to measure display distortions are cumbersome; the common basis for most is aligning in front of the display some transparency on which are drawn some accurately spaced grid of lines or pattern, with a replica of the same pattern generated and displayed. Ideally the two patterns should be coincident; if otherwise, any localized mismatches indicate distortion at that location, the greater the mismatch the greater the distortion. A "ball" chart is the standard pattern in TV linearity measurement; more usually square grid patterns are used. Neither precision nor accuracy is very good with these methods, particularly as parallax errors are present. The hardware and/or software to generate the equivalent patterns on the display may themselves introduce errors.

Measurement of defocussing implies spot-size measurement; this is difficult because where exactly the center of the spot is and what the spot boundaries are or how to define the spot provides problems. Short of observing individual spots over the display area with a microscope, no other methods seem available.

As defocussing causes loss of optical resolution, defocussing should be gauged by resolution, or loss of it, over the display area. Perhaps a test pattern such as in fig 3 can be generated over various sections of the display and tested for resolution, varying the separation between the spots in the "cross" until individual spots are resolved. This separation gives a measure of the resolution and hence the defocussing. It must however be noted that resolution depends on such factors as display brighteners, contrast and, in storage CRT's, the time of measurement after display generation.

Such measurements are further complicated by varying factors such as temperature, power supply ripples and pickup, which cause long and short-term drifts in spot locations. The short-term drift or "jitter" may be at display refresh rate or lower, causing the resultant "spot" to be defocussed, "it" being due to slightly differently located spots after each refresh cycle which are integrated by the eye or the device making the hard-copy (e.g. camera) (see fig. 3). Such a "jittered" point may appear up to twice the size of the actual spot size.

Occasionally optical resolution is confused with "position resolution" or registration resolution, which is merely a measure of the precision with which a spot can be positioned with respect to the adjacent spot. It is a measure of the fineness of the dot matrix or grid on which the resultant display is positioned, and has nothing to do with actual spot size. Thus a 10" x 10" display area may have 8192 x 8192 accessed locations, giving a position resolution of 0.0012", even though the spot diameter may be 0.003". Overlap of spots occurs and is necessary if sharply defined lines are required (fig. 3.)

MEASURING DISTORTION PRECISELY

Recently methods using Moire fringes have been suggested to measure TV display distortion but only as a qualitative tool ("is the TV display very distorted or not?"). More recently a detailed treatment of Moire patterns as applied to raster-scans displays and devices has been given [3], with examples shown. Very economical in realization, these methods enable geometric distortions to better than 0.1% of display height to be obtained.

Moire patterns are formed when two sets of lines, one set being on a transparent plane, are superimposed. If both sets of lines are identical and properly aligned, they overlap and it appears as though only one set is present. If the 2 sets of lines are misaligned (or else the 2 sets are not identical e.g., one set of lines may be distorted), some lines overlap or intersect. The loci of intersections appear more pronounced than other lines, and are termed "moire fringes". Knowing the geometry of the sets of lines, the ideal Moire fringes may be calculated. If one set of lines is distorted, then the resultant "distorted" Moire fringes when compared with the "ideal" Moire fringes enable the original distortion to be caluclated to great accuracy. Simple Moire fringes are shown in fig. 4.

The spacing of resultant Moire patterns is usually inversely proportional to the difference in the spacings of the lines forming the two sets of interacting lines. Keeping one set of lines fixed (the "reference" transparency), the other set of lines formed on the display can be made to have variable spacing with resultant variation in Moire fringe spacing. The precision of determining distortion can thus be varied; a "magnifying glass" of variable magnification, to examine distortion, results.

For measuring display distortion one set of lines consists of a set of slightly divergent lines (or in another method, 2 sets of orthogonal lines to form a regular dot matrix) on a transparency, which is then appropriately aligned and affixed in front of the display. The other set of lines (or dots) ideally consists of parallel equispaced lines (or else a regularly spaced dot pattern) generated on the display. (In raster-scan displays, one set of lines, the horizontal scanning lines are inherently present!). The display distortions make these lines (or dot matrix) distorted, with the resultant Moire fringes distorted. Simple expressions convert these Moire fringe positions to distortion coefficiants over the display area. Two separate methods are available with actual results of a fairly linear TV monitor (better than 1.5% linear over 80% of the display area) shown in fig. 5.

So long as a display is capable of generating and displaying equispaced parallel lines or a dot matrix, the Moire fringe methods are applicable.

From such distortion measurements maps of constant distortion contours can be drawn for the display area, an example of which is shown in Fig. 6; a separate map for vertical and horizontal distortion is necessary. Examinations of these maps provide the correction data necessary to give distortion free displays. More so, they indicate that in general distortion profiles along say one horizontal scanning line (in raster-scans) cannot be expressed analytically.

CORRECTING DISTORTION

Corrections to give distortion-free and focussed displays may be implemented by 3 separate methods, each of which progressively eliminates smaller and smaller corrections. These are:

(1) clever and artful design of electron gun assemblies, focussing and deflecting assemblies, tube shielding, ripple-free power-supplies etc., of the CRT concerned.

(2) correction by feeding in certain focussing and deflecting coil currents or voltages, the form of which is derived from analysis.

(3) the final correction of residual non-analytically describable distortions, by applying correction currents derived from stored correction-coefficiants or data, which is derived from accurate measurement of display distortion.

(1) Tube design

"Clever and artful design" because tube design and associated focussing- and deflecting-assemblies are still more of an art than science. Shapes of electrodes, windings and location of coil-assemblies, designed for certain optimum E-M field configurations, are still in the main designed by graphical field-plotting. Analytic expressions for aberrations etc. may indicate where E-M field distributions need be varied but cannot predict or indicate where or how stray fields between say the deflection and focus coils occur or how to compensate for them. Minimizing or elimination of these aberration can only be achieved through experienced design.

Carefully designed tubes, of low deflection half-angles ($\pm 20^{0}$ or less), with precision aligned yokes and focus coils, optically flat screens, and with linear current-sweep feeds, have distortion of 1% at best.

(2) Correction currents

Previously, expressions were given for geometrical distortion and beam length variation causing defocussing - both effects (or defects !) due mainly to geometrical reasons. From a simple study of these expressions, appropriate currents to feed the focus and deflection coils can be obtained to eliminate this form of distortion.

For geometrical distortion, the necessary deflection coil current for the y-deflection is

$$I_y' = I_y\{1 - k_1/k_2 \,.\, (L-\ell/4).(I_y^2 + I_x^2)\} \quad \ldots\ldots\ldots (5)$$

where I_y, I_x are linear sweep currents producing y- and x-deflection respectively.

A similar expression for I_x^{1}, giving the x-deflection applies.

For correct focussing, the focus coil current is

$$I_f \approx I_{DC} + k_3\,(I_y^2 + I_x^2) \quad \ldots\ldots\ldots\ldots (6)$$

where k_3 is a predictable constant depending on k_2 and on the focus coil.

I_{DC} is the current giving static focus.

The correction currents may be generated from blocks of analogue squarers and multipliers and adders. The accuracy of even the most economic of these circuits is adequate for the corrections necessary. If 5% distortion need be corrected, then if the analogue circuits are 0.5% accurate, the correction precision is about 0.03%. Other distortions still present usually are much larger than this. The factor $(I_x^2 + I_y^2)$ common to both corrections suggests circuit sharing.

Alternatively the necessary waveforms may be generated by functions generators consisting of networks of voltage sources, resistors and diodes; the required waveforms are generated by switching in the appropriate branches by the corresponding x- and y-coordinates. The correction waveforms being piece-wise-linear approximations, the precision of correction depends on the complexity of the networks and the storage capacity available for the breakpoints in the piece-wise-linear correction waveform.

(3) Non-Predictable Correction

After the above form of corrections has been carried out, distortions of the order of 0.1% still remain. These are due to the remaining aberrations, electron-gun and coil-assemblies defects etc. and can only be deduced from direct measurements, e.g., by the Moire patterns method. These distortion-coefficiants or rather the derived distortion correction-coefficiants, being non-analytically describable, need be stored in the graphics console memory; in one case (4) , the display is divided into a 64 x 64 matrix, with an x-and y-correction-coefficiant associated with each matrix point. During display generation these correction-coefficiants are read off and applied via D/A converters to the appropriate deflection coils. Alternatively using a function generator [5] as described above, not only pincushion distortion may be corrected, but the piece-wise-linear approximation can be shaped so as to correct the remainent distortion as well.

SUMMARY

Computer-driven displays, particularly for precision graphics applications, need have high resolution and high geometric linearity. The most common form of displays, CRT's, suffer from large distortions, some predictable, others not. Precise measurements of distortions need be made before corrections can be implemented; Moire pattern methods provide such a measuring tool. Ultra-linear, focussed displays can be achieved through careful tube design, the use of analytically-derived corrected focussing and deflection-coil currents and finally, by applying localized corrections generated from stored tables of correction-coefficiants or stored-piece-wise linear correction curves.

. . . .

REFERENCES.

[1] H. Moss - "Narrow Angle Electron Guns and Cathode Ray Tubes" in "Advances in Electronics and Electron Physics Supp. 3", Academic Press, New York 1968.

[2] V. C. Sobolewski - "Pincushion Distortion in Magnetically Deflected CRT's with Flat and Non-flat Screens", to be published.

[3] V. C. Sobolewski - "Using Moire Patterns to Determine the Distortion of Graphic Displays and Graphic Input Devices", Proc. IEEE, Vol. 58, 4, pp. 567-583, April, 1970.

[4] W. Huelskoetter and J. Kimlinger - "Ultra Precision Artwork Generated with a CRT Display", IEEE Internatl. ConV. Record pp. 382-383, 1969.

[5] R. D. Klensch and E. D. Simshauser - "The CRT in Phototypesetting Systems", IEEE Spectrum, Vol. 6, pp. 75-80, Sept. 1969.

. . . .

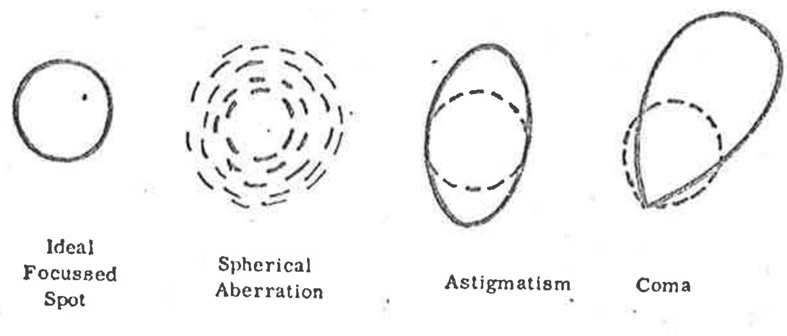


Figure 1(a). Some Spot Defocussing Aberrations

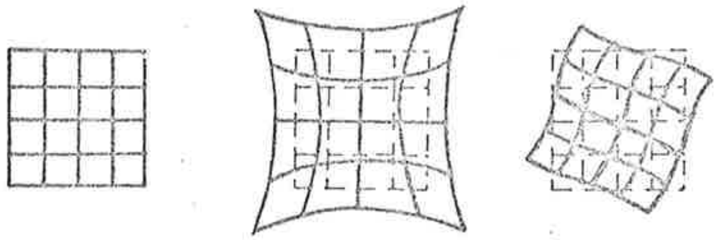
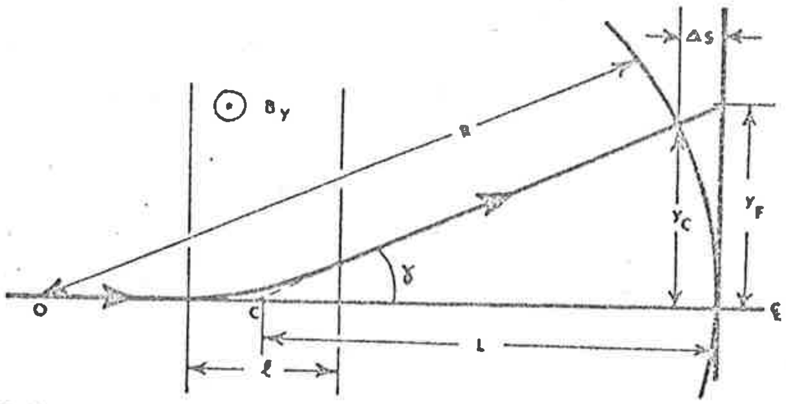


Image with no Geometrical Distortion

Pincushion Distortion

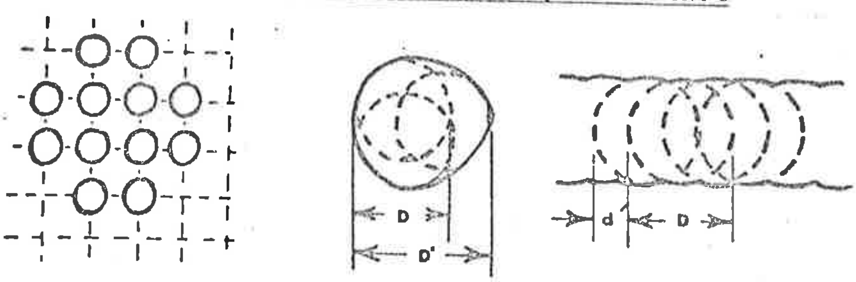
Magnetic Spiral Distortion

Figure 1(b). Geometrical Distortion of Display



Deflection occurs in the region of width 'l' where the magnetic field intensity B_y, at any instant is constant, and zero outside this region. The deflected electron beam (arrowed), when extended back to the axis, is assumed to originate at 'C', the 'centre of deflection': in fact 'C' shifts towards the screen with increasing γ. 'R' is the radius of curvature of the curved screen: 'S', evaluated for a particular γ, is used to calculate 'pincushion distortion' in curved screen CRT's.

Figure 2. Notation of Magnetically Deflected CRT's



Resolution Test Cross

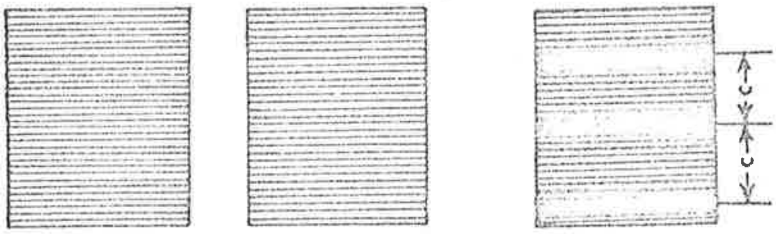
Effect of Jitter

D is true Spot Diameter; D' is apparent Spot Diameter

Position Resolution

d is Position Resolution

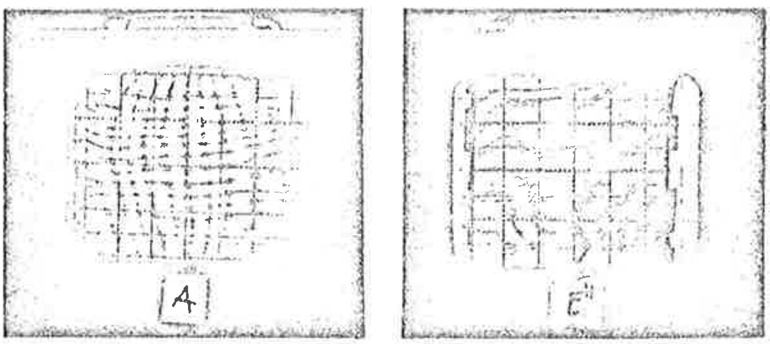
Figure 3. Some Resolution Concepts



Line spacing 'a' Line spacing 'b' M-fringe spacing 'C'

When two parallel line patterns of slightly different spacing are superimposed and aligned, Moire fringes occur of spacing C = (ab)/(b-a). Imagine one set of lines are generated on the display, and the other set on a transparency superimposed over the display; then the display distortion can be evaluated from simple relationships linking the ideal Moire fringe spacings with the actual Moire fringe spacings which are distorted (see reference [3]).

