Mildly Distorting Electronic Systems: Measurement, Design and Simulation

Jonathan B Scott

University of Sydney

To J Godfrey Lucas, David J Skellern, and Richard H Small

three wise men.

Preface

This work was half done before I could articulate the theme common to the diverse range of research it encompasses. This preface is where I identify the theme linking these otherwise-disconnected projects and, by the way, indulge in describing a personal view of present-day analog electronics.

In electronics, the vast majority of analyses are "small-signal", which is to say that a linear approximation is made, chiefly to simplify the work. These approximate methods give useful results—formulas for gain, dimensions of speaker boxes, component values to sustain oscillation, fractions for stable feedback loops—and have been doing so for decades, before computers, before precision instruments, in a romantic time for engineers, when instinct and experience was applied by an operator to making something work. Engineers seemed to resemble clock-makers, and knew charming quirks of the equipment they used empirically, if not theoretically.[1]

Way beyond "small signal", there is "large signal". Where something seriously overloads, a second way of looking at electronics is used. The exact definition of what happens is hard to summarise, and engineers use jargon words such as "clipping", "gain compression" or "slewing". Mathematically, things approach an asymptotic limit. A transistor reaches saturation and a node voltage approaches a limit, another transistor cuts off and the voltage on a capacitor slews at a constant rate, a magnetic core saturates and field virtually

[1] This charming view of the profession is held elsewhere, and has been put to good pedagogical use, for instance in "Stimulating by 'Faction' ", by R. A. Smith, in the proceedings of the *World Conference on Engineering Education*, Southhampton, 1993.

stops increasing, or a loudspeaker suspension moves beyond the limit of Hooke's law and cone motion falters.

In between is the region where the subtle distortions, insignificant in the small-signal case, arise. Sound systems distort, generators produce spurious components, and carriers cross-modulate in wideband amplifiers. The jargon has not settled, and this situation is termed "low-signal", "mildly-distorting", "weakly nonlinear" or other things. I prefer to use the adjective low-signal, because it positions the idea between large- and small-signal, and to distance the idea from the term weakly-nonlinear, with its existing and misleading connotations.

Analogue electronics is no longer the domain of "tinkerers" as hacker-engineers were called. The Philips Electronic Engineering kit and the radio ham have given way to the PC and the computer nerd. Small-signal and limiting approximations are no longer sufficient, just as slide rules and analog voltmeters are no longer good enough. These two situations are linked and this thesis is about one consequence of the link: The new precision of electronics engineering and a level of analysis—low-signal—that is both newly-possible and newly-demanded.

Statement of Originality

This work was carried out between 1990 and 1996, during my enrolment as a PhD student, in the Department of Electrical Engineering at the University of Sydney, and in the Department of Electronics, in the School of Mathematics, Physics, Computing and Electronics at Macquarie University. It was supervised by Anthony Parker and Mark Johnson.

Unless otherwise stated all work is my own. The DFT package, and in particular the algorithms for numerical integration and automatic interpolation and windowing decisions are original. The behavioural nonlinear model of a loudspeaker and the associated set of nonlinear driver parameters are original ideas. The idea of measuring the non-linearity of a loudspeaker driver by means of extracting small-signal (Thiele-Small) parameters for various displacements by means of applied dc is original. Some industrial assistance was provided by Elecoustics (Glenn Leembruggen), and some measured data was taken from work by Jonathan Kelly. The nonlinear audio load design is original. I used some measured data from a thesis by Greg Lemon. Some of the design innovations of the TDFD meter are original, in particular the use of booster-follower circuits to reduce residual IMD in the sources, and the use of passive notch prefilters before active filtering. Dirk Heuer, under my supervision, carried out prototype construction and machine level programming as part of his honours thesis. The pulsed-IV device characterisation system (APSPA) and its novel applications were designed in conjunction with Anthony Parker. Exactly who came up with original design features and novel applications is obscured by a marvellous and close co-operation: we apportion credit equally. Some of the APSPA software was written

by James Rathmell. The use of gain-derivative surfaces for visualising nonlinearity in a multidimensional fashion is original, though the name was provided by Danny Webster, a colleague at UCL. The design of the solid-state pseudo-valve amplifier including use of low desensitivity, transformer coupling, unipolar field-effect drive and Webster-node bootstrapping, is original. Mechanical construction was carried out by Adrian Vos.