Figure 4. Formation of Simple Moire Fringes



Two forms of Moire fringes are shown. In 'A' Moire fringes corresponding to the simple case shown in Figure 4 but with horizontal and vertical fringes simultaneously shown. In the complex Moire fringes shown in 'E', the broken line traces out directly the graph of vertical distortion on a certain linear scale. Both examples of course show distorted Moire fringes due to display distortion.

Figure 5. Moire Fringes from an actual Raster-scan Display (from reference [3]).

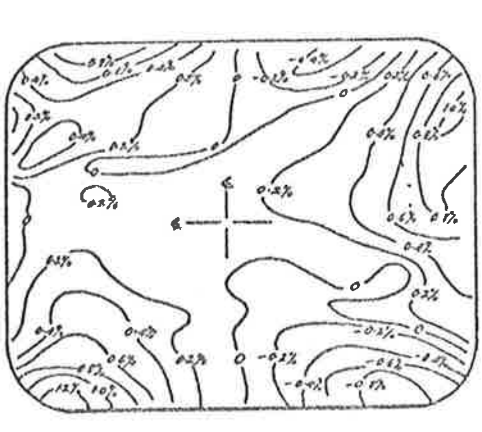


Figure 6. Contour Map showing Horizontal Distortion (o. 2% Distortion Contours)

A.13.2 REPRINT OF ITEM IN "COMPUTER GRAPHICS", WINTER, 1969-70.

A QUARTERLY REPORT OF SIGGRAPH-ACM

COMPUTER GRAPHICS

VOL. 3 NO. 4 WINTER 1969-70

ASSOCIATION FOR COMPUTING MACHINERY
ACM
FOUNDED 1947

INTERNATIONAL ISSUE

INTERNATIONAL REPORTS

CONTENTS

Page

I AUSTRALIA

A. The Commonwealth Scientific and Industrial Research Organization (CSIRO) Division of Computer Research P. O. Box 109, Canberra City, A.C.T. Australia 2601 — 15

Dr. G. N. Lance, Director, submits a brief description of the work which is currently being undertaken at their Research Division.

B. University of Adelaide Adelaide, South Australia 5001 — 21

Mr. V. C. Sobolewski, Electrical Engineering Dept., furnishes two items. One deals with work being done on a graphics store and direct I/O capabilities; while the other deals with the precise determination in display area distortion of TV monitors and TV cameras using Moire patterns. He states that these methods are applicable to any other graphic display or equipment capable of generating a parallel line structure.

II CANADA

A. University of Toronto L. Mezei, Associate Professor — 22

B. International Business Machine, Canada Ltd. 36 King St. E Toronto 210, Ontario Mr. William J. Bacon, Manager, Scientific Applications — 23

III ENGLAND

A. Atlas Computer Laboratory Science Research Council Berkshire, England Mr. F. R. A. Hopgood — 23

B. Imperial College of Science and Technology, Center for Computing and Automation Prince Consort Rd, London SW7 Dr. W. S. Elliott, Professor of Computing — 23

Sobolewski, V. C. (1969-70). Graphics input/output interface and store & Precise determination of faster distortion of graphic devices. *Computer Graphics, 3*(4), 21.

NOTE:

This publication is included in the print copy
of the thesis held in the University of Adelaide Library.

A.13.3 INVITATION TO PARTICIPATE AS A PANEL MEMBER TO DISCUSS "COMPUTER GRAPHICS" AT THE A.C.M. CONFERENCE, NEW YORK.

Unfortunately due to lack of travel funds, the invitation could not be taken advantage of.

1133 AVENUE OF THE AMERICAS
NEW YORK, N.Y. 10036
(212) 265-6300

ACM Association for Computing Machinery

January 30, 1970

Reply to: Code 745
Dept. of the Navy
Naval Ship Research
and Development Center
Washington, D. C. 20007

Mr. V. C. Sobolewski
Electrical Engineering Dept.
The University of Adelaide
Adelaide, South Australia, 5001

Dear Mr. Sobolewski:

The Special Interest Group for Computer Graphics is sponsoring a panel on computer graphics, at the 1970 ACM Conference to be held in New York City, September 1-3. The topic to be discussed by the panel will be: Computer Graphics in the 1970s.

I should like to extend an invitation to you to participate in this panel. We would appreciate your views on the state of the art of computer graphics in your country, identification of problem areas, and what are your goals for the 1970s.

An early response will allow us time to coordinate the many details associated with organizing such a panel. A number of your colleagues from Canada, England, France and Germany, as well as the United States, have been invited to participate in the panel.

Sincerely yours,

Mrs. Jackie Potts, Chairman
SIGGRAPH, Capital Region

cc:
S. Matsa

A. 20

1133 AVENUE OF THE AMERICAS
NEW YORK, N.Y. 10036
(212) 265-6300

ACM Association for Computing Machinery

April 16, 1970

Reply to: Code 745
Department of the Navy
Naval Ship Research
and Development Center
Washington, D. C. 20007

Dr. V. C. Sobolewski
Electrical Engineering Department
The University of Adelaide
Adelaide, South Australia, 5001

Dear Dr. Sobolewski:

Thank you for tentatively accepting our invitation to participate in the International Panel on Computer Graphics at the ACM National Conference, which will be held in New York at the Hilton Hotel from September 1-3. I hope that you have been successful in obtaining the necessary travel funds and will definitely be a panelist.

The theme for the panel will be "Computer Graphics in the Seventies". After a brief review of the Graphic Conference being held in Brunel, England, this month, each panelist will be allocated twenty minutes in which to give a resume of the national outlook for Computer Graphics in his country.

In the event you can attend the Conference, please send me a two hundred word abstract of your talk and a bibliographical sketch by June 15. If I can be of any assistance, please feel free to contact me.

I am looking forward to meeting you.

Sincerely yours,

Mrs. Jackie Potts, Chairman
SIGGRAPH, Capital Region

JP/ej

cc: S. Matsa

A.13.4 COPIES OF DESIGN NOTES ACCEPTED FOR PUBLICATION.

Accepted by "PLANAR", published by FAIRCHILD SEMICONDUCTORS, Aust.

1. "LINEAR VOLTAGE CONTROLLED ASTABLE AND MONOSTABLE USING THE uL 914".

2. "PULSE LEADING-AND LAGGING-EDGE INDICATORS".

by Victor C. Sobolewski,
Electrical Engineering Department,
University of Adelaide,
Adelaide, South Australia.

Submitted to "Planar"

Accepted 20th October 1969.

A.205

[illegible]CHILD [illegible]IA PTY. LTD.

420 MOUNT DANDENONG ROAD, CROYDON, VICTORIA. P.O. BOX 151, CROYDON 3136. CABLE: FAIRCHILD MELBOURNE.
TELEPHONE: 722 4131

JWH/JM

20th October, 1969.

Mr. V. C. Sobolewski,
Electrical Engineering Dept.,
The University of Adelaide,
Adelaide,
South Australia. 5000.

Dear Sir,

Thank you for your correspondence of September 22nd and your interest in including this material in Planar.

Our engineers have thoroughly read the article and, in conjunction with them, can I make the following requests?

1. Could you summarize the article down to about 750 words? We could then supply complete information to any requesters (with your permission of course).

2. Would you provide a short personal history? This helps to bring the article closer to the reader and is also good for the author.

I have taken photo-copies of your article and return your original notes.

Thanking you for your interest and looking forward to hearing from you.

Yours sincerely,

JOHN W. HOUSTON
Marketing Services Supervisor

"LINEAR VOLTAGE-CONTROLLED ASTABLES AND MONOSTABLES"

by V.C. Sobolewski,
Electrical Engineering Department,
University of Adelaide.

INTRODUCTION:

Simple, efficient voltage-controlled oscillators and voltage-controlled time-delays can be made using the Fairchild UL914. The control voltage-output relationship is extremely linear over a wide range (see Figs.4,5,7).

Both circuits are very similar in operation to the usual collector-coupled astable and monostable circuits, the only difference being a transistor inserted between the timing capacitor and resistor, in common-base connection (transistors "Q_3", "Q_4" in Fig.2), which provides a constant-current source for the discharge of the timing capacitor.

This constant current "I_C", say from "Q_3", is clearly given by

$$I_C = \alpha\Big]_{Q_3} \cdot \frac{(V_{BB} - V_{BE}\Big]_{Q_3} - V_B)}{R_{BB_2}} \qquad .....(1)$$

Similarly for "Q_4".

Due to the symmetry of the circuit in Fig.2, only the right half of the circuit need be considered.

Across "C_2", the usual relationship "$i = C_2 \cdot \frac{dV}{dt}$" holds. For a constant current "I_C" through "C_2" this can be rewritten as

$$T = C_2 \cdot \frac{V_x}{I_C} \qquad .....(2)$$

where "T" is the time interval associated with the voltage change "V_x", and "V_x" is the voltage variation across "C_2".

During the quasi-stable state of "Q_2", "V_x" is clearly given from Fig.3 by

$$V_x = (V_{CC} - V_{CE\ sat}]_{Q_1}) - (V_{BE\ sat} - V_{BE\ cutin})_{Q_2} \qquad .....(3)$$

where "$V_{BE\ sat}$" etc., include the voltage drops across "R_{B_2}" due to "I_{C_2}".

Substituting for "I_C" and "V_x" into (2) results in

$$T_2 = \frac{R_{BB2} \cdot C_2}{\alpha]_{Q_3}} \cdot \frac{(V_{CC} - V_{CE\ sat}]_{Q_1}) - (V_{BE\ sat} - V_{BE\ cutin})_{Q_2}}{V_{BB} - V_{BE}]_{Q_3} - V_B} \qquad .....(4)$$

which is the fundamental timing equation for both circuits.

VOLTAGE-CONTROLLED ASTABLE (FIG.2)

For the astable, the quasi-stable period of transistor "Q_1" is given by a similar expression as (4). The repetition rate is then "$f = \frac{1}{T_1 + T_2}$", and assuming $R_{BB1} = R_{BB2} = R_{BB}$, $C_1 = C_2 = C$ and identical transistors Q_3, Q_4 and noting that $Q_3 = Q_4 \approx 1$, and making $V_{BB} = V_{in}$, the control voltage, results in

$$f = \frac{1}{2R_{BB} C} \cdot \frac{V_{in} - \overline{V_{BE}}\ Q_3 - V_B}{V_{CC} - V_{CE\ sat}]_{Q_1} - (V_{BE\ sat} - V_{BE\ cutin})_{Q_2}} \qquad .....(5)$$

Keeping "V_{CC}", and "V_B" constant, (5) is of the required form

$$f = K_1 \cdot V_{in} - F_o \qquad .....(5a)$$

Any deviations from linearity are due to the various transistor voltages "V_{BE}", "V_{CE}" etc. varying non-linearly with increasing "V_{in}".

A typical "f vs Vin" curve is of the form as in Fig.4, while two specific cases are shown in Fig.5.

The circuit operates so long as "Vin" provides adequate collector current "$I_C]_{Q_3}$" (and hence base current "I_{B_2}") to keep "Q_2" in saturation (and similarly "Q_1" need be kept in saturation). Thus it is required that

$$I_b]_{Q_2} \geqslant \frac{I_C]_{Q_2}}{h_{FE}]_{Q_2}} \qquad \text{.....(6)}$$

which with a slight approximation reduces to

$$Vin - V_{BE}]_{Q_3} - V_B \geqslant \frac{R_{BB2}}{R_{C2}} \cdot \frac{V_{CC} - V_{CE\ sat}]_{Q_2}}{h_{FE}]_{Q_2}} \qquad \text{....(6a)}$$

It can also be seen from the circuit that

$$V_B + V_{BE}]_{Q_3} = V_{CE}]_{Q_3} + V_{BE\ sat}]_{Q_2} \qquad \text{.....(7)}$$

From (7), minimizing "$V_{CE}]_{Q_3}$" (by keeping "Q_3" in saturation) minimizes "V_B" (typically "$V_{B\ min}$" ≈ 0.3 - 0.4V), and hence from (6a) minimizes "Vin". With R_{BB} = 12K, "Vin_{min}" is approx. 2.6 -2.7V.

The <u>lower frequency limit</u> corresponds to "Vin_{min}". Multiplying both sides of (6a) by

$$\frac{\alpha]_{Q_3}}{2R_{BB_2} \cdot C_2} \cdot \frac{1}{(V_{CC} - V_{CE\ sat}]_{Q_1}) - (V_{BE\ sat} - V_{BE\ cutin})_{Q_2}}$$

results in

$$f_{min} \geqslant \frac{\alpha]_{Q_3}}{C_2 \cdot R_{C2} \cdot h_{FE}]_{Q_2}} \cdot \frac{(V_{CC} - V_{CE\ sat}]_{Q_1})}{(V_{CE} - V_{CE\ sat}]_{Q_1}) - (V_{BE\ sat} - V_{BE\ cutin})_{Q_2}} \qquad \text{....(8)}$$

$$\geqslant \frac{1}{R_{C2} \cdot C_2 \cdot h_{FE}]_{Q_2}} \qquad \text{....(8a)}$$

which is independent of supply voltages etc.

The upper frequency limit (for reasonable linearity) occurs when the halfperiod of the output frequency (T_1 or T_2) equals or becomes comparable with the period for the overshoots at "$V_{BE}]Q_2$" to decay to their steady state value ($V_{BE\ sat}]Q_2$), and for the collector voltage of the OFF transistor "Q_1" to fully reach "V_{CC}" (see Fig.3).

The time constant of the overshoots and rise time is (ref.2)

$$\tau' = C_2\ (R_{C1} + R_{B2} + r_{bb_2}) \qquad .....(9)$$

where "r_{bb}" is the base spreading resistance of "Q_2". The steady-state values (less than 1% from the final value) are reached in $t \geqslant 4.5\tau'$.

Hence the maximum linear half-period is

$$T_{max} \geqslant 4.5\ C_2\ (R_{C1} + R_{B_2} + r_{bb_2}) \qquad .....(10)$$

giving

$$f_{max} \leqslant \frac{1}{9\ C_2\ (R_{C1} + R_{B_2} + r_{bb_2})} \qquad .....(11)$$

The maximum useful linear range is then (from 8(a) and (11))

$$\frac{f_{max}}{f_{min}} \leqslant \frac{R_{C1} \cdot hFE]Q_2}{4.5\ (R_{C1} + R_{B_1} + r_{bb_1})} \qquad .....(12)$$

which is purely dependent on UL914.parameters. With typical values substituted ($hFE]Q_2 \approx 50$, $r_{bb} \approx 100\Omega$), the linear range is approximately 6:1. In practice for ±1% linearity limits this is approximately 3-4:1. The input voltage range is from about 2.6-2.7V to about 6.5-7V. Two specific cases are shown in Fig.5.

VOLTAGE-CONTROLLED MONOSTABLE. (FIG.6)

Removing "R_{BB_1}", "C_1" and "Q_4" and connecting the collector of "Q_2" directly to "R_{B_1}", the previous circuit results in a monostable.

A positive trigger pulse (>1.2V) is fed into "R_{B_1}". The quasistable state, initiated by the trigger pulse is given by (4). Keeping "V_{BB}" constant ($\approx$ 3.6V) and varying "V_{CC}", the control voltage, results in (4) taking the form of

$$T = K_2 . V_{in} - T_o \qquad .....(13)$$

The deviations from linearity are again due to transistor voltages varying non-linearly with increasing "Vin". As previously, the circuit operates satisfactorily when "Q_2" is kept in saturation during the stable state, and "Q_1" during the quasistable state.

For the quasistable state then

$$I_b]_{Q_1} \geqslant \frac{I_C]_{Q_1}}{h_{FE}]_{Q_1}}$$

from which it can be shown that

$$V_{in\ min} \geqslant V_{BE\ sat}]_{Q_1} \qquad .....(14)$$

and is thus of the order of 0.6-0.7V.