Contents

1 Introduction **1**

1.1 Chapter Synopses 2

2 Conventional Theory and Background Material **5**

2.1 Definition of Terms 6

2.2 Volterra Analysis of Distortion 7

2.2.1 Harmonic Distortion 8

2.2.2 Intermodulation Distortion 9

2.2.3 Cross-modulation Distortion 9

2.3 Linearity Measurement 11

2.3.1 Total Harmonic Distortion 11

2.3.2 Two-tone Intermodulation Tests 12

2.3.3 Third-order Intercept Point 13

2.3.4 TDFD Intermodulation Method 16

2.4 Design for Low Distortion 16

3 Spectral Analysis of SPICE Data 18

3.1 The Traditional Use of .FOUR 19

3.2 Unconventional .FOUR Analyses 21

3.2.1 Viewing The Noise Floor 21

3.2.2 Use of Two Stimulating Frequencies 21

3.2.3 Example 22

3.3 FFT Post-processing 25

3.3.1 Example 26

3.4 Windowing Non-harmonic Signals 28

3.4.1 Example 29

3.5 A Flexible Post-processor using DFTs 30

3.5.1 The Discrete Fourier Transform of Non-equispaced Data 31

3.5.2 Example 33

3.6 Application of DFT to Measured Time-voltage Data 37

3.6.1 Measurement 1 38

3.6.2 Measurement 2 39

3.7 Complex Circuit Simulation Example 40

3.8 Conclusion 42

4 Characterising Loudspeaker Driver Nonlinearity 44

4.1 Method 46

4.2 The Nonlinear Model 47

4.3 Driver Characterisation 49

4.3.1 Impedance Sweeps 50

4.3.2 Current-displacement Curve 50

4.3.3 Zero-displacement Bl 51

4.3.4 Computation 51

4.4 Example 52

4.4.1 Measurements 53

4.4.2 Computed Functions 53

4.5 Verification of Nonlinear Parameters 53

4.5.1 Distortion Measurements 57

4.6 Conclusion 59

5 A Total Difference-Frequency Distortion Meter 60

5.1 Motivation 61

5.2 Potential of TDFD 62

5.3 Conventional TDFD Meters 63

5.4 Ultimate THD and TDFD Meter Performance 64

5.5 A New TDFD Design 64

5.5.1 Detail 65

5.6 Verification of Performance 68

5.6.1 Calibration for Very Low Level Non-linearities 69

5.7 Examples of Use 71

5.8 Conclusion 73

6 Arbitrary Pulsed Semiconductor Parameter Analyser 75

6.1 History 75

6.1.1 Split-model Response 77

6.2 The APSPA Measurement System 78

6.3 Novel Applications 83

6.3.1 Selecting Pulse Timing 83

6.3.2 S-parameters as a Function of Bias Point 88

6.3.3 Pulse Profiling 88

6.3.4 Breakdown Characteristics 91

6.3.5 Observing Charge-migration Effects 92

6.3.6 Arbitrary Trajectories 93

6.3.7 Channel Temperature Measurement 94

6.3.8 Derivative Extraction 98

7 Gain-derivative Surfaces **99**

7.1 Introduction 99

7.2 The Valve Amplifier Dilemma 100

7.2.1 Measurement System 100

7.2.2 Amplifier Tests 103

7.2.3 Conclusion 120

7.2.4 A Pseudovalve Circuit 121

7.3 GaAs FET Amplifier Design 124

7.3.1 RF-TDFD Measurements 126

7.3.2 Device GDS 128

7.3.3 The Webster Design 132

7.3.4 Results 134

7.4 Conclusion 135

8 Concluding Remarks **137**

A Specifications of VXI-based APSPA **139**

B APSPA MTT-S Presentation **141**

C APSPA User's Manual **155**

C.1 Hardware Installation 155

C.2 Software Overview 157

C.3 Command-line Options 158

C.4 The SET File 159

D Nonlinear Loudspeaker-like Load **165**

D.1 Introduction 165

D.2 Typical Impedance Characteristics 169

D.3 Typical Non-linear Characteristics 169

D.4 Aim of the New Load Circuit 174

D.5 Equivalent Circuit 174

D.5.1 Component Selection in the Linear Subcircuit 175

D.5.2 Component Selection in the Non-linear Subcircuits 176

D.6 Results 179

D.6.1 Power Dissipation 180

D.7 PSPICE Listing 182

List of Figures

2.1 Block diagram of a fundamental-notch type of THD meter. 11

2.2 Plot of output and IMD levels as a function of input level, showing the third-order intercept point, with logarithmic (dB) scales. 14

2.3 Plot of output and IMD levels as a function of input level, showing extrapolation to the third-order intercept point, for a typical modern mixer. . . . 15

3.1 Comparison of results of some Fourier analyses carried out with two fundamental frequencies, using the .FOUR command. 22

3.2 A simple BJT circuit used to illustrate the use of .FOUR. 23

3.3 PSPICE source file for BJT circuit. 24

3.4 Fourier output from PSpice's PROBE for the circuit of figure 3.3. The traces are both of collector voltage, V(coll), one using timesteps limited to a maximum of 10μs, and the other using PSpice-selected timesteps. 27

3.5 Circuit of a simple audio oscillator. 30

3.6 Time-domain display of the startup transients that occur in the audio oscillator circuit of figure 3.5. Note that only the (symmetrical) positive half of the waveform is shown for clarity. 31

3.7 FFT spectrum analysis of audio oscillator carried out using Probe, with and without raised-cosine (Hanning) windowing. 32

3.8 Circuit of a simple RF, BJT oscillator. 34

3.9 Spectrum obtained with DFT post-processor, including windowing, of the collector voltage of the circuit of figure 3.8. Insets show expansions of the time-domain data, including an expansion to show the shape of a single cycle. 36

3.10 Spectrum of the output of a BJT amplifier computed with DFT from a DSO trace dump, and compared with similar SPICE output. The 'scope was set to 2ms/div, with about 5 divisions of deflection, and returned 4096 8-bit data points. The inset shows the fundamental peak, whose shape is set by the window function and the ≈20ms time record length. 38

3.11 Analysis of 8192 samples of a $2V_{p\text{-}p}$ TDFD test signal. Data was originally sampled at 10Msps and decimated to 39ksps before analysis. Insets show relevant TDFD signal peaks at ≈3960, ≈4120, ≈8080 and ≈12040Hz. . . . 39

3.12 Circuit of an example amplifier using bipolar junction transistors (BJTs). . 40

3.13 Comparison of simulated and measured THD as a function of output level for the circuit of figure 3.12. Note that the vertical scale is linear (%) in the top half of the graph and log (dB) in the lower for clarity. 41

3.14 Comparison of a simulated and measured output spectrum for the circuit of figure 3.12 delivering ≈20V peak into 8Ω. 42

3.15 Comparison of a simulated and measured output spectrum for the circuit of figure 3.12 delivering ≈5V peak into 8Ω. 43

4.1 Equivalent nonlinear circuit of the driver. 47

4.2 Diagram of equipment used to measure driver impedance for various dc coil currents. 49

4.3 Diagram of setup for measuring cone displacement against dc coil current. . 50

4.4 Plot of data used for parameter extraction of the Peerless 831921 driver. (Impedance data has been decimated for clarity.) 54

4.5 Plots of the nonlinear parameters extracted for the 831921 driver: L_{e1}, L_{e2}, R_{Le}, R_{et}, C_{mes} and L_{cet} against cone displacement. Note that R_{et} and L_{cet} are box-modified equivalent parameters. 55

4.6 Behavioural model of the nonlinear driver as implemented in SPICE. 56

4.7 Block diagram of the equipment setup for measuring driver acoustic and current distortion. 56

4.8 Comparison of predicted and measured driver current distortion at 36Hz. . 57

4.9 Comparison of predicted and measured driver current distortion at 54Hz. . 58

5.1 Block diagram of a conventional TDFD meter showing the signals present. . 63

5.2 Simplified block diagram of the new TDFD meter. 65

5.3 Detailed block diagram of TDFD meter. 66

5.4 Photograph of the prototype TDFD meter. The various analog sections are visible on the main PCB. A commercial PCB carrying the display and microprocessor is hidden behind the front panel. The toroidal power transformer is visible on the rear panel. 68

5.5 Residual TDFD in prototype meter. 71

5.6 Distortion of an example amplifier measured with the prototype meter. . . 72

5.7 Comprehensive set of measurements made with the prototype. Note that the levels of this standard circuit are below those that may be optained with conventional THD systems. 73

5.8 TDFD versus time for the same buffer amplifier as measured in figure 5.6. . 74

6.1 Typical location of time constants associated with GaAs devices and regions covered by conventional measurement systems are shown graphically on a "speed/level" plane. 76

6.2 Block diagram of the APSPA measurement system. 80

6.3 Layout of the instruments in the VXI rack. 81

6.4 Video image of the assembled VXI APSPA system. 81

6.5 Graphical representation of the measurement sequence finding history-dependence error in a device. 84

6.6 Repeat of some data from figure 6.5 with the points numbered in order of measurement to show the pseudo-random sequencing. 85

6.7 History-dependence error surface for a M/A-Com KF44 MESFET. 86

6.8 Comparison between S_{21} measured under pulsed conditions with parameters measured under conventional, dc conditions. 87

6.9 A device drain characteristic with points marking where S-parameters might be measured to assist in predicting low-signal linearity behaviour. 89

6.10 Drain-current as a function of time for an HP AT44101 FET, plotted on a logarithmic scale. 90

6.11 Characteristics of a GaAs FET measured *significantly* outside the device safe-operating-area (SOA). 91

6.12 Characteristics of a Monofast FET before and a few seconds after exposure to field-induced stress. The graph shows also the response to the stress stimulus. 92

6.13 Drain trajectory of the Monofast FET in figure 6.12 before and after field-stressing with the stimulus simulating operation with a reactive load. 93

6.14 Characteristics of a Triquint FET measured under short pulsed conditions from two different bias points, compared to dc characteristics. Capital Qs mark the bias points. 95

6.15 Magnitude of FET S_{21} measured under static conditions, and isothermal (pulsed) conditions with two substrate temperatures. 96

6.16 Plot of output conductance (g_{ds}) and its first (g'_{ds}) and second (g''_{ds}) derivatives, measured on a Monofast GaAs FET. 97

7.1 Block diagram of the gain-derivative automated measurement system. . . . 101

7.2 A set of gain-derivative surfaces with no DUT in place. The four surfaces depict the fundamental amplitude, the amplitude of second- and third-order components, and the THD. 102