For the stable state, it is required that

$$I_b]_{Q_2} \geqslant \frac{I_C]_{Q_2}}{h_{FE}]_{Q_2}}$$

i.e.

$$\frac{V_{CC} - V_{BE}]_{Q_3} - V_B}{R_{BB2}} \geqslant \frac{V_{in} - V_{CE\ sat}]_{Q2}}{h_{FE}]_{Q_2} \cdot R_{C2}} \qquad .....(15)$$

from which the maximum value of "R_{BB}" can be found if "V_{in}" is at its maximum value.

As "$V_{BE\ sat}\rfloor Q_2$ - $V_{in} - V_{CE\ sat}\rfloor Q_1$" appears at the base of "Q_2" (Fig.3) then,

$$V_{in} - V_{CE\ sat}\rfloor Q_1 - V_{BE\ sat}\rfloor Q_1 \leqslant B\ V_{EBO}\rfloor Q_2 \qquad .....(16)$$

Hence for typical values for the UL914, this makes $V_{in\ max} \approx 6.5V$, and from (15) limits $R_{BB} \leqslant 12.5K$.

Analogously to the case of the astable, the various overshoots and finite rise times limit the linear range of operation. A maximum trigger rate can be defined for linear operation which is

$$T_{TRIGGER} \geqslant T_{max} + 4.5\,\tau' \qquad .....(16a)$$

from which the maximum duty cycle (linear range) is

$$(1 - \frac{4.5\,\tau'}{TRIGGER\ MAX}) \times 100\% \qquad .....(17)$$

If "T_{max}" is made to occur at "$V_{in} \approx 6.5V$" this gives a maximum duty cycle of about 80-85%. The maximum linear range, defined by "$\frac{T_{max}}{T_{min}}$" is obtained by substituting "$V_{in\ min}$" and $V_{in\ max}$" into (4). With "$V_{in\ min} \approx 0.7V$", "$V_{in\ max} \approx 6.5V$," this gives "$\frac{T_{max}}{T_{min}} \approx 18\text{-}20:1$"; for $\pm 1\%$ linearity it is about 8:1. Two cases are shown in Fig.8.

"V_{bmin}" is obtained from (7) but evaluated at "$V_{BE\ sat}\rfloor Q_2$) $V_{in\ max}$" giving "$V_{b\ min} \approx 0.6 - 0.7V$."

The output at "A" varies with "V_{in}". Using the Superposition Theorem,

$$Vout = \frac{(V_{in} - V_{CE\ sat}\rfloor Q_2) R_{B_1}}{R_{B_1} + R_{C2}} + \frac{V_{BE\ sat}\rfloor Q_2 \cdot R_{C2}}{R_{B_1} + R_{C2}} \qquad .....(18)$$

$$\approx 0.4\ V_{in} + 0.5 \qquad \ldots\ldots(18a)$$

when typical values are substituted.

REFERENCES:

(1) BIDDLECOMB, R.W. "Latest Multivibrator Improvement", Electronics, Vol.36, pp.64-65, April, 1963.

(2) MILLMAN, J. and TAUB, H. "Pulse, Digital and Switching Networks" McGraw Hill, New York, 1965, Chapter 11.

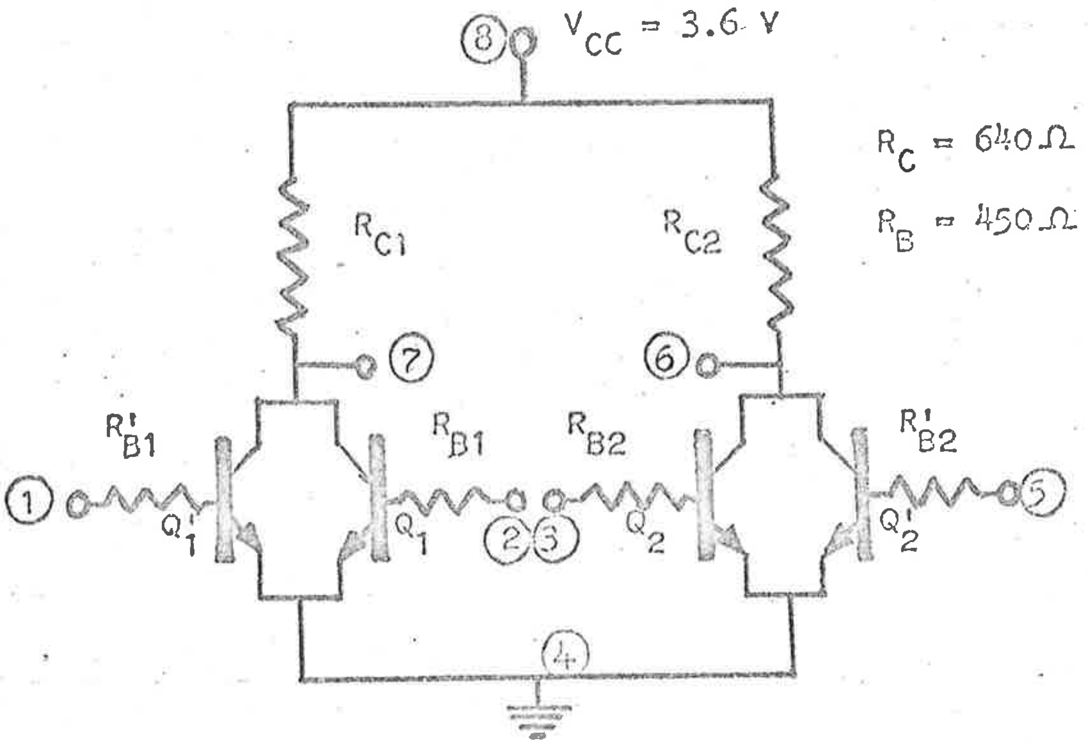


Figure 1. Circuit of Fairchild uL 914 with typical component values. Numbers 1-8 refer to pin numbers.

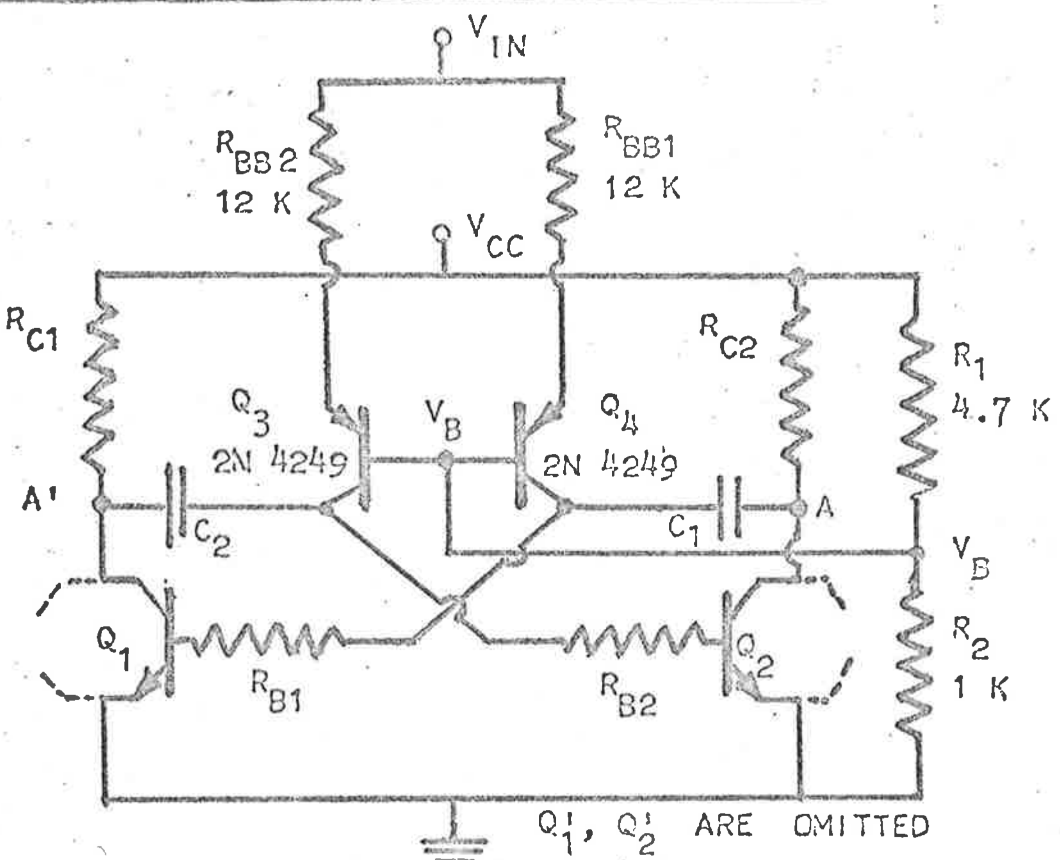


Figure 2. Voltage Controlled Astable using a uL 914.

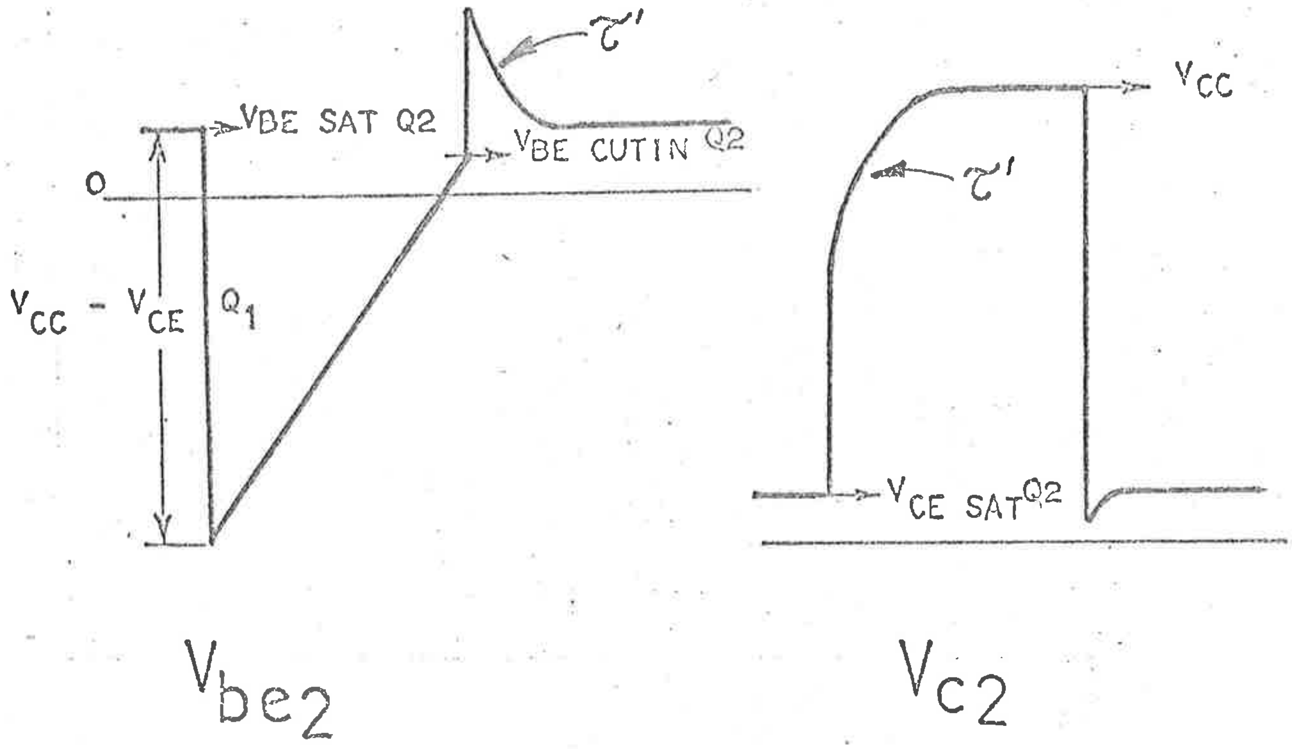


Figure 3. Typical voltage waveforms at Q_2. $\tau' = C_2 . (R_{C1} + R_{B1} + r_{BB2})$.

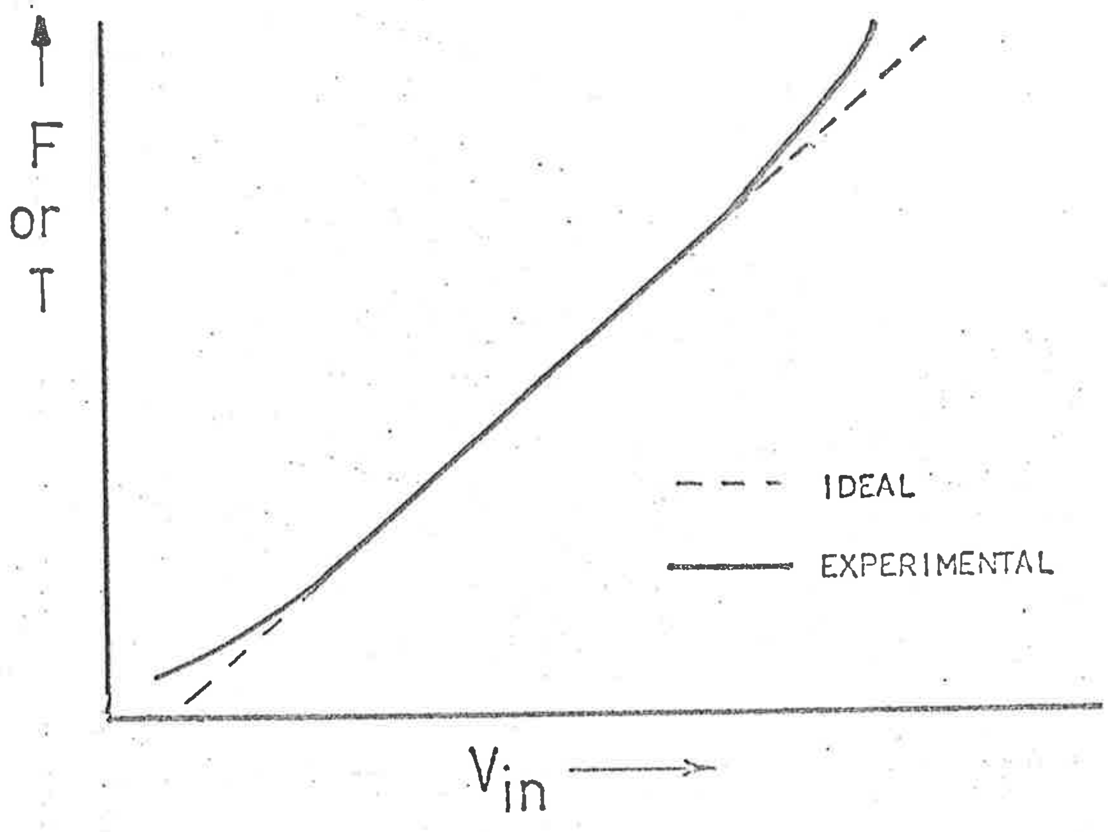


Figure 4. Typical Input-Output characteristic.