7.3 Fundamental, second-, third-order and THD surfaces for a Playmaster "Sixty-Sixty" with a loudspeaker load. 103

7.4 Fundamental, second-, third-order and THD surfaces for a Playmaster "Sixty-Sixty" with an 8Ω resistive load. 104

7.5 Fundamental, second-, third-order and THD surfaces for the Monarch amplifier with a loudspeaker load. 106

7.6 Fundamental, second-, third-order and THD surfaces for the Monarch amplifier with an 8Ω resistive load. 107

7.7 Fundamental, second-, third-order and THD surfaces for the Lenard prototype amplifier with a resistive load. 108

7.8 Fundamental, second-, third-order and THD surfaces for the Lenard prototype amplifier with a loudspeaker load. 109

7.9 Fundamental, second-, third-order and THD surfaces for the Lenard prototype amplifier with a loudspeaker load, after replacement of the valves. . . . 110

7.10 Fundamental, second-, third-order and THD surfaces for the VTA 30 amplifier with a resistive load. 112

7.11 Fundamental, second-, third-order and THD surfaces for the VTA 30 amplifier with a loudspeaker load. 113

7.12 Fundamental, second-, third-order and THD surfaces for the VTA 80 amplifier with a resistive load. 114

7.13 Fundamental, second-, third-order and THD surfaces for the VTA 80 amplifier with a loudspeaker load. 115

7.14 Fundamental, second-, third-order and THD surfaces for the Audio Note amplifier with a resistive load. 116

7.15 Fundamental, second-, third-order and THD surfaces for the Audio Note amplifier with a loudspeaker load. 117

7.16 Fundamental, second-, third-order and THD surfaces for the Vacuum State preamplifier with a 10kΩ resistive load. 118

7.17 Fundamental, second-, third-order and THD surfaces for the NAD preamplifier with a 10kΩ resistive load. 119

7.18 Fundamental, second-, third-order and THD surfaces for the Fender guitar amplifier with a loudspeaker load. 122

7.19 Circuit of the final stages of the Fender guitar amplifier. 123

7.20 Circuit of the author's "pseudo-valve" amplifier, based on the Fender amplifier circuit. 123

7.21 Fundamental, second-, third-order and THD surfaces for the author's pseudovalve amplifier with a loudspeaker load. 124

7.22 Fundamental, second-, third-order and THD surfaces for the author's pseudovalve amplifier with a resistive load. 125

7.23 Circuit of the bi-notch filter structure used in the combiner filters in the RF-TDFD system. 127

7.24 Circuit of the combiner filters used to isolate sources in the RF-TDFD system.128

7.25 Equipment block diagram for automated RF-TDFD measurement. 129

7.26 Gain surface of an NE33284 HEMT with 50Ω load. A contour plot of the surface is projected onto the XY-plane for clarity. 129

7.27 Second-order intermodulation (IM2) surface for the same transistor and conditions as figure 7.26. 130

7.28 Third-order intermodulation (IM3) surface for the same transistor and conditions as figure 7.26. 131

7.29 A circuit suitable for realising a gain stage with desired distortion characteristics fixed by the derivative superposition method. 132

7.30 Fundamental ($\triangle$), even-order ($\diamond$) and odd-order ($\bigtriangledown$) intermodulation components plotted against gate-source potential for a single device (open symbols) and the designed amplifier (solid symbols). 133

7.31 Fundamental ($\triangle$), even-order ($\diamond$) and odd-order ($\bigtriangledown$) intermodulation component power plotted against input power for a single device (open symbols) and the 4-device amplifier (solid symbols). 136

B.1 Photograph of the author and his supervisor at the HP booth at the 1996 IEEE MTT-S exhibition, with the prototype APSPA system. 143

B.2 Photograph of the prototype APSPA system as demonstrated at the 1996 IEEE MTT-S exhibition. 144

B.3 Display of results from DEMO 1. 145

B.4 Display of results from DEMO 2. 146

B.5 Display of the first comparison result from DEMO 3, showing the dissimilar characteristics resulting from use of different pulse widths. 147

B.6 Display of the second comparison result from DEMO 3, showing the breakup using shuffled-order measurements. 148

B.7 Display of the third comparison result from DEMO 3, showing a "clean" agreement between sequential and shuffled-order measurements. 149

B.8 Display of results from DEMO 4. 150

B.9 Display of results from DEMO 5. 151

B.10 Display of results from DEMO 6. 152

B.11 Display of results from DEMO 7. 153

B.12 Display of results from DEMO 8. 154

C.1 Template .SET file for APSPA. 160

C.2 APSPA timing diagram showing the relationship between timing parameters. 161

D.1 Impedance of a typical, 3-way, commercial loudspeaker. 166

D.2 Impedance of a single-driver system. 167

D.3 Impedance of a 2-way loudspeaker fitted with impedance-compensating networks. 168

D.4 Measurement setup used to determine current linearity of loudspeakers. . . 170

D.5 Harmonic currents and THD in the 3-way loudspeaker. 171

D.6 Harmonic current distortion in the single-driver enclosure. 172

D.7 Harmonic current distortion in the loudspeaker with impedance compensation. Note the relatively flat *fundamental* current. 173

D.8 The circuit of the non-linear loudspeaker-like load. 174

D.9 Circuit used for numerical determination of suitable nonlinear elements. . . 177

D.10 A piecewise-linear approximation to the quadratic V-I characteristic. 178

D.11 Non-linear load circuit modelling a single-driver enclosure. 180

D.12 Simulated (lines) and measured (symbols) impedance of the non-linear load. 181

D.13 Simulated fundamental and harmonic currents, and THD, of the non-linear load. 182

D.14 Measured fundamental and harmonic currents, and THD, of the non-linear load. 183

D.15 Power dissipation in several resistances plotted against frequency with 25Vrms drive for the single-driver model circuit of figure D.11. 184

Chapter 1

Introduction

This thesis describes a number of quite separate achievements with a common theme. The common link is concern with distortion that arises from weak nonlinearity in "low-signal" operation. The adjective low-signal distinguishes the situation from a large-signal one, where overloads such as clipping or slew-rate limiting cause gross distortion, or a small-signal one, where nonlinearity is ignored.

In this work I show that low-signal techniques may be routinely used in electronic engineering. The outcome of the research is to demonstrate the viability of measuring, modelling, simulating, and presenting low-signal performance—to show it worthwhile considering low-signal operation in normal engineering work. It is beyond the scope of this work to produce any particular product; however, an audio meter has been designed and prototyped, a tool (comprising a model and procedure) for analysisng loudspeaker drivers has been developed, a GaAs FET measurement instrument system has been built and is being shown by Hewlett-Packard, and a number of refereed publications based partly or solely on original work associated with this thesis have appeared.

1.1 Chapter Synopses

Chapter 2 defines terms and provides a review and tutorial on the present understanding of nonlinearity. Much of this information has appeared in *Analog Electronic Design*, J B Scott, Prentice-Hall, 1991.

Chapter 3 covers methods developed to allow enhanced spectral analysis of data from computer simulation packages, in particular SPICE in various manifestations, and certain results obtained by means of those methods. The algorithm, embodied in a program, is presently being used in at least one commercial company (M/A-COM) and one research program at another university (University College, London). A description has appeared in the literature as "Modern Guide to Spectral Analysis with SPICE", J. Scott and A. Parker, *IEEE Circuits and Devices Magazine*, vol. 11, no. 5, pp. 10-16, September 1995, and "Distortion Analysis using SPICE" Jonathan B Scott and Anthony E Parker, *Journal of the Audio Engineering Society*, vol. 43, no. 12, December 1995, pp 1029–1040.

Chapter 4 describes the determination, and verification by means of behavioural modelling, of a set of parameters characterising the nonlinearity of a loudspeaker driver. The model originally appeared in "Loudspeaker Impedance Non-linearity and Bi-wiring", Jonathan Scott and Glenn Leembruggen, 95^{th} *Convention of the Audio Engineering Society*, Melbourne, August, 1993, and the unprecedented success of this method has been reported in "New Method of Characterising Driver Linearity", Jonathan Scott, Jonathan Kelly, and Glenn Leembruggen, *Journal of the Audio Engineering Society*, vol. 44, no. 4, April 1996.

Chapter 5 describes an instrument for measuring Total Difference-Frequency Distortion (TDFD), according to the recent Audio Engineering Society (AES) standard. The design has been reported in the literature, see "High Performance Total Difference Frequency Distortion Meter", Jonathan Scott and Dirk Heuer, *Journal of the Audio Engineering Society*, vol. 42, no. 6, June 1994, pp 483–489. This meter is believed to have the highest resolution of any distortion measuring instrument reported to date.

Chapter 6 identifies the need for pulsed measurement of devices fabricated in III-V semiconductor material in order to facilitate adequately precise modelling. Next, an Arbitrary Pulsed Semiconductor Parameter Analyser (APSPA) system, developed to carry out such measurement, is described. Some novel applications are presented. This material has attracted considerable international interest and has appeared in

"Pulsed-bias/Pulsed-RF Device Measurement System Requirements", Jonathan Scott, Mohamed Sayed, Paul Schmitz and Anthony Parker, *24th European Microwave Conference*, Cannes, September 1994, pp951–961;

"Trends in Device Characterisation: A Pulsed Semiconductor Parameter Jonathan Scott and Anthony Parker, Analyser System for III-V FETs", *Journal of Electrical and Electronics Engineering, Australia*, vol. 14, no. 3, September 1994, pp. 196–205;

"RF, Pulsed-IV, Device Measurement System in VXI", Jonathan Scott, Anthony Parker and Mohamed Sayed, *Workshop on Applications of Radio Science*, Canberra, June 25–27, 1995;

"Method to Determine Correct Timing for Pulsed-I/V Measurement of GaAs FETs", Anthony Parker and Jonathan Scott, *Electronics Letters*, vol. 31, no. 19, 14 September, 1995, pp1697–1698.

"Determining Timing for Isothermal Pulsed-bias S-parameter Measurements" Anthony Parker, Jonathan Scott, James Rathmell and Mohamed Sayed, *IEEE MTT-S International Symposium Digest*, San Franscisco, June 17–24, 1996;

"New Applications for Pulsed/Isothermal Test System", Anthony Parker, Jonathan Scott, James Rathmell and Mohamed Sayed, *ARFTG*, invited paper, San Franscisco, June 24, 1996, and

"Modelling and Characterisation of GaAs Devices", A. E. Parker and J. B. Scott, G. Machado (ed), *Low-power HF microelectronics*, chapter 6, IEE, ISBN 0 85296 874 4, UK, 1996.

Chapter 7 describes a new method of 'visualising' the linearity performance of amplifiers, and demonstrates that informative measurement of audio amplifier linearity demands the use of a realistic load. The method appears in:

"Gain-derivative Surfaces: Trying to Visualise 'Valve Sound'," Jonathan Scott and Adrian

Vos, 6^{th} *Regional Convention of the Audio Engineering Society*, September 10–12, Melbourne, 1996,

while the design of the load is covered in Appendix D, and has appeared as:

"A Nonlinear Load for Power Amplifier Testing", Jonathan B. Scott and Greg Lemon, 5^{th} *Convention of the Audio Engineering Society*, Sydney, April, 1995.

An application of the technique to RF travelling-wave amplifiers appears in:

"Control of Circuit Distortion by the Derivative Superposition Method", Danny R Webster, Jonathan B Scott, and David G Haigh, *Microwave and Guided Wave Letters*, vol. 6, no. 3, March, 1996, pp123–125.

Chapter 2

Conventional Theory and Background Material

Between small-signal and large-signal operation lies the operating region characterised by mild nonlinearity. Operation here is variously described as "low-signal", "mildly-distorting", or "weakly nonlinear" operation. Mild nonlinearity and the subtle distortion associated with it form a subject that is not well covered, in any practical sense, in the literature. Involved, mathematical presentations are available,[8] but do not find application in routine, industrial situations.

Virtually all textbooks present methods of linearised, small-signal analysis. Most textbooks describe gross distortion, notably clipping (voltage limitation, or voltage overload), usually in association with load-lines and/or device transfer characteristics.[1, 2] Many discuss slew rate limiting (current limitation, or current overload), but fewer identify its cause, and fewer still identify it as an analogue of clipping; for example see the well-respected textbook [3]. Some books cover nonlinear analysis, but this tends to be limited to elementary power-series discussions associated with measurement of linearity: see for example [4]. The author has summarised some commonly-accepted ideas elsewhere.[5]

Many journal papers discuss the measurement of nonlinearity, and a number address

design with attention to low-signal distortion. These tend to be specific to certain fields, noticeably audio, and more recently microwave and fibre-optics. The reader must be wary, however, as many appealing publications are less than rigorous[6], while some mix misleading statements along with insight.[7]

This chapter presents the current understanding of distortion embodied in such journal papers. The less-general aspects of most of these papers is beyond the scope of this work, but many will be analysed in more detail in the appropriate chapter: for example, papers relevant to the design of loudspeaker driver linearity such as [24], [25], [26], [27], and [28] are discussed in chapter 4.

2.1 Definition of Terms

Distortion is correctly defined as any process whereby the output of a system is caused to differ from its input in some way apart from a scaling factor and an offset. Importantly, distortion may be divided into linear and nonlinear categories. Linear distortion can simply be viewed as the result of filtering, intentional or otherwise. Most usually, however, the term distortion is used to refer solely to the more subtle and more damaging *nonlinear* distortion. This will be the case here. When the distinction is important it will be made clear in context.

Nonlinear distortion results when a signal passes through a nonlinear system. The distortion of a signal may be identified by the addition of new frequency products to the signal; products, in general, which were not present in the original input.

Use of the term "small-signal" implies that the distortion products are insignificant, and can be neglected. Small-signal analysis thus means linear analysis, or analysis of a linear model of a circuit or system. The term "large-signal" usually refers to situations involving gross distortion or "overload". In this thesis, the term "low-signal" will be used to denote a level of signal sufficient to make nonlinear products significant, but that may be well below the level associated with gross overload, and insufficient to move devices out of

their normal operating region. This level of distortion is denoted by various terms in the literature including "mild nonlinearity". For instances see [63, 64, 31].

2.2 Volterra Analysis of Distortion

In the *midband frequency range* of a *time-invariant* system, it is possible and convenient to describe nonlinear distortion by modelling the relevant electronics mathematically in terms of the Taylor's series, invoking Fourier theory and superposition, and then manipulating a Volterra expansion.[1] This is virtually the only treatment found in the general electronics literature.