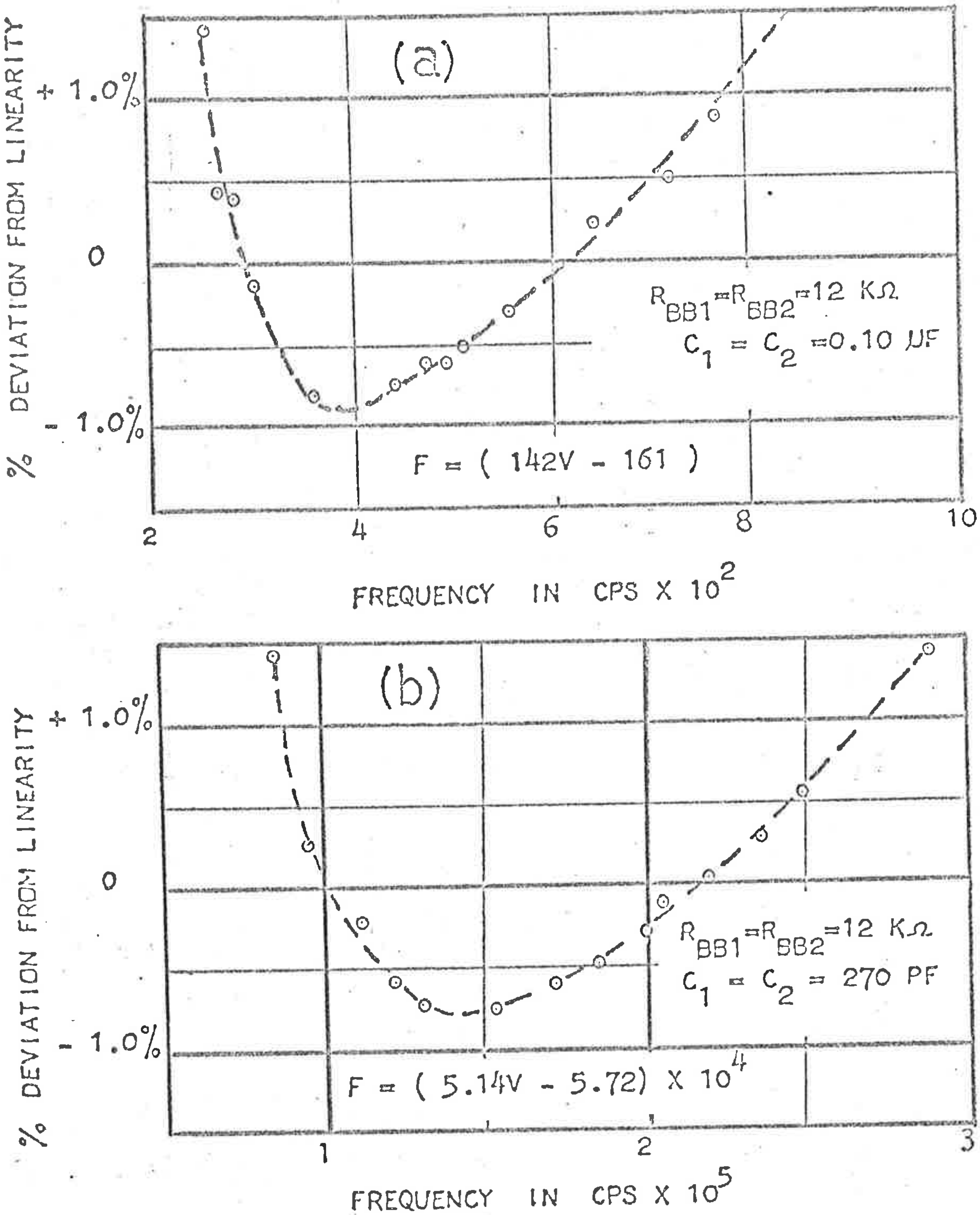


Figure 5. Output Frequency of V.C. Astable shown in Figure 2 indicating the deviation from linearity.

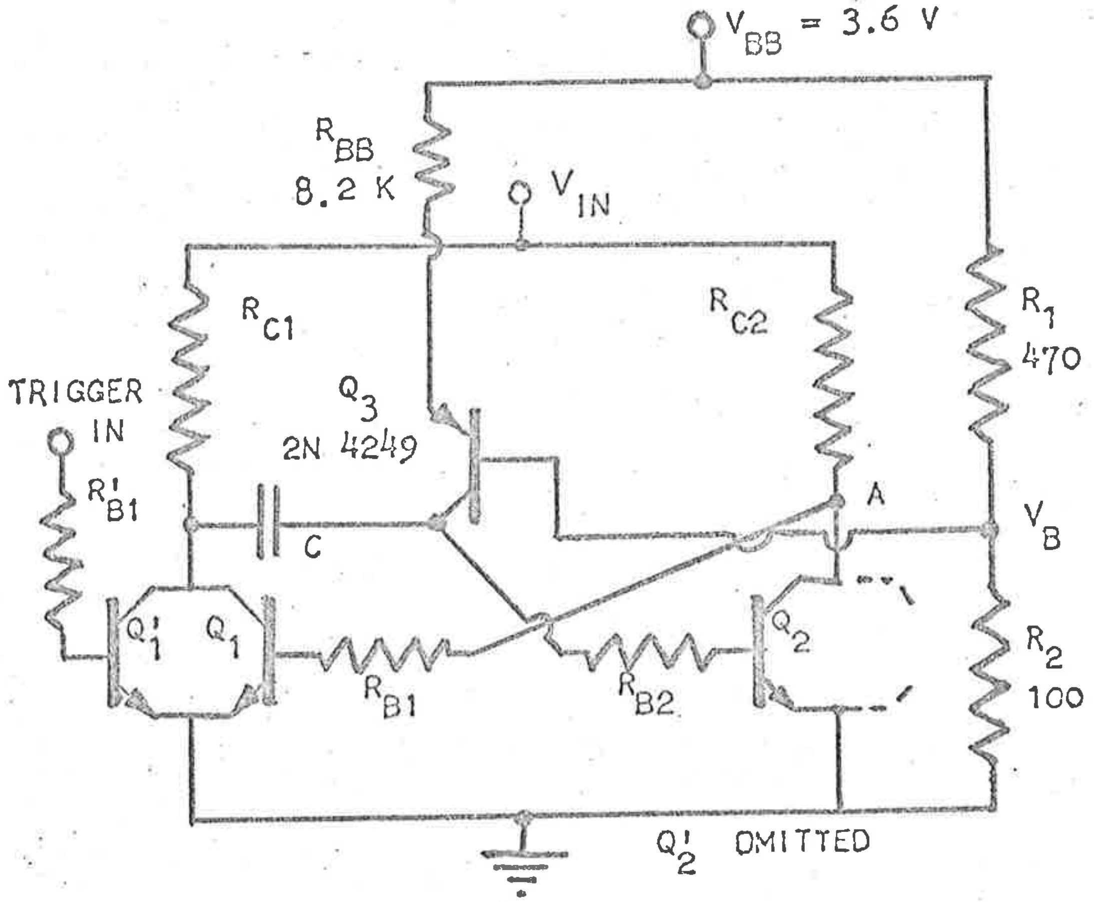


Figure 6. Voltage Controlled Monostable using a uL 914. Output is at "A", the collector of Q_2.

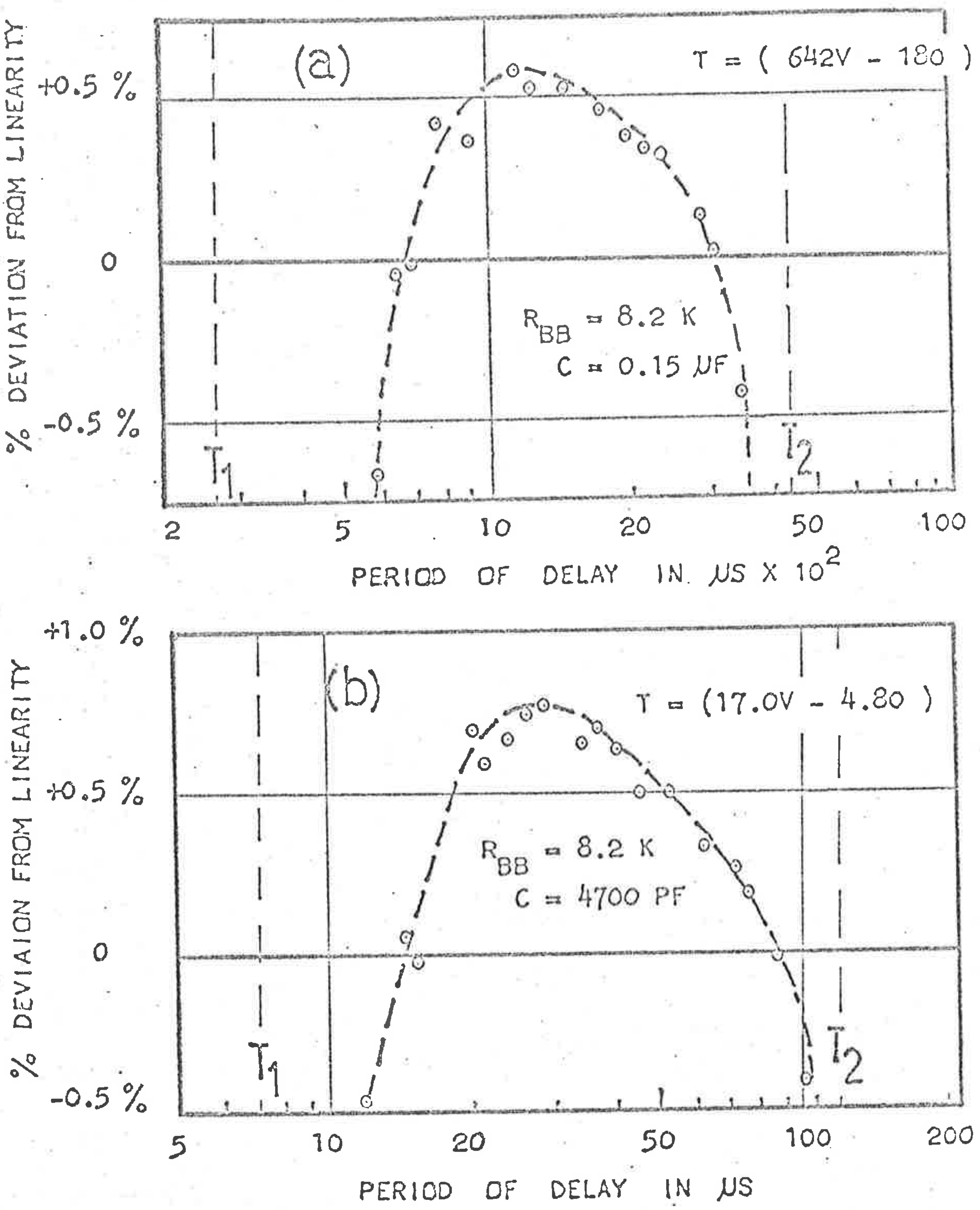


Figure 7. Pulse width Output of the V.C. Monostable shown in Figure 6 indicating the deviation from linearity.

PULSE LEADING- AND LAGGING-EDGE INDICATOR

by V.C. Sobolewski
Electrical Engineering Department,
University of Adelaide.

1. LAGGING-EDGE INDICATOR (fig. 1)

For the "lagging-edge indicator" (i.e. level change from "High"-to-"Low" voltage), "R_1" and "R_2" are so chosen as to keep the particular half of a UL914 just in saturation (typically $V_B > 0.8$-$0.9V$), and such as to minimise the supply current.

"C" in conjunction with "R_1", "R_2" and "Zin" form a differentiating ("High-pass") network. For an output pulse-width of 50-60ns, "C" is of the order of 20-30pF. (The pulse width is limited by the fall time of the input pulse, and is delayed by the usual switching transmission delay). The differentiated pulse switches the UL914 OFF, generating the positive output pulse of 3.6V amplitude.

2. LEADING-EDGE INDICATOR (fig. 2)

In a "leading-edge indicator" (i.e. level change from "Low" to "High" voltage), the input is first inverted (via the other half of the UL914) and the result fed into the above circuit again.

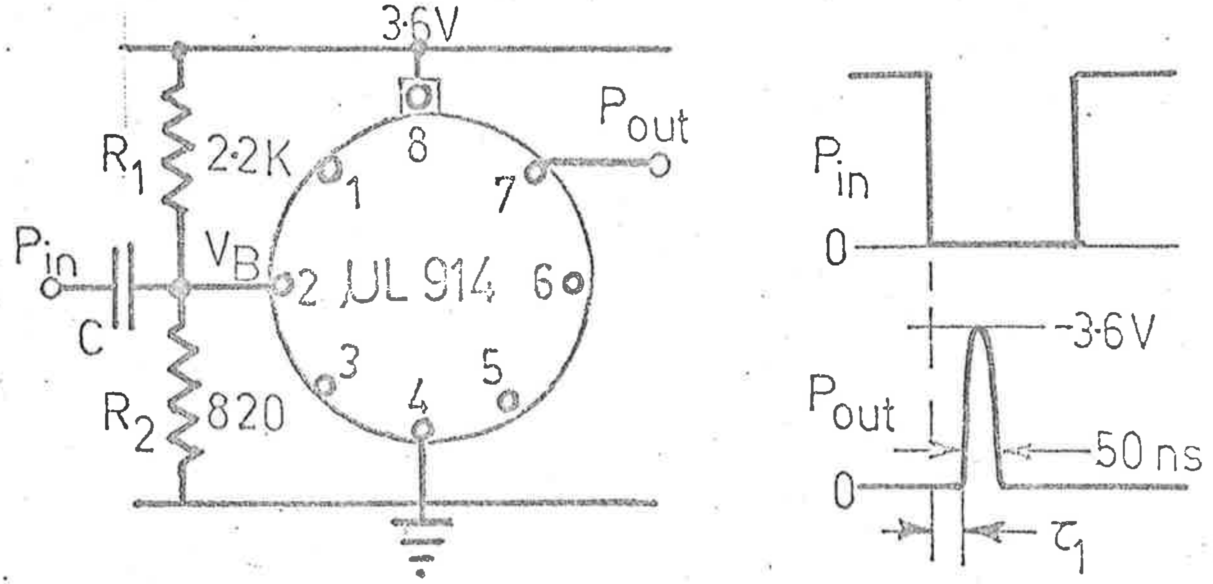


Figure 1. "Lagging Edge Indicator". "τ_1", the total propagation delay, is half the propagation delay of a uL 914.

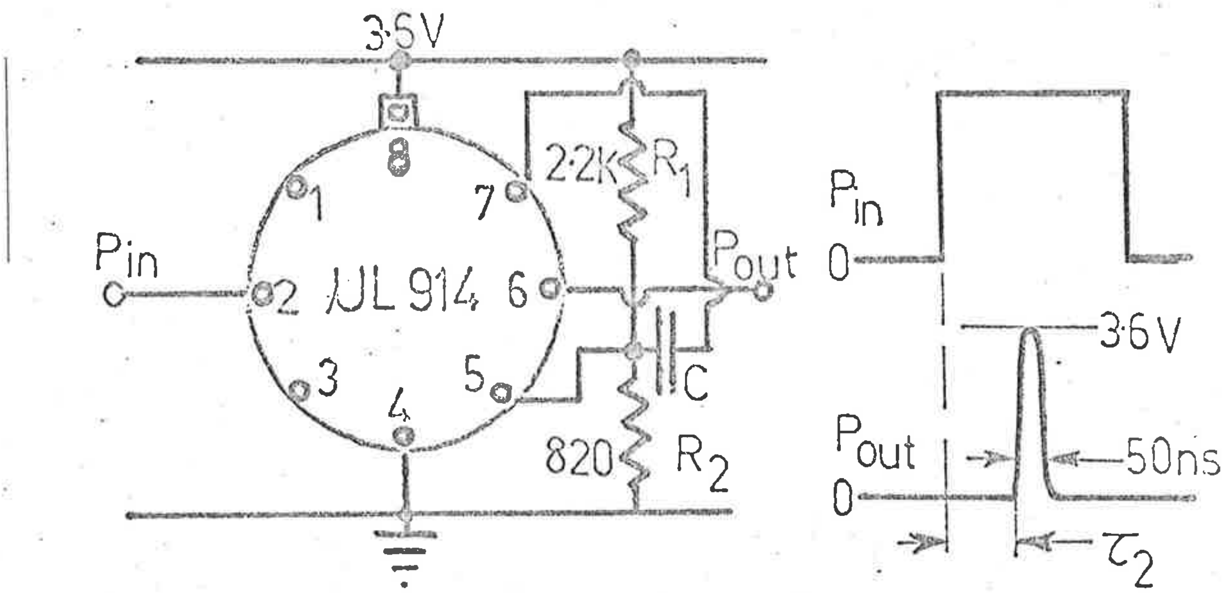


Figure 2. "Leading Edge Indicator". $\tau_2 \approx 2\tau_1$.

A.13.5 COPY OF SHORT DESIGN NOTE ACCEPTED FOR PUBLICATION.