Consider an electronic system that has a transfer function $\mathcal{T}$. Let the input signal be E_i. Then define the output signal by:

$$\mathcal{T}(E_i(t)) = E_o(t) \tag{2.1}$$

One can analyse the effect that the system will have on any signal by analysing the effect it has on each component of the signal separately. Taylor's method of expansion allows any function to be approximated to arbitrary accuracy as a polynomial series. Hence one may replace $\mathcal{T}$ with a polynomial, confident that the coefficients are determinable, and analyse the original system's effect on a signal by analysing the polynomial's effect upon a sine-wave or the set of sine-waves comprising a signal. This leads to

$$E_o(t) = a_0 + a_1 E_i(t) + a_2[E_i(t)]^2 + a_3[E_i(t)]^3 + ... \tag{2.2}$$

where the input signal $E_i(t)$ may be replaced by the sum of the appropriate set of sine-wave signals. When $a_i = 0$ for $i > 1$—that is, when only the constant and linear terms are not zero—the system is said to be linear.

Several interesting manifestations of distortion may be analysed using this approach.

[1] The distinction between a Taylor's series, derived from a Taylor expansion, and a Volterra series is important. The Volterra series is a special case of the Taylor's series where the independent variable is substituted by a single-frequency sinewave, or generally a sum of such waves, the expansion carried out, and the dependent variable represented with all terms of like frequency collected, the coefficients then representing the total contribution at their associated frequency.

2.2.1 Harmonic Distortion

Consider an input signal of the form:

$$E_i = v_1 \cos(\omega_1 t) \tag{2.3}$$

Substituting into 2.2, and given that a_2 is not zero, $E_o(t)$ will contain the term

$$a_2 v_1^2 \cos^2(\omega_1 t) = \frac{1}{2} a_2 v_1^2 [1 + \cos(2\omega_1 t)] \tag{2.4}$$

while if $a_3 \neq 0$

$$a_3 v_1^3 \cos^3(\omega_1 t) = \frac{1}{2} a_3 v_1^3 [\cos(\omega_1 t) + \cos(\omega_1 t)\cos(2\omega_1 t)] \tag{2.5}$$

which, when expanded, and ignoring phase, includes the term

$$\frac{1}{4} a_3 v_1^3 \cos(3\omega_1 t) \tag{2.6}$$

and so on for each non-zero a_i in the series. Thus, in a nonlinear system a single input frequency can give rise to an output containing not only the input frequency but multiples of it. This is referred to as harmonic distortion, since the new frequencies occur at multiples (harmonics) of the input frequency.

The observation that harmonic distortion occurs leads to the most common test of a system's distorting capacity: the measure of total harmonic distortion, or THD. This is the figure most commonly quoted in specifications of audio amplifier distortion. THD is defined as the fraction of the RMS sum of all output harmonic voltage components divided by the RMS amplitude of the fundamental sine-wave output component.

Using this definition of THD, and denoting the distortion present in a signal x as $\mathcal{D}(x)$, gives

$$\mathcal{D}(V) = \frac{\sqrt{\sum Harm^2[V]}}{Fund[V]} \tag{2.7}$$

It is possible, in any given situation, to express $\mathcal{D}(V)$ in terms of the Volterra expansion, but this is not generally useful. The reverse is not, in general, possible.

2.2.2 Intermodulation Distortion

Harmonic distortion, and the measure of THD, arise from the consideration of nonlinear response to a single sinewave. Similarly, intermodulation distortion describes the nonlinear response to two pure tones. Various measures will be seen to arise from this observation.

Consider an input signal to a system of the form:

$$E_i = v_1 \cos(\omega_1 t) + v_2 \cos(\omega_2 t) \tag{2.8}$$

Considering only the second-order term in the power series expansion with input E_i given by equation 2.8,

$$\begin{aligned} a_2 E_i^2 &= a_2[v_1 \cos(\omega_1 t) + v_2 \cos(\omega_2 t)]^2 \\ &= a_2 v_1^2 \cos^2(\omega_1 t) + a_2 v_2^2 \cos^2(\omega_2 t) + 2a_2 v_1 v_2 \cos(\omega_1 t)\cos(\omega_2 t) \end{aligned} \tag{2.9}$$

The last term is the intermodulation (IM) term. This IM term becomes:

$$a_2 v_1 v_2[\cos((\omega_1 + \omega_2)t) + \cos((\omega_1 - \omega_2)t)] \tag{2.10}$$

Thus two signals give rise to terms in their own sum and difference frequencies.

The third-order term expands to give a similar component:

$$\cdots [\cos(\omega_1 \pm 2\omega_2)t + \cos(2\omega_1 \pm \omega_2)t] \tag{2.11}$$

and so forth. In general, when an n^{th}-order polynomial is used, it can be shown that all possible sums and differences of any n-permutations of available terms can occur.[8]

2.2.3 Cross-modulation Distortion

Cross-modulation distortion (CMD) is responsible for the transfer of sidebands from one modulated carrier to another. Consider the input:

$$E_i = v_1(1 + \mathcal{M}\cos(\omega_m t))\cos(w_1 t) + v_2 \cos(\omega_2 t) \tag{2.12}$$

where the first term is the interfering signal and the second is a second carrier, considered without modulation.

Consider the third order term:

$$3a_3v_1^2v_2\{(1+\mathcal{M}\cos(\omega_m t))^2\cos^2(\omega_1 t)\}\cos(\omega_2 t) \tag{2.13}$$

Squaring the section within braces (note that all the terms therein are of power 2) results in detection of the modulation associated with the interfering signal; subsequent multiplication by the $\cos(\omega_2 t)$ term modulates it onto the other carrier, so:

$$\frac{3}{2}a_3v_1^2v_2[1+2\mathcal{M}\cos(\omega_m t)+\mathcal{M}^2\cos^2(\omega_m t)][1+\cos(2\omega_1 t)]\cos(\omega_2 t) \tag{2.14}$$

which includes the term:

$$=\frac{3}{2}a_3v_1^2v_2[1+2\mathcal{M}\cos(\omega_m t)+\mathcal{M}^2\cos^2(\omega_m t)]\cos(\omega_2 t) \tag{2.15}$$

from which it can be seen that there will be modulation on the second carrier, and in particular some of this modulation is the same as that on the first, or interfering, carrier:

$$=3a_3v_1^2v_2\mathcal{M}[\cos(\omega_m t)]\cos(\omega_2 t) \tag{2.16}$$

Adding the signal term in ω_2—that is the a_1 term and the a_3 term, and ignoring $\mathcal{M}^2$ terms,

$$(a_1v_2+\frac{3}{2}a_3v_1^2v_2)\cos(\omega_2 t)+3\mathcal{M}a_3v_1^2v_2\cos(\omega_m t)\cos(\omega_2 t) \tag{2.17}$$

which reduces to

$$(a_1v_2+\frac{3}{2}a_3v_1^2v_2)[1+\mathcal{M}_c\cos(\omega_m t)]\cos(\omega_2 t) \tag{2.18}$$

where cross-modulation factor

$$\mathcal{M}_c=\frac{3a_3v_1^2\mathcal{M}}{a_1+\frac{3}{2}a_3v_1^2} \tag{2.19}$$

If v_1 is large, this is approximately $2\mathcal{M}$.

The practical implication of this is that a strong signal can impose its modulation on another carrier in the presence of nonlinearity, but the nonlinearity must have a contribution at least at the third-order term of the polynomial.