Accepted by "Electronic Design", New York.

"LOW VOLTAGE-RAIL VOLTAGE CONTROLLED MONOSTABLE"

by Victor C. Sobolewski,
Electrical Engineering Department,
University of Adelaide,
Adelaide, South Australia.

Submitted to 'Electronic Design'

Accepted November 3rd 1969
'Ideas for Design'

A.221

Electronic Design

HAYDEN PUBLISHING COMPANY, INC. 850 THIRD AVENUE, NEW YORK, N.Y. 10022 PLAZA 1-5530

November 3, 1969

Mr. Victor C. Sobolewski
Research Student
Electrical Engineering Dept.
The University of Adelaide
Adelaide, S.A. 5000
Australia

Re: Your Idea for Design

SEP-15

Dear Sir:

I have reviewed your Idea for Design entry

"Low Voltage-Rail Voltage..."

and am happy to accept it for publication in ELECTRONIC DESIGN.

Unfortunately, publication will be delayed for a few months due to the large backlog of Ideas for Design we now have on hand. Prior to publication, though, you will receive a copy of the final edited version of your material for approval. In addition, you will receive payment for your entry shortly.

You are now eligible for the "Most Valuable of Issue Award," as determined by our readers. The announcement of the winning entry appears approximately six to eight weeks after its initial publication. All winning entries become eligible for the "Idea of the Year" award of $1000.00.

Thank you again for sending us your Idea for Design and please keep us in mind when you have another.

Cordially,

Don Mennie
Idea for Design Editor

"LOW VOLTAGE RAIL VOLTAGE CONTROLLED MONOSTABLE"

The operation of the circuit is very similar to the usual collector-coupled monostable, the only difference being the transistor "Q_3" which, with the "hard" voltage source "V_B" (due to "R_1" and "R_2"), provides a constant current source to discharge the timing capacitor "C." The constant current is

$$I_C = \alpha\rbrack_{Q_3} \cdot \frac{V_{CC} - V_{BE}\rbrack_{Q_3} - V_B}{R_{BB}} \qquad .....(1)$$

The quasi-stable state defined by the compound transistor "$Q_2 - Q_2'$" is easily found to be of duration

$$T = \frac{R_{BB}C}{\alpha\rbrack_{Q_3}} \cdot \frac{(V_{in} - V_{CE\ sat}\rbrack_{Q_1}) - (V_{BE\ sat} - V_{BE\ cutin})Q_2\text{-}Q_2'}{V_{CC} - V_{BE}\rbrack_{Q_3} - V_B} \qquad .....(2)$$

which is of the form

$$T = K.V_{in} - T_o \qquad .....(3)$$

The deviations from linearity are due to the various transistor voltages "V_{BE}", "V_{CE}" etc. varying non-linearly with "V_{in}".

The poor linearity at low "V_{in}" is due to this and also due to the difficulty in locating precisely the leading and falling edges of the output pulse. The lower limit of "V_{in}" occurs when it approaches "$V_{BE}\rbrack_{Q_1}$" and is of the order of 0.3V.

The upper linearity limit is caused by the incomplete recovery of the voltage at the base of "Q_2" to "$V_{BE\ sat}\rbrack_{Q_2}$", and the collector voltage of "Q_1" to "V_{in}", when the next trigger pulse is applied. However these recovery transients are short, giving good linearity up to 97% of duty cycle.

The absolute upper limit for "V_{in}" is determined by "BV_{EBO}" of "Q_2"; as the initial OFF voltage appearing at the base of "Q_2" (at the trigger pulse instant) is "$V_{BE\ sat}]_{Q_2} - V_{in} - V_{CE\ sat}]_{Q_1}$", then

$$V_{in\ max} \leqslant BV_{EBO}]_{Q_2} + V_{CE\ sat}]_{Q_1} + V_{BE\ sat}]_{Q_2} \qquad .....(4)$$

The Darlington connection effectively increases "$V_{in\ max}$" by a factor of 2 over a single transistor.

The circuit operates so long as "V_{CC}", "R_{C2}", "R_{B1}" provide adequate base current for "Q_1" to saturate up to "$V_{in\ max}$" during the quasi-stable state - from this "R_{C1}" is obtained.

Similarly during the stable state it is required that,

$$V_{BE}\ Q_1 \leqslant V_{BE\ cutin}\ (\approx 0.4\ V) \qquad .....(5)$$

For some "V_{CC}", "R_{C2}" and "R_{B1}" can be obtained.

With a full range of "V_{in}" from 0.2-0.3V → 25V, the linear range (±1%) is of the order of 100:1. For the widest range, "V_B" is to be kept to a minimum; this can be obtained from

$$V_B + V_{BE}]_{Q_3} = V_{CE}]_{Q_3} + V_{BE\ sat}]_{Q_2 - Q_2'} \qquad .....(6)$$

The output amplitude is constant and is of the order of 1.7V for the values shown (it is obtained by superposition of the voltages "V_{CC}", "$V_{BE}]_{Q1}$"). The circuit thus is compatable with I.C.'s.

. . . .

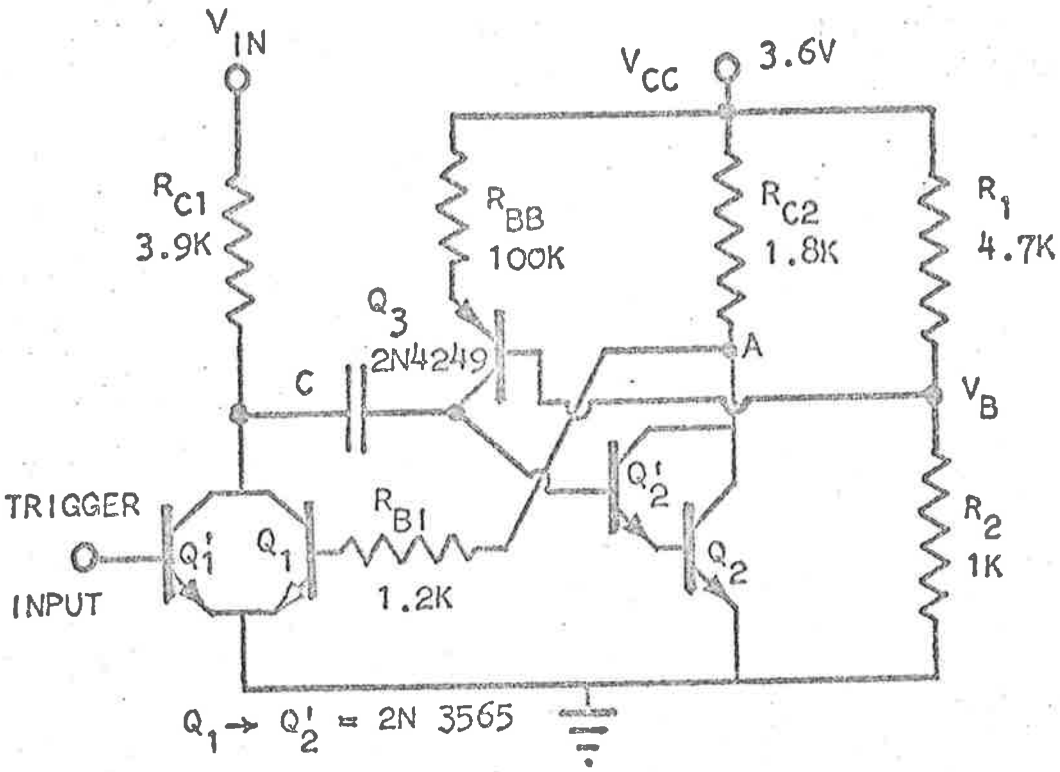


Figure 1. Linear Voltage Controlled Monostable.

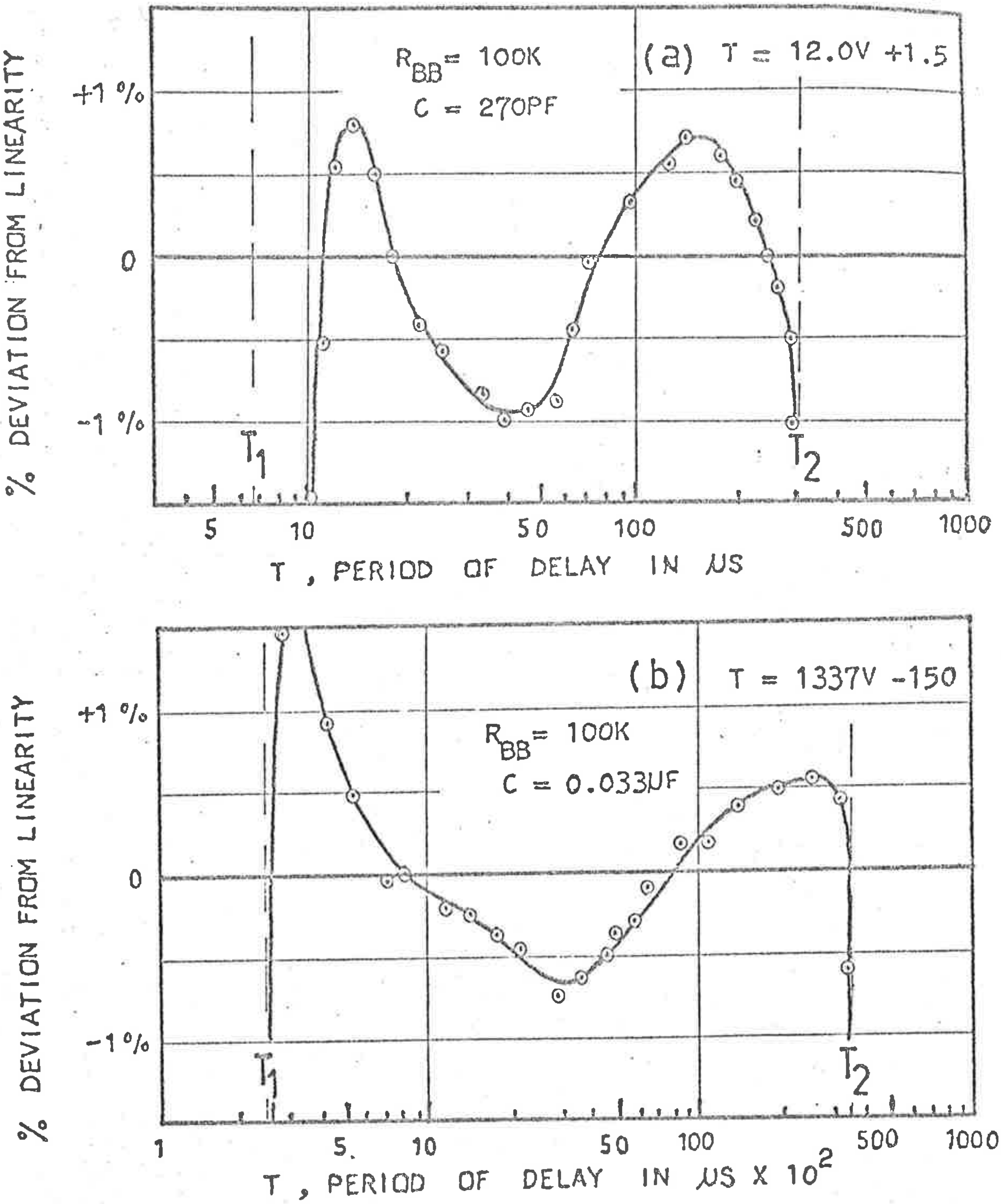


Figure 2. "T_1" is the delay at $V_{in\ min} \approx 0{\cdot}3$ V. "T_2" is the Trigger Pulse period and corresponds to $V_{in\ max} \approx 25$ V.

A.13.6 MISCELLANEOUS LETTERS

DEPARTMENT OF THE NAVY
NAVAL TRAINING DEVICE CENTER
ORLANDO, FLORIDA 32813

IN REPLY REFER TO:
53:JCM
16 November 1970

Mr. V. C. Sobolewski
Department of Electrical Engineering
University of Adelaide
Adelaide, South Australia

Re: Using Moire Patterns to Determine the Distortion of Graphic Displays and Graphic Input Devices, IEEE Vol. 58, No. 4, April 1970

Dear Mr. Sobolewski:

Your well written article was of considerable interest. I especially appreciate your reference to my letter in the IEEE Proceedings.

Mr. Bob Phillips, the head of an associate electronic research laboratory here at NTDC told me he has exchanged letters with you. He is using our joint literature toward improving television generated visual displays for training purposes.

My further work in the area of image plane distortion is a publication in the September 1970 SMPTE Jounral. (Society of Motion Picture and Television Engineers). (Copy enclosed.)

I look forward to seeing your name in print again and also to the application of the principles we have uncovered.

With sincere respect,

J. C. McKechnie
Electronic Research Engineer
Visual Simulation Laboratory

TDC 5216/9

NAVAL TRAINING DEVICE CENTER
ORLANDO, FLORIDA 32813
AREA CODE 305
TELEPHONE: 841-5611

IN REPLY REFER TO:
52:ENP
10 June 1970

Mr. V. C. Sobolewski
Department of Electrical Engineering
University of Adelaide
Adelaide, South Australia

Dear Sir:

Your paper entitled "Using Moire Patterns to Determine the Distortion of Graphic Displays and Graphic Input Devices" on pages 567 through 583 of the April 1970 Proceedings of the Institute of Electronic and Electrical Engineers has been read by the members of this laboratory with great interest, as one of the main interests here is television techniques.

This note is a request for reprints of your article. If you can spare them, two or three reprint copies would be appreciated. Further, photographic reproductions of the plates used in figures 8, 10, and 15 would be of use if such are available. It is realized that this request transcends the usual courtesy responsence of an author. Should there be a charge for these reprints, please advise. Do not send until you have received a formal purchase request.

Finally, may I congratulate you on the publication of your article; this is certainly an extension of measurement much needed in this phase of electronics, and long delayed.

Thank you.

Sincerely yours,

E. N. PHILLIPS
Head, Electronics Laboratory
By direction of
the Commanding Officer

REFERENCES

GENERAL BIBLIOGRAPHY

(B.1) R. D. Parslow, R. W. Prowse, R. Elliot Green eds., "Computer Graphics - Techniques and Applications", Plenum Press, London and New York, 1969.

(B.2) H. R. Luxenberg, R. L. Kuehn, "Display Systems Engineering", McGraw-Hill, New York, 1968.

(B.3) B. Kazan, M. Knoll, "Electronic Image Storage", Academic Press, New York, 1968.

(B.4) H. H. Poole, "Fundamentals of Display Systems", Spartan Books, Washington DC, 1966.

(B.5) N. Stadtfeld, "Information Display Concepts", Tektronix, Beaverton, Oregon, Jan, 1969.

(B.6) F. Burgess, ed., "Out-line Computing Systems", American Data Processing Inc., 1965.

(B.7) D. Secrest and J. Nievergett eds., "Emerging Concepts in Computer Graphics," 1967 U. of Illinois Conference W.A. Benjamin New York, 1968.

(B.8) F.F. Kuo ed., "System Analysis by Digital Computer", John Wiley, New York, 1967.

REFERENCES

(1) I. E. Sutherland,"Computer Graphics", Datamation Vol. 12, 5, May 1966, pp 22-27

(2) T. P. Dankowski, A. G. Lippert,"Computer Graphics", (in two parts) Machine Design, Pt. 1, April 17, 1969 pp 226-232, Pt. 2 May 1, 1969 pp148-155

(3) E. L. Jacks,"Observation on a Graphic Console System", in"Emerging Concepts of Computer Graphics" - U. of Illinois Conference (W.A. Benjamin, New York 1968) pp 131-170

(4) S. H. Chasen,"The Role of Man-Computer Graphics in the Design Process", Locheed-Georgia Research Laboratories Report, Feb. 1969.

(5) C. Strachey,"Time-sharing in Large, Fast Computers", Proc. 1st IFIP Conf., June 1959, paper B.2.19.

(6) R. M. Fano,"The MAC system: the Computer Utility Approach", IEEE Spectrum, Vol. 2, Jan. 1965 pp 56-64

(7) R. M. Davis,"A History of Automated Displays", Datamation, Vol. 11, 1, Jan. 1965 pp 24-28

(8) W.H.P. Leslie, "Computer Graphical Design Aids", J. Roy. Aero. Soc., Vol. 71, 4, April 1967 pp 237-244

(9) A. van Dam, "Computer Driven Displays and their Use in Man/Machine Interaction", in Advances in Computers, Vol. 7, (Academic Press, New York 1966) pp 239-290

(10) E. L. Jacks,"The DAC-1; A Computer System for the Study of Graphical Man-Machine Communications", G-M Eng. Jnl, Vol. 12, 2, 2nd Quarter 1965 pp 2-8

(11) G. J. Devene, B. Hargreaves, D. M. Walker,"The DAC-1 System", Datamation, Vol. 12, 6, June 1966 pp 37-47

(12) E. L. Jacks, "A Laboratory for the Study of Graphical Man-Machine Communications", Proc F.J.CC, Vol. 26, 1964 pp 343-350

(13) B. Hargreaves et al., "Image Processing Hardware for Man-Machine Graphical Communications", Proc. F J C C Vol. 26, 1964 pp 363-386

(14) T. R. Allen, J. E. Foote, "Input/Output Software capability for a Man-Machine Communication and Image Processing System", Proc F J C C, Vol. 26, 1964 pp 387-396

(15) F. N. Krull, J. E. Foote, "A Line Scanning System Controlled from an On-line System", Proc. F J C C, Vol. 26, 1964 pp 397-410

(16) M. P. Cole, P. H. Dorn, R. Lewis, "Operational Software in a Disc Oriented System", Proc. F J C C, Vol. 26, 1964 pp 351-362

(17) Special Issue "Computer-Aided Design", Lockheed-Georgia Quarterly, Summer 1965.

(18) W. A. McConnell, "Computer Graphics: An Assembly Line for Information", American Documentation, July 1968 pp 278-285 (Ford Approach)

(19) W. A. Fetter, "Computer Graphics in Communications", (McGraw-Hill, New York 1966) (Boeing approach)

(20) S. A. Coons, "An Outline of the Requirements for a Computer-Aided Design System", Proc S J C C 1963 Vol. 23, pp 299-304

(21) D. T. Ross, J. E. Rodriguez, "Theoretical Foundations for the Computer-Aided Design System", Proc S J C C 1963, Vol. 23, pp 305-322

(22) D. T. Ross, "The AED Approach to Generalized Computer Aided Design", Proc ACM Natl. Meeting 1967 pp 367-385

(23) "The Computer as a Design Tool - the MIT approach",

Electro-Technology, Nov. 1963. pp 112-115

(24) R. Stotz, "Man Machine Console Facilities for Computer Aided Design", Proc. S.J.C.C, Vol. 23, 1963 pp 323-328

(25) J. E. Ward, "Display Hardware for Dynamic Man-Machine Interaction", 20th Annual A F C E A Conv., June 1966

(26) I. E. Sutherland, "Sketchpad - A Man-Machine Graphical Communication System", Proc. S J C C, 1963, Vol. 23, pp 329-346

(27) T. E. Johnson, "Sketchpad III - A computer Program for Drawing in Three Dimensions", Proc. S J C C 1963, Vol. 23, pp 347-353

(28) L. G. Roberts, "Machine Perception of Three Dimensional Solids", Lincoln Labs Tech. Report No 315 M.I.T. 1963

(29) W. R. Sutherland, "on-line Graphical Specification of Computer Procedures", Lincoln Labs. Tech. Report No. 405 M.I.T., May 1966.

(30) I. E. Sutherland, "Computer Inputs and Outputs", Sci. Am., Vol. 215, 3, Sept. 1966 pp 86-96

(31) G. J. Culler, B. D. Fried, "An On-line Computing Center for Scientific Problems", TRW Report M19-3U3, 1963, Cleveland, Ohio.

(32) "The TRW Two-Station On-Line Scientific Computer", AD 609-720 CFSTI Report, Dec. 1964.

(33) "Special Issue on Computers" Proc. IEEE, Vol. 49, 1, Jan 1961.

(34) W. F. Bauer, D. L. Gerlough, J. W. Granholm, "Advanced Computer Applications", pp 296-304 in (33).

(35) R. T. Loewe, R. L. Sisson, P. Horowitz, "Computer Generated Displays" pp 185-195 in (33).

(36) "Special Issue on Computers," Proc IEEE, Vol 54, 12, Dec. 1966.

(37) "Special Issue on Computer-Aided Design," Proc IEEE Vol. 55, 11, Nov. 1967.

(38) J. Diebold, "What's Ahead in Information Technology," Harvard Bus. Rev., Vol. 43, 5, Sept-Oct. 1965. pp 76-82

(39) A. B. Kamman, "The Uses of Display Terminals for Business Applications," Comp. Aut. April 1970 pp 15-17

(40) M. L. Dertouzos, "CIRCAL; On-Line Circuit Design," Proc. IEEE Vol. 55, 5, May 1967 pp 637-654

(41) M. L. Dertouzos, "An Introduction to On-line Circuit Design," pp 1961-1971 in (37).

(42) H.C.So, "OLCA; An on-line Circuit Analysis System" pp 1954-1961 in (37).

(43) D. Christiansen, "Computer-aided design: part 1," Electronics, Vol. 39, 19, Sept. 19, 1966 pp 110-123

(44) "An hour's work with an interactive computing system," Electronic Design, Feb. 1., 1969 p 32

(45) E. Wolfendale, "REDAC - A Computer-Aided design System incorporating Optimization and Information Retrieval", paper presented at the IREE Convention, Sydney, May 1969.

(46) N. Lindgren, "Human Factors in Engineering: part II," IEEE Spectrum, Vol. 3, 4, April 1966 pp 62-72

(47) F.F. Kuo, W. G. Magnuson, W. J. Walsh, "Computer Graphics in Electronic Design," Datamation, March 1969 pp 71-79

(48) J. Katzenelson, "AEDNET: A Simulator for Non-linear Networks," Proc. IEEE, Vol. 54, 11, Nov. 1966 pp 1536-1552

(49) D. S. Evans, J. Katzenelson, "Data Structures and Man-Machine Communication for Network Problems," Proc. IEEE, Vol. 55, 7, July 1967 pp 1135-1144

(50) R. A. Dixon, J. E. Ironside, "On-line Computer Aided Control System Design Using a Graphical Display", Intl. Conf. on Comp.Aided Design, Southampton, I.E.E, London. pp 140-151

(51) J. A. Weaver, "Circuit Specification on A Computer Display", ibid. pp 579-588

(52) A. De Mari, "On-line Computer Active Network Analysis and Design in Symbolic Form", Conference on Computerized Electronics, Cornell U. New York, August 1969, Session II - paper No. 4.

(53) T. J. Kobylavz, "Non-linear Network Computer Analysis", in (B.8) Chapt 6. pp 197-240

(54) C. O. Harbourt, "Computer-aided design: part 5 Doing a model job", Electronics, Vol. 40, 2, Jan 23, 1967 pp 82-87

(55) M. E. Daniel, "Develop useful general models", Electronic Design, Sept. 27, 1969 pp 68-72

(56) F. S. Taylor. "The U K Scene", pp 85-89 in (B.1)

(57) C. Machover. "Computer Graphics in the Unites States", pp 61-83 in (B.1)

(58) Computer Graphics - ACM Siggraph Newsletter, Fall 1969-70. p 9

(59) D. J. Theis, L. C. Hobbs, "Low Cost Remote CRT Terminals", Datamation, June 1968 pp 22-29

(60) "Graphic Data Tablets" The Electronic Engineer, Nov. 1969 pp 50-54

(61) "IDIIOM - The Interactive Display Systems", Data Sheets - Information Display Inc. Mt Kisco, N.Y.

(62) R. J. Houldin, "IBM Graphic Display System", Information Display, Sept/Oct. 1966 pp 16-20

(63) "LSD-1 - Line Drawing System - 1": Data Sheets, Evans and Sutherland Computer Corp., Salt Lake City, Utah.

(64) J. E. Bryden , "Design Consideration for Computer Driven CRT Displays", Computer Design, March 1969 pp 38-46

(65) R. L. Kuehn and H. R. Luxenberg, "Visual Experience and Colorimetry", pp 102-136 in (B.2)

(66) R. L. Kuehn, "Image Analysis", pp 137-164 in (B.2)

(67) "Model 1500A Artwork Generator", Data Sheets - SEACO Computer Display Inc. Garland, Texas.

(68) pp 294-307 in (B.3)

(69) O. Goncarovs , "Creation of contrast on a Cathode Ray Tube", Information Display Feb. 1970 pp 29-31

(70) G. R. Kaelin , "Flat-Panel Displays", paper presented at the 11th National Symposim on Info. Dis. plan Soc. Info Display 1970 (reprint)

(71) R. T. Loewe, "System Design, Coding, Formats, and Programming", pp 24-69 of (B.2)

(72) pp 27-288 of (B.4)

(73) W. H. Tew Jr., "Displaying Engineering Data on a colour CRT", WESCON 1969, IEEE paper 18/2

(74) C. Machover, "Graphic CRT Terminals - characteristis of commercially available equipment ," Proc FJCC, 1967 pp 149-159

(75) pp 94-95 of (B.5)

(76) R. W. Watson , "Design of a General Purpose Graphic Terminal for a Time-sharing System", Tech. Report No 138-68, Shell Dev. Co., Emeryville, California, Jan. 1968.

(77) V. C. Sobolewski, "Using Moire Pattern to Determine the Distortion of Graphic Displays and Graphic input Devices," Proc IEEE Vol 58, 4, April 1970 pp 567-583

(78) W. Huelskoetter, J. Kimlinger, "Ultra-Precision Artwork generated with CRT display", IEEE Intl. Convention

Record 1969 pp 382-383

(79) pp 290-291 of (B.4)

(80) J. E. Wurtz "Dry Process Materials for Hard Copy Printout from CRT's", Information Display, March/April, 1969 pp 45-52

(81) R. Rinder, "Hard Copies from CRT Terminals", Datamation, Vol. 15, 4, April 1969 pp 64-66

(82) H. H. Poole, "Computer Display System Tradeoffs", pp 41-59 in (B.1)

(83) pp 73-75 in (B.4)

(84) pp 74-85 in (B.5)

(85) M. H. Lewin, "An Introduction to Computer Graphics Terminals", Proc. IEEE, Vol. 55, 9, Sept. 1967 pp 1544-1552

(86) M. Macauley, "Low Cost Terminals Using Television Techniques - A Design Study", Proc. IREE (Aust), Vol. 29, 9, Sept. 1968 pp 307-312

(87) J. Standeven et al., "Multiconsole-computer display system utilizing television techniques", Proc. IEE, Vol. 115, 10, Oct. 1968 pp 1375-1379

(88) G. A. Rose, "Intergraphic" - A Microprogrammed Graphical Interface Computer", IEEE Trans. Elec. Comp., EC-16, 6, Dec. 1967 pp 773-784

(89) D. R. Evans, "Computer Graphics Hardware Techniques", pp 7-15 in (B.1)

(90) P. E. Oberbeck, Hybrid Circuit Design of the Artrix Graphical Processor", Dept of Computing Science Tech. Report No 220, U. of Illinois Jan. 1967.

(91) M. L. Dertouzos and H. L. Graham, "A Parametric Graphical Display Technique for On-line use", Proc. F.J.C.C. Vol. 29, 1966 pp 201-209

(92) R. Elliott Green, "Present-day Computer Graphics Research", pp 207-216 in (B.1)

(93) R. A. Banker, "Techniques of Dynamic Display", Control Engineering, Feb. 1960 pp 100-106

(94) C. Machover, "Converting Data to Human-interpretable form", Data Systems Design, Sept. 1964 (reprint)

(95) R. W. Asher, H. Martin, "Cathode Ray Devices," pp 237-276 of (B.2)

(96) R. E. Mahan, "Data Display - System Characteristics", Instruments and Control Systems, March 1970 pp 81-84

(97) W. Bial, "A Computer Controlled CRT Display System", Electronic Engineering, June 1966 pp 354-359

(98) M. L. Dertouzos, "Character Generation from resistive storage of time derivates", Proc. F J C C, 1969, Vol. 25, pp 561-568

(99) E. N. Chase, "Computer-Driven Line-Drawing Displays - A User's View", Data Processing Magazine, Jan. 1968 pp 24-27

(100) F. W. Mumma, "Progress with Displays", Data Processing Magazine, Oct 1969. pp 34-42

(101) D. B. Parker, "Graphical Communications in On-Line System", pp 89-100 in (B.6)

(102) S. Bird, "Computer Graphics Software Techniques", pp 17-27 of (B.1)

(103) chapts 17-19 of (B.4)

(104) J. A. Davis, "Recent Advances in CRT Display Devices," in "Recent Advances in Display Media", NASA SP-159 NASA 1968 pp 25-39

(105) E.A. Hilborn, "Status, Trends and Predictions of

Display Devices", WESCON 1969, IEEE paper 18/1

(106) L. C. Hobbs, "Display Applications and Technology", pp 1870-1884 in (36)

(107) J. B. Armstrong et al, "Advanced Display Techniques", Instruments and Control, Dec. 1968 pp 69-71

(108) W. Merel, "The Electroluminescent Cross-Grid Display", Electro-Technology, Jan. 1966 pp.81-82

(109) H. G. Blank, J. A. O'Connell, M. S. Wasserman, "Non-linear Resistors Enhances Display Panel Contrast", in "Optoelectronic Devices and Circuits" ed. S. Webber McGraw Hill, 1964. pp 188-193

(110) C.P. Crocetti et al., "Light valves, Lasers and Electro-luminescent Panels", pp 381-388 of (B.2)

(111) pp 313-320 in (B.3)

(112) D. L. Bitzer, H. G. Slottow, "The Plasma Panel- A Digitally Addressable Display with Inherent Memory", Proc F J C C, Vol. 29, 1966 pp 541-547

(113) D. L. Bitzer, H. G. Slottow, "The Plasma Panel - A new device for direct display of Graphics", pp 3-28 in (B.7)

(114) J. Hanlet, R. W. Haas, "Carrier Injection Luminescence" in "Recent Advances in Display Media", NASA SP-159, NASA 1968, pp41-49

(115) H. Borden, R. Kniss, "Strobing makes longer LED alphanumeric displays practical", Electronics, Vol. 43, 5, March 2, 1970 pp 126-135

(116) H. B. Tilton, "Principles of 3-D CRT Displays", Control Engineering, Feb. 1966 pp 74-78

(117) W. F. Sharpe, "The Economics of Computers", Columbia U. Press, New York, 1964 pp 363-441

(118) "T-4002 Graphic Computer Terminal"-Data Sheets, Tekronix Inc., Beaverton, Or. 1970.

(119) D. N. Keast, "Survey of Graphic input Devices", Machine Design, August 3, 1967 pp 114-120

(120) E. M. Thomas, "GRASP - a graphic Service system", Proc. ACM Natl. Meet., 1967 pp 395-402

(121) C. M. Cundall, "Cathode Ray-Tube Display System", Electronics and Power, March 1968 pp 115-120

(122) R. L. Fulton, " Visual Input to Computers", Datamation, Vol. 8, 8, August 1963 pp 27-40

(123) G. A. Rose, ""Light Pen" Facilities for Direct View Storage Tubes - An Economical Solution for Multiple Man-Machine Communications", Trans. IEEE Elect. Comp, EC-15, August 1965 pp 637-639

(124) L. Gallenson, "A graphic tablet display console for use under time-sharing", Proc. FJCC, Vol. 31 1967, pp 689-695

(125) M. R. Davis, T.O. Ellis, "The Rand Tablet - A Man-Machine Graphical Communication Device", Proc. FJCC Vol. 26, 1964 pp 325-331

(126) J. F. Teixerira, R.P. Sallen, "The Sylvania data Tablet: A new approach to graphic data input", Proc. SJCC, Vol. 32, 1968 pp 315-321

(127) "Graf-pen", Data sheets, SAC, Southport Conn.