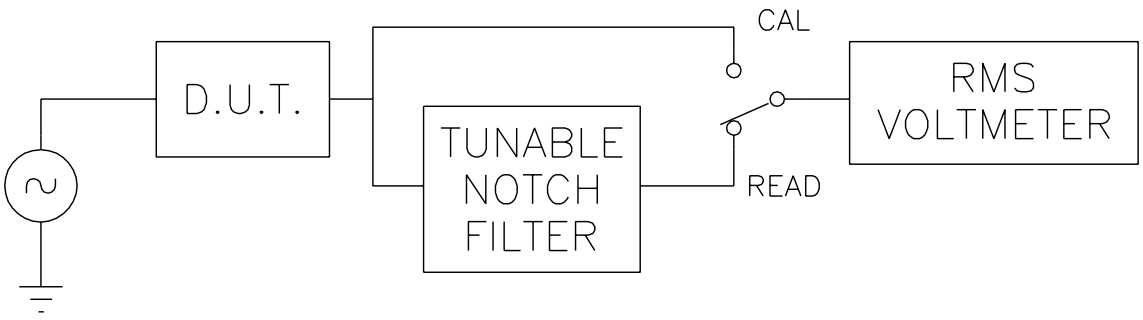

Figure 2.1: Block diagram of a fundamental-notch type of THD meter.

2.3 Linearity Measurement

Characterising linearity—the measurement of distortion—is fundamentally a job of spectrum analysis. A number of tests are traditionally used. Conventional tests usually return a single scalar number, and are carried out at a single fixed frequency and level.

2.3.1 Total Harmonic Distortion

Total Harmonic Distortion (THD) is defined in terms of the output magnitude of a single-frequency signal and its multiples present in the output, when the input is the single stimulus tone, as noted above. Measurement can be straightforwardly carried out with a spectrum analyser. Digital spectrum analysing instruments, such as the HP3561 dynamic signal analyser, will do this calculation automatically.

THD is commonly measured by passing the output signal through a notch filter to remove the fundamental, and subsequently measuring the RMS voltage of the residual in an appropriate bandwidth. The block diagram of such an instrument is given in figure 2.1. This actually measures "THD plus noise", but the two are the same if the noise level is sufficiently low as is common.

The AWA F242 meter offers a dynamic range of 80dB, in a bandwidth from 20Hz to 100kHz, by notching any single tone from this bandwidth to a depth of 80dB, and using a true-RMS meter with 100kHz bandwidth. The meter returns the measure of "ratio of fundamental to noise and distortion". With the usual assumption of low noise, this will be

THD. When the distortion is relatively small, below a few percent as is usually the case, the amplitude of the fundamental at the output is assumed to be the same magnitude as the total output signal. With this assumption, THD is simply the ratio of the RMS voltage without the notch filter to that with it switched in. Hence the measurement is obtained directly without calculation from a meter display decalibrated to read 1 without the notch filter.

2.3.2 Two-tone Intermodulation Tests

The two-tone intermodulation test is a test often used in RF applications to assess the intermodulation occurring in receivers and transmitters. In the case of a receiver, the distortion is important in assessing how well the receiver will resist CMD, and how well it will resist picking up a signal, which is really the difference frequency between two signals to which it is not tuned but which impinge upon the input. In the case of a transmitter simultaneous handling more than one signal or handling wideband signals, it is important as a measure of how much signal is emitted at frequencies that are not those fed to the transmitter input (the so-called "spurious" output).

The basic idea is that two frequencies are fed in, and a search is made for a third-order intermodulation frequency. In the case of a receiver, two frequencies of equal level are typically injected. Call these f_1 and f_2. The frequencies selected are usually frequencies to which the receiver could be tuned, but to which it is not. If there is the potential for intermodulation (IMD) effects, frequencies of $2f_1 - f_2$ and $2f_2 - f_1$ will be present. If the difference frequency $f_1 - f_2$ is small, say 100 kHz, it is easy to tune the receiver to the frequencies $2f_1 - f_2$ and $2f_2 - f_1$, because these IMD terms and the two fundamentals will all fall in a relatively narrow bandwidth readily spanned even by a narrow-band system. The amplitude of the input frequencies is increased until the IMD signal is perceptible above the receiver noise. The level at which this occurs is called the IMD level, and the ratio of this to the receiver input noise floor is the IMD dynamic range, the figure of interest.

In general, two-tone IMD is defined as the ratio of the level of each fundamental to the level of one of the two nearest, third-order IMD products:

$$IMD = \frac{|f_1|}{|2f_1 - f_2|} \tag{2.20}$$

2.3.3 Third-order Intercept Point

When specifying the distortion performance of RF power amplifiers and mixers, a concept called "intercept point" is often used. When measuring THD in audio systems the distortion is measured at some given, single level. This may well be full power, which is to say at the place where naïve expectation is for the distortion to be worst. Using the intercept point definition gives a linearity measure independent of operating level.

In amplifiers, the output power should be proportional to the input power and the gain constant. However, the third-order distortion products arising can be expected to vary as the *cube* of input power. This assumes that in Taylor's expansion, the third-order term dominates higher-order terms. Clearly the distortion level gains rapidly on the desirable output power as the input power is increased.

There would have to be some point at which the IMD products overtook the linear ones if this case continued. In fact, the system will usually overload before this point is reached, but the point exists theoretically, and can be found by extrapolation, as shown in figure 2.2. The apparent power level at which the intersection of the two lines occurs is the "third-order intercept power", "third-order intercept point", "IP3", "3OIP" or just the "intercept point", and this can be expressed as the input or output level. The IMD power level is customarily measured using the two-tone test described above. This test, and the resulting figure of merit are used to define the distortion performance. Note that as the slopes of the lines are known, the intercept point alone tells all about the graph, so it is possible to figure out what the IMD level will be for any input power level once you have the intercept point.

The test can even be applied to mixers. Suppose two signals of identical amplitude at

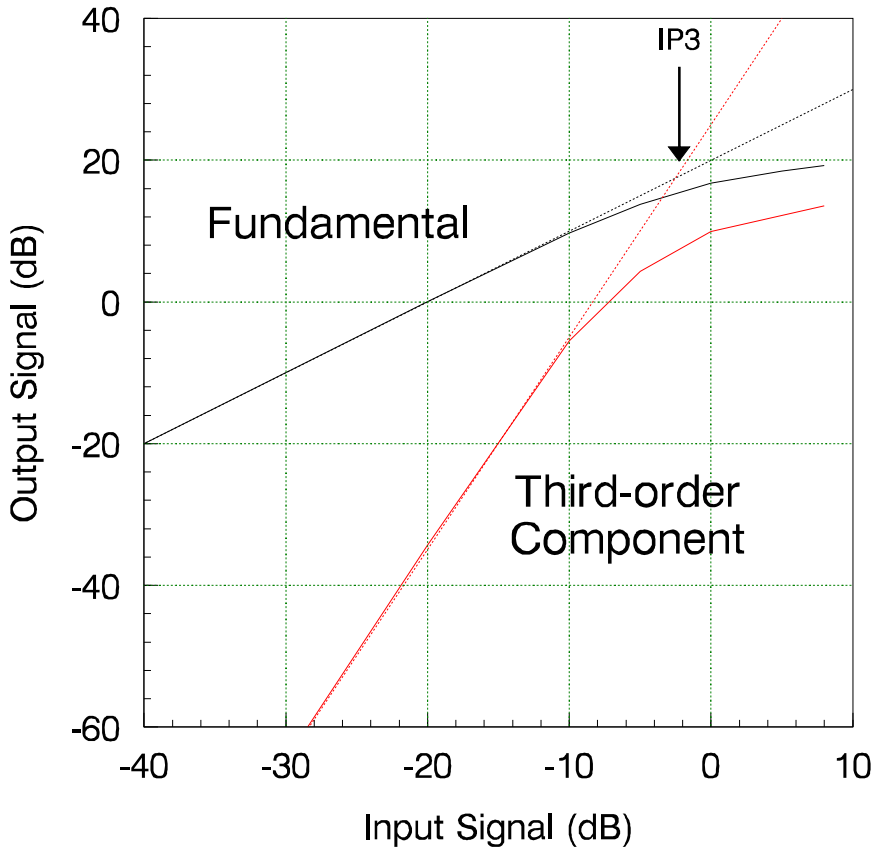

Figure 2.2: Plot of output and IMD levels as a function of input level, showing the third-order intercept point, with logarithmic (dB) scales.

frequencies ω_1 and ω_2 are present on the RF input of a mixer, and the LO signal is at ω_0. The IP3 for a mixer is defined as the input amplitude of the (equal) signals at ω_1 and ω_2 that would (in the absence of gain compression in the circuit) give equal output amplitude of the wanted output signals at $\omega_1 \pm \omega_0$ and $\omega_2 \pm \omega_0$ and the unwanted third-order signals at $2\omega_1 - \omega_2 \pm \omega_0$ and $2\omega_2 - \omega_1 \pm \omega_0$.