(128) "Ecricon", Data sheets, Shintron Co. Inc. Cambridge Mass.

(129) pp 65-70 of (B.5)

(130) B. M. Gurley, C.E. Woodward, "Light-pen links Computer to Operator", Electronics, Vol. 32, 47, Nov. 20, 1959 pp 85-87

(131) ""Photopen" Light Sensing System",- Data Sheets TC-198, Sanders Associates Inc., Nashua, New Hamp.
(132) L. G. Roberts, "The Lincoln Wand", Proc. FJCC, Vol. 29, 1966 pp 223-227
(133) A. M. Jessup, C. S. Wallace, "A cheap Graphic Input Device", Aust. Comp Jnl., Vol. 1, 2, May 1968 pp 95-96
(134) pp 190-201 in (B.3)
(135) J. S. Weiss, "Inexpensive Electronic Memories", paper presented at 11th National Symposium on Info. Display, Soc. Info. Display, 1970 (reprint).
(136) R. H. Norman, "Memories I: Staying ahead of the game", Vol. 41, 22, Electronics Oct. 28, 1968 pp 106-108
(137) D. W. Moore, "Memories II: scant room for improvement", ibid. pp 109-112
(138) pp 53-54 of (B.5)
(139) R. A. Tracy, "Delay Lines - key to low cost in keyboard machines", Electronics, Vol. 42, 12, June 9, 1969 pp 114-117
(140) D.C. Vimari, "A Choice of serial memories", Electronics, Vol. 42, 21, Sept. 15, 1969 pp 122-125
(141) pp 202-205 in (B.3)
(142) R. H. Anderson, "A Simplified Direct-Viewing Bistable Storage Tube", IEEE Trans. on Elec. Dev., ED-14, 12, Dec. 1967 pp838-844
(143) C. N. Winningstad "The Simplified Direct Viewing Bistable Storage Tube in Computer-Output Applications", 8th Natl. Info. Display Symp, May 1967, pp 129-136
(144) " ARDS-100A",- Data Sheets, Computer Displays Inc., Waltham, Mass.

(145) R. H. Stotz, " A new Display Terminal", Computer Design, April 1968 pp 80-86

(146) M.A. Ruben "Low Cost Graphics", pp 91-97 in (B1)

(147) pp 243-247 of (B.3)

(148) M.D. Marsh, "Where they are Today ... Display and Storage Tubes", Electronic Industries, April 1966 pp 54-61

(149) J. W. Esch et al., "ARTRIX Final Report", Dept. of Computing Science No 238, U. of Illinois, June 1967.

(150) J. E. Esch, "System Design for Artrix Graphical Processor", Dept of Computing Science, No 219, Jan. 1967

(151) A. D. Berg, R. W. Smith, "A possible Electron Image Information Store", Proc. 9th Meeting AGARD Avionics Panel, Paris, Sept. 1965, NATO. pp 105-112

(152) A.E. Huston, J. Balderstone, D.A. Proctor, "An image Storage Tube", ibid. pp 113-126

(153) W. L. Erickson, T. M. Soller, "Computer-Driven Displays", 1965 Intl. Conv. Rec., Pt. 3, pp 72-84

(154) S. Davis, "Display Processing Subsystems for Computer Data Displays", Computer Design, May 1969 pp 50-55

(155) S. Sherr, "Generalized Equations for Display System Design", 9th Int. Info. Display Symposium, 1968 pp 148-157

(156) W. H. Ninke, "Graphic -1: A Remote Graphical Display Console System", Proc. FJCC, Vol. 27, 1965 pp 839-846

(157) C. Christensen, E.N. Pinson "Multi-function Graphics for a large Computer System", Proc. FJCC, Vol. 31, 1967 pp 697-711

(158) "Computer Graphics : Vol. 1 NASA-MSFC Digital Computer Graphics Program", CFSCTI N68-16206 Dec. 1967.

(159) B. Davis - Private communication

(160) D. J. Theis, L. C. Hobbs, "Mini-computers for Real-time applications", Datamation, Vol.15, 3, March 1969 pp 39-61

(161) F. J. Corbato, V.A. Vyssotsky, "Introduction and Overview of the Multics System", Proc. FJCC, Vol. 27, 1965 pp 185-196

(162) A. R. Rundle, "Interactive Software Techniques", pp 29-39 in (B.1)

(163) W. F. Miller, A.C. Shaw, "A Picture Calculus", pp 29-39 in (B.7)

(164) H. E. Kulsrud, "A General Purpose Graphics Language", Comm. ACM, Vol 11, 4, April 1968 pp 247-254

(165) J. N. Hamburg, J. Ironside, G. C. Barney, "An economical Display System", Comp. Bull., Sept. 1969 pp 314-322

(166) N. E. Wiseman, "A note on compiling display file from a data structure", Comp. J., Vol. 11, Aug. 1968 pp 141-147

(167) pp 72-74 in (B.5)

(168) S. Pardee, P. E. Rosenfeld, "A Remote Time Share Terminal with Graphic Output Capabilities", Bell Telephone Labs, N.J. (reprint).

(169) J. H. Morrissey, "Quiktran System", pp 116-126 in (B.6)

(170) "On-line Circuit Design on Commercial Time-shared Systems", pp 314-322 in (B.4)

(171) N. E. Wiseman, H. U. Lemke, J. O. Hiles, "Pixie: A New Approach to Graphical Man-Machine Communication", Int 1. Conf on Comp aided Design, Southhampton, 1969, IEE, London pp 463-471

(172) R. B. Talmadge, "Structuring Compilers for On-line Systems", pp 105-115 in (B.6)

(173) I B M Systems Journal, Nos. 3/4, 1968.

(174) M. H. Lewin, "A Magnetic Device for Computer Graphic Input", Proc. FJCC, Vol. 27, 1965 pp 831-838

(175) Computers and Automation, Dec. 1969 p 65

(176) A. M. Hlady, "A touch Sensitive X-Y Position Encoder for Computer Input", Proc FJCC, Vol. 35, 1969, paper 119.

(177) P. W. Woo, "A proposal for Input of Hand Drawn Information to a Digital System", IEEE Trans. Elec. Comp., EC-14, Oct 1964 pp 609-611

(178) E. A. Johnson, "Touch Display - A Novel Input/Output Device for Computers", Electronics Letters, Vol. 1, 8, Oct. 1965 pp 219-220

(179) J. T. Locascio et al., "Obtaining Light Pen Versatility", Information Display, Nov/Dec. 1967 pp 36-39

(180) D. R. Haring, "The Beam Pen - A Novel High Speed Input/Output device for CRT Display Systems", Proc. FJCC, Vol. 27, 1965 pp 847-855

REFERENCES DEALING PRIMARILY WITH COMPUTER-AIDED CIRCUIT DESIGN

(181) G. W. Zobrist, "Thinking of getting into CAD?" Electronics, March 30, 1970 pp 98-108

(182) "Electronics guide to CAD Programs", Electronics, April 13, 1970 (Supplement Foldout).

(183) C. Pottle, "The Past Development, Present Problems, and future Potential of Computer-Aided Network Design", Proc. NEC, 1969 pp.716-719

(184) "Panel discussion on Computer-Aided Design", Proc. IEEE, Vol. 55, 11, Nov. 1967 pp 1775-1786

(185) C. Pottle, "Comprehensive Active Network Analysis by digital Computer - A State Space Approach", Proc. 3rd Annual Allerton Conf. on Circuit and System Theory, 1965 pp 659-688

(186) "Automated Circuit Analysis Course No. 881", April 3 - 7, 1967. Lecture Notes U. of California, Berkeley, Calif.

(187) W.W. Happ, "NASAP: Present Capabilities of a Maintained Program", IEEE Intl. Conv. Record, N.York, 1967 Pt. 5. pp 64-88

(188) R. A. Sibley, "A new programming language PL/1", Proc. 20th ACM Natl. Conf, 1965 pp 543-563

(189) A. F. Malmberg, F. Cornwell, F. Hofer, "NET-1 Network Analysis Program - 7090/94 Version", Rept - LA - 3119, 1964, Los Alamos Scientific Labs, Los Alamos. N. Mexico.

(190) D. Christiansen, "Computer-aided design: part 6 - Comparing the "Big Two" programs", Electronics, Feb. 6, 1967 pp 74-90

(191) F.H. Branin, "Machine Analysis of Networks and its Application", IBM Rept. TR00855, March 1962, IBM, Poughkeepsie, N. York.

(192) F. H. Branin "A New Method for Steady-State A C Analysis of RLG Networks", IBM Rept TR00 1391, Jan. 1966, IBM Poughkeepsie, N. York.

(193) "Electronic Circuit Analysis - User's Manual", IBM File No. 1620-EE-02X, 1965, IBM, White Plains, N. York.

(194) R. Carpenter, W. Happ, "Computer-aided design: Part 3 Analysing circuits with symbols", Electonics Vol. 39, 25, Dec. 12, 1966. pp 92-98

(195) S. R. Sedore, "SCEPRE gives designers new Power", Electronic Design, March 15, 1967 pp 230-233

(196) H. W. Mathers, "SCEPTRE - (System for Circuit Evaluation and Prediction of Transient Radiation Effects)", IBM Rept. No: 67-825-2122, May 1967, IBM Electronics System Centre, Onego, N.Y. .

(197) D. A. Calahan, N. E. Abbott, "Stability Analysis of Numerical Integration", Conf. Record 10th Midwest Symp. Circt. Theory, Lafayette, 1967 (IEEE), paper I-2.

(198) C. Pottle, "Comprehensive Active Network Analysis by Digital Computer - A State Space Approach", Res. Rept EERL 59, Sept 1966, Cornell U.

(199) "CIRCUS - A Powerful Aid to the Circuit Designer", Boeing Co.,1967, Seattle, Wash.

(200) "CIRC - A Computer program for circuit design engineering", - SD8 Application Bulletin 64-79-07B SOS, Santa Monica, Calif., Aug 1967.

(201) E. Wolfendale, "Computer aids for electronic equipment design", - reprint from "Computer Aided Design", Autumn 1968.

(202) "Redac - Summary of Programmes", Racal Research Ltd., Tewkesbury, Gloustershire.

(203) W. L. Morris, R. Ranalli, "Today's Computer Aided Design Programs and Application Capability", Proc. NEC, 1969 pp 720-725

(204) D. F. Dawson et al., "Computer-Aided Design of Electronic Circuits - A User's View Point", pp 1946-1954 in (37)

GENERAL REFERENCES DEALING WITH MAN-COMPUTER GRAPHICS.

(205) J.C.R. Licklider, W. E. Clark, "On-Line Man-Computer Communications", Proc. SJCC, Vol. 21, 1962 pp 113-126

(206) S. L. Smith, "Man-computer Information Transfer", Electro Technology, Aug. 1963 pp 83-87

(207) D. Holstein, "Are you ready for Computer-aided designs", Product Engineering, Nov. 23, 1964 pp 66-76

(208) "Man-Machine Interaction: Engineering Design - Symposium Summary", Proc IFIP, Congress '65 (Spartan Books) pp 476-478

(209) "Man-Machine Interaction: Remote Consoles and Displays - Symposium Summary", ibid. pp 504-509

(210) "Man-Machine Interation: Graphic Data Processing Symposium Summary", ibid. pp 578-583

(211) I. Abzug, "Graphic Data Processing", Datamation, Vol. 11, 1, Jan. 1965 pp 35-37

(212) D. E. Rippy, D. E. Humphries, "MAGIC - A Machine for Automatic Graphics Interface to a Computer", Proc FJCC, Vol. 27, 1965 pp 819-830

(213) A. H. Vorhaus, "General Purpose Display System", Datamation, Vol. 12, 6, July 1966 pp 59-63

(214) M. D. Prince, "Man-computer Graphics for Computer-Aided Design", Proc. IEEE, Vol. 54, 12, Dec. 1966 pp 1968-1708

(215) J. H. Argyris, P. C. Patton, "A look into the Future - How computers will influence Engineering", J. Roy. Aero. Soc., Vol. 71, 4, April 1967 pp 244-270

(216) G. D. Hornbuckle, "The Computer Graphics User/Machine Interface", IEEE Trans Hum. Fact, HFE-8, 1, March 1967 pp 17-20

(217) "Picture looks brighter for Computer Displays", Electronics, Vol. 41, 22, Oct. 16, 1967 pp165-171

(218) R. A. Siders, "Computer-aided design", IEEE Spectrum, Nov. 1967 pp 84-92

(219) R. H. Stotz, "Directions in Time-sharing Terminals", IEEE Computer Group News, May 1968 pp 12-18

(220) G.A. Rose, "Graphical Terminals and the Computer Utility", Proc IREE (Aust), Vol. 29, 11, Nov. 1968 pp 386-389

(221) W. L. Frank, "Software for Terminal-oriented Systems", Datamation Vol 14, 6, June 1968 pp 30 -34

(222) H. J. Genthner, "Interactive Computer Graphics", Computers and Automation, Nov. 1968 pp 14-17

(223) Workshop, "Impact of Terminals", IEEE Computer Group News, March 1969 pp 23-28

(224) J. Lafuente, "Add graphic display to your computer", Electronic Design, April 1, 1969 pp 64-67

REFERENCES NOT PREVIOUSLY REFERRED TO WITH MAN-COMPUTER GRAPHIC APPLICATIONS

(225) R. W. Mann, "The 'CAD' Project", Mechanical Engineering, May 1965 pp 41-43

(226) "Computer Graphics", Data and Control Systems, Feb. 1967 pp 32-33

(227) J. H. Kanner, "CAI - The New Demonology", Datamation, Vol. 14, 9, Sept. 1968 pp 38-40

(228) K. C. Knowlton, "Computer-Animated Movies", pp 343-370 in (B.7)

(229) H. S. Rasmussen, "Graphic-aided Design of a Graphic Device", 9th Natl. Symp. Info Display, pp. 61-68

(230) B. Herzog, "Computer Graphics for Designers", pp 189-230 of (B.7)

(231) R. Khol, "Computer Stress Analysis", Machine Design, Nov. 21, 1968 pp 136-145

(232) "Advanced NC goes to work", Machine Design, Dec. 19, 1968 pp 130-132

(233) "Computer-aided design at ICL", Data Processing, Sept.-Oct., 1969 pp 489-495

(234) J. R. Logan, "Computer Graphics in Electronic Circuit Design", Computers and Automation, Nov. 1968 pp 28-29

(235) C. J. Fisk, D. L. Caskey, L. E. West, "ACCEL: Automated Circuit Card Etching Layout", pp 1971-1982 in (37)

(236) A. Spitalny, M. J. Goldberg, "On-line Graphics Applied to Layout Design of Integrated Circuits", pp 1982-1988 in (37)

(237) J. S. Koford et al., "Using a Graphic Data processing System to Design Artwork for Manufacturing Hybrid Integrated Circuits", Proc. FJCC, Vol. 29, 1966 pp229-246

(238) J. C. Simek, C. J. Tunis, "Handprinting Input Device for Computer Systems", IEEE Spectrum, July 1967 pp 72-81

(239) A. Rosenfeld, "Picture Processing by Computer", Computing Surveys, Vol. 1, 3, Sept. 1969 pp 146-174

(240) A. M. Noll, "The digital computer as a creative medium", IEEE Spectrum, Oct. 1967 pp 89-95

(241) R. W. Conn, "Digitized Photographs for illustrated computer output", Proc. SJCC, Vol. 30, 1967 pp 103-106

(242) "Computers and Automation" Annual Computer Art Contest August issues, 1966, 1967, 1968, 1969.

(243) I.G. MALOFF, D.W. EPSTEIN, "Reflective Optics in projection T.V.", pp 250-269 in "TELEVISION Vol. 4", RCA Labs Princeton N.J. 1947.