A problem with this measure is the dependence for a "sensible" answer on the cubic nature of the IMD level. If the assumption of dominance of the third-order term in the Taylor expansion was unjustified, the concept becomes rather meaningless. Figure 2.3 shows the kind of difficulty that can arise.

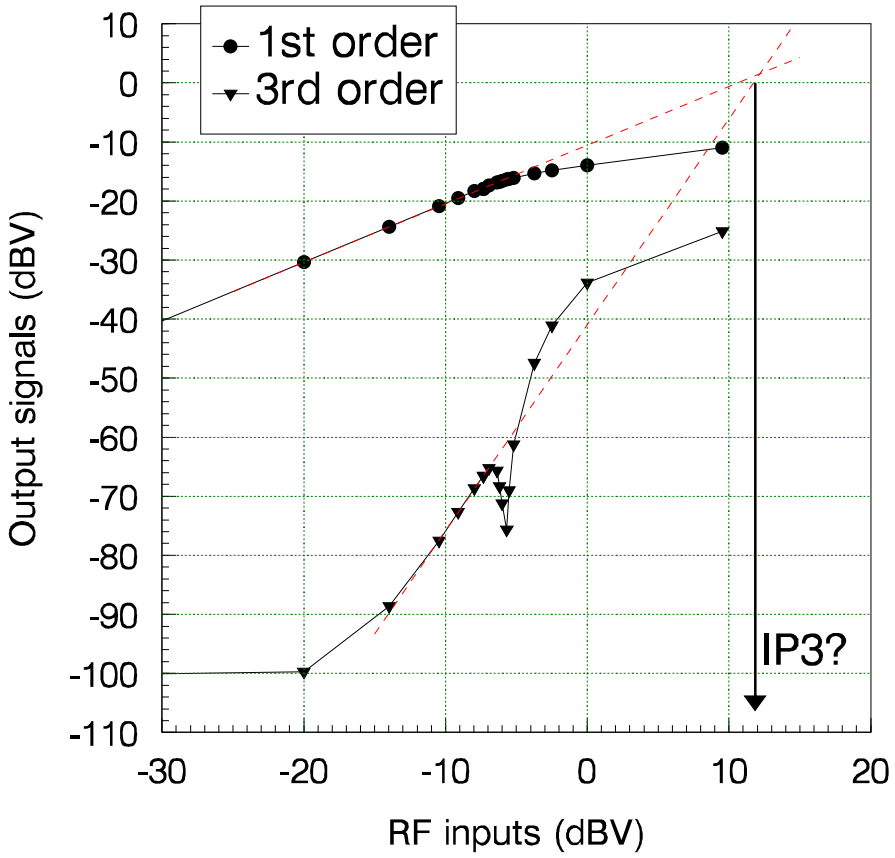

Figure 2.3: Plot of output and IMD levels as a function of input level, showing extrapolation to the third-order intercept point, for a typical modern mixer.

2.3.4 TDFD Intermodulation Method

More recently, IMD tests have been adopted as a standard through the Audio Engineering Society (AES), which is the body responsible for audio standards in the USA.

The total difference-frequency distortion (TDFD) test method is an ingenious audio distortion test, proposed in 1983 by Neville Thiele, then of the Australian Broadcasting Corporation studios in Sydney.[9] The test injects equal amplitude sine-wave signals at $f_1 = 2f_0$ and $f_2 = 3f_0 + \delta$ into an audio system, and then expects signals in a narrow bandwidth, about f_0. The distortion is defined as twice the RMS sum of the distortion products at $f_0 + \delta$ and $f_0 - \delta$, divided by the arithmetic sum of the two input test signals.

This test has the recommendations that:

- the source signals do not themselves have to be particularly pure, as their harmonics will not interfere provided they are somewhat below the level of the fundamentals;
- the signals sought for simple RMS metering may be selected by a narrow bandpass filter, which means a low noise floor (noise is proportional to bandwidth);
- values read out are usually within a couple of dB of those encountered in the conventional definition of distortion (this varies a little with the coefficients of the polynomial), so that no new scale of typical values need be used, and comparisons with old measurements are immediately meaningful.

2.4 Design for Low Distortion

There are two general approaches available for reducing circuit nonlinearity:

1. application of feedback, and
2. choice of circuit topology.

The first method, the application of feedback, relies on the observation that the greater the desensitivity (excess or spare gain) in a feedback loop, the more closely the transfer function of an amplifier approaches that set by the feedback components rather than that of the amplifier itself. Typically the feedback components are passive and can be very linear elements. Unfortunately, the extent to which feedback can be applied is limited by the onset of instability.

"Circuit topology", as a method of minimising distortion, simply means the arrangement of active components, the use of various configurations such as differential pairs in preference to other alternatives, etc.

Chapter 3

Spectral Analysis of SPICE Data

SPICE has become a defacto standard for analog circuit simulation. It is now applied even to valve circuits.[10] It copes well with gross distortion mechanisms such as clipping (voltage overload) and slew limiting (current overload) that result when devices saturate or cut off. Although SPICE is the premier, industry-standard, analog electronic simulation program, it is notable for its lack of successful application to the problem of simulating low-signal nonlinearity. This region is important in modern engineering, being responsible for harmonic and intermodulation distortion in high-fidelity music reproduction electronics, and cross-modulation between carriers in wideband RF amplifiers, for example.

SPICE has not been popular in this role in the past for several reasons. The chief cause was that models of devices were crude, and did not represent real-world behaviour with adequate accuracy. This objection is now vanishing, as models improve.

Low-signal distortion is characterised by measures, described in chapter 2, that involve the separation of signal components in the frequency domain. Nonlinear analysis using SPICE thus requires a spectral analysis of the large-signal (.TRAN) analysis output. Similar (time-voltage) data is also provided by Digital Sampling Oscilloscopes (DSOs) used to measure systems, and the same analysis tool can be used with data from either source.

3.1 The Traditional Use of .FOUR

The usual method of distortion analysis in SPICE is to perform a transient analysis and then execute a .FOUR command. This performs a Discrete Fourier Transform (DFT) at a specified frequency and its harmonics, in order to obtain an estimate of the magnitude of the spectral content at those frequencies. It then returns a calculation of the THD. The method of carrying out this procedure and examples of the output obtained are described both in SPICE manuals and common textbooks such as [11].

It is common experience for such results to be misleading, however. A rule of thumb states that high confidence in the results of the .FOUR command may be gained by repeating it with a longer duration transient analysis, and with a smaller timestep. If the two answers are the same, there was no problem. This is similar to the laboratory technique to assure that a measurement has been made at "small signal" levels: repeat at a higher or lower level, and check that the answer is the same.