(244) P.C. JANSEN, "Television Optics", Electronics Application Bulletin, pp 61-76, pp 95-103, pp 161-179 (parts 1-3 Vol. 11, 1950.

(245) H. RINIA, H.J. DEGIER, P.M. van ALPHEN, "Home Projection Television Pt. 1", Proc. IRE, Vol. 36, 3, March 1948.
pp 395-400

(246) H.W. LEVERENZ, "An Introduction to Luminescence of Solids", John Wiley New York 1950. pp 357-359

(247) Philips Data Handbook - Pt 5 "Screen Phosphors and Industrial CRT's" pp 1-71, Phillips EC & M Division May 1970.

(248) G.I. COHN, H.M. MUSAL, "Transient Response of Phosphors", Trans. Comp. Parts, CP-5, 2, June 1958.
pp 90-101

(249) D.G. FINK ed. "T.V. Engineering Handbook", McGraw-Hill 1957.
p 5-16

(250) P. VLAHOS, "Film-Based Projection Systems", pp 297 in chapt. 9 in "Display Systems Engineering" ed. H.R. Luxenburg and R.L. Kuehm: McGraw-Hill, New York 1968.

(251) S.V. FORGUE, "The Storage Orthicon and its Application to Teleran", RCA Review, Vol. 7, 4, December 1946.
pp 633-650

(252) R.W. REDINGTON, "The Transient Response of Photoconductive Camera Tubes Employing Lon Velocity Scanning", Trans Electron Devices, ED-5, 7, July 1957. pp 250-225

(253) L.J. Van der POLDER, "Target-Stabilization Effects in TV Pickup Tubes", Phillips Res. Repts, Vol. 22, 1967.
pp 178-207

(254) L.J. Van der POLDER, "Beam Discharge Lag in Television Pick-up Tube", in "Advances in Electronics", Vol. 28A, 1969 Academic Press, London. pp 237-245

(255) L. HEIJNE, "Photoconductive properties of Lead Oxide Layers", Phillips Res. Rept. Supp. 4, 1961. chapt. 8 pp 140-145

(256) R.C. NEUHAUSER, L.D. MILLER, "Beam Landing Errors and Signal Output Uniformity of Vidicons", J. Svc. Mot. Pict. TV Eng., Vol 67, March 1958. pp 149-153

(257) J. CASTLEBERRY, R.H. VINE, "An Improved Vidicon Fowsing Deflecting Unit", JSMPTE, Vol. 68, April 1959. pp 226-229

(258) H.G. LUBSZYNSKI, J. WARDLEY, "Some Problems of Resolution in Low-Velocity Camera Tubes", Proc. IEE, Vol. 110, 6, June 1963. pp 59-63

(259) RCA Data Sheets on Projection Kinescopes, RCA Tube Division, RCA, Harrison, N.J.

(260) Philips Data Handbook Pt. 5 "VIDICON SPECIFICATION-Spurious signal Specification for Vidicon Tubes", Philips EC. and M Division, May 1970.

(261) R. CLARK-JONES, "On the Detective Quantum Efficiency of T.V. Camera Tubes", JSMPTE, Vol. 68, July 1959. pp 462-466

(262) P.K. WEIMER, S.V. FORGUE, R.R. GOODRICH, "The Vidicon Photoconductive Camera Tube", RCA Review, Vol. 12, 1951. pp 306-313

(263) P. MERTZ, "Perception of T.V. Random Noise", JSMPTE, Vol. 54, Jan 1950. pp 8-34

(264) Y. KIUCHI, "The Persistant Lag in the Vidicon", TOSHIBA REVIEW, Vol. 15, 1958. pp 920-926
(abstract only - paper in Japanese yet!)

(265) MACHOL R.E. ed, "System Engineering Handbook", McGraw-Hill 1968. pp 128-129

(266) "Handbook of Mathmatical Functions", ed. M. Abramovitz and I. Stegun: Dover 1965. p 972

(267) R.R. LAW ,"Contrast in Kinescopes", in "Television Vol. III (1938-1941)": RCA Labs. 1944. pp 294-324

(268) S. DEUTSCH, "Theory and Design of T.V. Receivers", McGraw-Hill, New York 1951. pp 495-496

(269) Philips Data Handbook - Part 5 ,"Data Sheets on MW13-38", 5 pps. Phillips E C & M Division, May 1967.

(270) E.M. NOLL, "TV for Radiomen", McMillan Co,New York 1949. chapt 10. pp 312-352

(271) H. G. BOYLE, E.B. DOLL, "Compact Projection T.V. System", Electronics,Vol. 30, April 1948. pp 72-77

(272) W. E. BRADLEY, E. TRAUB, "A New T.V. Projection System", Electronics, Vol. 29, Sept 1947. pp 84-89

(273) Test Report - "Phillips Projection System" , Wireless World , Vol. 56, 10, Oct 1950. pp 365-367

(274) P. VLAHOS, "Selection and Specifications of Rear Projection Screens", JSMPTE , Vol. 70, 2, Feb 1961. pp 89-95

(275) E. W. D'ARCY, G. LESSMAN, "Objective Evaluation of Projection Screens", JSMPTE , Vol. 61, 6, Dec 1953. pp 702-720

(276) A.J. HILL , "A First-Order Theory of Diffuse Reflecting and Transmitting Surfaces", JSMPTE, Vol. 61, 1, 1953. pp 19-23

(277) R.R. LAW, I.G. MALOFF, "Projection Screens for Home Television Receivers", Jour. Opt Soc. Am,Vol. 38, 6, June 1948. pp 497-502

(278) J. MORGAN ,"Introduction to GEOMETRICAL and PHYSICAL OPTICS", McGraw Hill, New York, 1953. pp 36

(279) "Applied Optics and Optical Engineering", ed R. KINGSLAKE Vol. I : "Light Sources for Optical Devices", CARLSON F.E, CLARK C. N., chapt 2. Academic Press, New York 1965. pp 43-109

(280) EDMUND Catalog, Number 681, Edmund Scientific Co, New Jersey.

(281) P, ROLLS , "Photographic Optics" in "Photography for the Scientist", ed. C.E. Engel, Academic Press, London 1968. pp 69-172

(282) G.V. NOVOTNY , "FIBER OPTICS for Electronic Engineers", in "Opto-Electronic devices and Circuits", McGraw Hill 1964.

(283) N.S. KAPANY , "FIBER OPTICS - Principles and Applications", Academic Press, New York 1967.

(284) D.F. CAPELLARO, "Optical Systems using Fiber Optics" in "LIGHT and HEAT SENSING", 6th. AGARD AVIONICS PANEL MEETING, PARIS, July 1962. pp 311-324

(285) W.P. SIEGMUND ,"Fiber Optics - Principles Properties and Design Considerations," ibid. pp 265-296

(286) "Designing with CROFON Light Guides", E.I. Du PONT de NEMOURS and Co Inc. Wilmington , Delaware.

(287) Data Sheets on "Standard Fiber Optic Light Guides", Electro-FiberOptics Corp., Worcester, Mass.

(288) "IRE Standards on Television: Method of Measurement of Aspect Ratio and Geometric Distortion, 1954", Proc. IRE, Vol. 42, 6, July 1954. pp 1098-1103

(289) "SOFT-COPY CONFERENCE", Advance Notice from National Bureau of Standards, Washington , Aug 1970.

(290) J.R. HEFELE, "A precision Solid State TV Picture Monitor", Hewlett-Packard Inc. Feb 1968. p 2-8

(291) G. OSTER, M. WASSERMAN, G. ZWERLING, "Theoretical interpretation of moire patterns", J. Opt Soc. Am, Vol. 54, 2, February 1964. pp 169-175

(292) A.J. HENK, "Scan Linearity measurements without tears", Wireless World, Vol. 72,9, September 1966.
pp 449-454

(293) Letraset Catalogue 1968, London, Letraset Ltd.

(294) V.K. ZWORYRIN et al. "Electron Optics and the Electron Microscope", John Wiley, New York 1945, chapts 1. 14-19.

(295) R.K. LUNEBERG, "Mathematical Theory of Optics": Supplementary Notes on Electron Optics, U. of Calif. Press, Berkeley 1964. pp 373-410

(296) W. GLASER, "Grundlagen der Electroneoptik", Springer Verlag, Vienna 1952. Chapts. 16, 22. ("Foundations of Electron Optics").

(297) B. PASZKOWSKI ,"Electron Optics", Iliffe Books, London 1968, chapt. 9 (particularly p. 281).

(298) H. MOSS, "Narrow Angle Electron Guns and Cathode Ray Tubes", Academic Press, New York 1968.
pp 111-116

(299) H.G. LUBSZYNSKI, et al. ,"New All-Electrostatic Vidicon", Proc IEE, Vol. 116, 3, March 1969. pp 339-347

(300) J. WARDLEY, F.W. JACKSON ,"A 13 mm. All Ectrostatic Vidicon" in "Advances in Electronics", Vol. 28A, 1969.
pp 247-251

(301) Chapter 7 of (297) above. pp 238-263

(302) "Deflection Yokes-Focus Coils" - Data Sheets and Catalogue - Syntronic Instruments Inc. Addison, Ill., 1968.

(303) J.L. DIVILBISS, "Illiac III Scanner Analog Circuits", Rept. 320, April 1969, Dept. Comp. Science, U. of Illinois, Urbana.

(304) D. FINK ed. "T.V. Engineering Handbook", McGraw Hill, New York 1958. p 6-3

(305) pp 478 - 479 in (294).

(306) V. JARES, B. NOVOTNY ,"Two Methods for the Determination of the Imaging Properties of Electron-optical Systems with a Photocathode", in "Advances in Electronics" Vol. 28A, 1969, Academic Press. pp 523-533

(307) T. SOLLER, M. STARR, G.F. VALLEY eds.,"Cathode Ray Tube Displays", Vol. 22, Radiation Lab. Series, McGraw Hill, New York 1948. pp 305-308, 338-40

(308) B.F. WADSWORTH, "Deflection Non-linearities and Distortions in PEPR's CRT Display", PEPR Eng. Note 19, April 1963, Lincoln Labs., M.I.T.

(309) P.K. WEIMER, A. ROSE, "The Motion of Electrons Subject to Forces Transverse to a Uniform Magnetic Field", Proc. IRE, Vol. 35, 11, Nov 1947. pp 1273-1279

(310) S. FRANCO ,"Dynamic Focussing and Pincushion Correction Circuits for the Illiac III Display System", File 805 Aug. 1969. Dept. Comp Science, U of Illinois, Urbana.

(311) Celco Data Sheets -(Available Pincushion and Defocussing Correction circuits): CELCO , Upland, Calif.

(312) G.M. GLASFORD, "Fundamentals of T.V. Engineering", McGraw Hill, New York 1955. pp 39-41

(313) O.H. SCHADE, "Electro Optical Characteristics of T.V. Systems - Part II", RCA Rev, Vol. 9, 2, June 1948.

(314) D. FINK ed.,"TV Engineering Handbook", McGraw Hill, New York 1958. chapts. 11, 15, 16.

(315) R. F. SHEA, "Amplifier Handbook", McGraw Hill, New York 1966. chapt. 25.

(316) Linear Integrated Circuit D.A.T.A. Book, Spring 1970, D.A.T.A. Orange, N.J.

(317) "I/C OP-AMP SELECTION CHARTS", The Electronic Engineer, Vol. 29, 7, July 1970. pp 57-61

(318) Digital Integrated Circuit D.A.T.A. Book, Spring 1970, D.A.T.A., Orange, N.J.

(319) "TRANSISTOR/TRANSISTOR MICROLOGIC INTERGRATED CIRCUITS", Appl. Bull APP-131: Data Sheets, Fairchild SC, Mountain View, Calif. 1967.

(320) G.A. ROBINSON, "Operation of Vidicons in Unusual Environmental Conditions", J. SMPTE, Vol. 70, 4, April 1961. pp 282-83

(321) "TH Series Displays": Data Sheets, Ball Brothers Research Corp., Boulder, Colorado, March 1969.

(322) R.E. Johnson, "The Application of Vidicons in Modern T.V. Systems", Info. Display, Vol. 7, 3, March 1970. pp 29-35

(323) J.E. WORTZ "High Resolution Cathode Ray Tubes" Info. Display, Vol. 4, 3, May/June 1967. pp 39-45

(324) R.A. BELL, "CRT Phosphor Selection", Instruments and Control Systems, March 1970. pp 86-90

(325) "Monitor Type 2788CG - Inspection and Maintenance Manual", Pye Ltd, Cambridge, England: Nov 1956.

(326) "Sony T.V. Camera Model CVC-100 - Technical Manual" Sony Corp., Tokyo, Japan.

(327) R. W. Biddlecomb, "Latest Multi-Vibrator Improvement", Electronics, Vol. 36, 4, April 1963. pp 64-65

(328) J. MILLMAN, H. TAUB, "Pulse, Digital and Switching Networks", McGraw Hill, New York 1965. chapter 11

(329) "MC 1520 - Op Amp": Data Sheets, Motorola Inc.

(330) J.K. Clapp, "An Inductance-Capacitance Oscillator of Unusual Frequency Stability", Proc. IRE, Vol. 36, 3, March 1948 pp356-358

(331) Proc. Photoconductivity Conference, Atlantic City, Nov 1954. John Wiley, New York 1956. p V11

(332) S.P. ZELENOBORSKIY, "Storage in Camera Tubes with Photoconductive Targets", Radio Engineering and Electronic Physics, 7, July 1962. pp 1110-1119

(333) L. HEIJNE "Photoconductive Properties of Lead Oxide Layers", Phillips Res. Rept. Supp, 4, 1961 chapt. 8

(334) R.H. BUBE, "Phosphors and Photoconductors", Proc NEC, Vol. 12, 1956. pp 733-739

(335) R.H. BUBE ,"Photoconductivity", RCA Eng. Vol. 4, Dec 1959/Jan 1960. pp 28-32

(336) A. ROSE ,"An Outline of some Photoconductive Processes", RCA Review, Vol. 9, Sept 1951. pp 362-414

(337) A. ROSE, M.A. LAMPERT , "Gain-Bandwidth Product for Photoconductors", RCA Review, Vol. 3, March 1958. pp 57-68

(338) P. K. WEIMER ,"Television Camera Tubes - A Research Review", in "Advances in Electronics", Vol ,19 Academic Press, New York. pp 3 -437

(339) M.A. LAMPERT ,"Simplified Theory of Space Change Limited Currents in an Insulator with Traps", Phys. Rev., Vol. 103, 6, Sept 15, 1956. pp 1648-1656

(340) A. ROSE ,"Space Change Limited Currents in Solid", Phys. Rev., Vol. 97, 6, March 1955. pp 1538-1544

(341) R.W. REDINGTON , "Maximum Performance of High Resistivity Photoconductors", Jnl. Appl. Phys., Vol. 29, 2, Feb 1958, pp 189-193

(342) A. ROSE, M.A. LAMPERT ,"Photoconductor Performance, Space-Change Currents and the Steady-State Fermi Level", Phys. Rev., Vol. 113, 5 March , 1959. pp 1227-12235

(343) NEZU K., NAITO S. ,"Economical Display Generation of a Large Character Set", Proc FJCC, Vol. 35, 1969. pp 569-573

(344) Catalog and Specification Sheets - CINTRA Physics International Mountain View, Calif., 1969 .

(345) "Handbook and Catalog of Operational Amplifiers", Burr-Brown Research Corp., Tex., 1969 .

pp 33-34

(346) "Ferrox-cube Transformer Pot Cores" - Data Sheets in MULLARD TECHNICAL HANDBOOK, Vol 6, "MAGNETIC COMPONENTS", Mullard Ltd., London, UK, 1966 .

(347) "μA 723 Precision Voltage Regulator", Application Notes, Fairchild Semiconductors, Mountain View, Calif., 1968.

. . . . . .

Sobolewski, V. C. (1970). Using moire patterns to determine the distortion of graphic displays and graphic input devices. *Proceedings of the IEEE, 58*(4), 567-583.

https://doi.org/10.1109/PROC.1970.7695