Why the need to allow this "settling time" in the transient run? (The .FOUR command uses the transient analysis output data over the last period of the specified fundamental, that is from TSTOP to TSTOP-TPERIOD, discarding the earlier part of the results.) Using data covering exactly one period of the fundamental, it performs no windowing of the data, and will give erroneous results if any signal is present that is not periodic in this interval. Some SPICE texts, eg. [12], do discuss this, but do not make clear the implications. The most obvious is that it is assumed that the signal contains only components that are harmonics of the specified ("fundamental") frequency. At first, this seems to demand only that all sources in the circuit generate waveforms that are periodic with the same period, that is either the period of the fundamental or a sub-multiple of it. However, there is a more subtle aspect. When a transient analysis is carried out, the initial conditions, which occur at time zero ($t = 0$), are virtually never the same as the conditions one period later; this is only the case in the absence of reactive elements (unless great trouble is taken to precisely set their initial conditions!). In mathematical terms, the solution of the differential circuit equations consist of the sum of a periodic part and a decaying, transient part (the so-called

homogenous and heterogenous solutions). The transient part requires some time to decay to a negligible level—typically a couple of cycles for small distortions. Consequently, it is generally a very bad idea to perform a transient analysis whose duration is only one period. It is necessary to allow the circuit to relax to provide the data that is suitable for .FOUR, so the transient analysis should run for longer than one cycle, typically several.

Some texts do not describe this problem at all, and even carry out .FOUR examples with the analysis duration set to exactly one period. They do this on circuits without reactive elements where it does not matter, which is a shame, as most circuits of interest have capacitors at least! To be fair, [11] and [13] both gloss over .FOUR because they are moving to the FFT facility of a post-processing program as a "better" method. However, treatment of the more general approach has little more depth either in these texts, or the post-processing program's manuals[15, 14].

The .FOUR command has another potential weakness. It assumes that conditions have been set to ensure that it is provided with sufficient information (number of data) in the transient analysis data it uses, namely the last period's worth. In practice, a minimum of about 100 points is needed for a low-resolution harmonic analysis, and at least 1000 points in order to obtain results comparable to an analogue distortion meter or spectrum analyser. SPICE usually interpolates the data onto an equispaced grid before performing the DFT. If the point density is too low, numerical inadequacy will pollute the answer. In addition, interpolation can itself introduce "noise". These considerations may demand that the timestep be forced down, and thus the total number of points up. This can make the calculation quite long, especially if several cycles must be passed first in order to remove transient settling effects.[1] The simplest method of checking is as described above—if a run with double the point density gives the same answer, the point density was sufficient.

[1]Note that the increase in point density will *not* hasten the decay of the transient solutions, which are a natural property of the circuit. Increasing the point density only gives the interpolator more data with which to work. Unfortunately, the increase in point density also increases the effort involved in computing the earlier results that will subsequently be discarded.

3.2 Unconventional .FOUR Analyses

It is possible to perform Fourier analysis, and obtain an estimate of THD simply by following the rules of section 3.1. This will be adequate for many applications. This section will describe more techniques for using .FOUR; section 3.3 will discuss differences when using so-called FFT post-processing packages. However, if you wish to have a solid understanding of discrete Fourier analysis, there is no substitute for reading a tutorial or a textbook on the subject. An interesting tutorial can be found in [16], while [17] and [18] are good references.

3.2.1 Viewing The Noise Floor

The .FOUR command normally computes the spectral content only at the frequencies whose magnitudes you specify. Thus there is no feel given for the "noise floor", or returned level at frequencies where a zero result is expected. This level can be very useful: If high, it can indicate problems that cast doubt on the reliability of the analysis. If all is working well in a spectral calculation, the noise floor is typically low, set by the resolution of the numbers provided to the spectral calculation routine—the quantisation noise. In various versions of SPICE, the dynamic range (ratio of signal to numerical noise) coming out of the .FOUR analysis varies around 80–120dB. You can force .FOUR to provide you with some intermediate results by organising it to operate at a lower frequency—a sub-multiple of the required answers—and to compute several times the number of harmonics actually required.

3.2.2 Use of Two Stimulating Frequencies

It is possible to obtain a spectral analysis using .FOUR when two (or more) frequencies are involved in a circuit. If all frequencies in a circuit are multiples of the fundamental specified to .FOUR, then the analysis will be effective and reliable. If all frequencies supplied to a circuit (by sources within it) are multiples of the specified "fundamental", then the

sum and difference products (intermodulation or cross-modulation products) must also be multiples.

3.2.3 Example

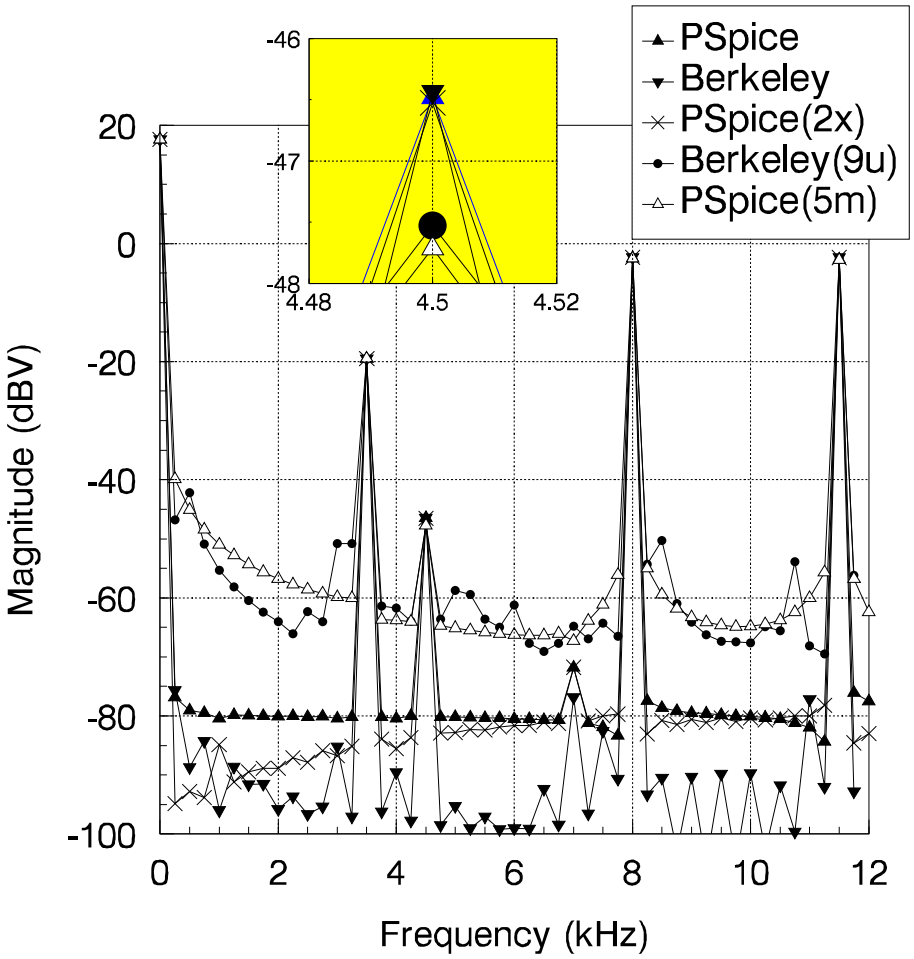

Figure 3.1: Comparison of results of some Fourier analyses carried out with two fundamental frequencies, using the .FOUR command.

Figure 3.1 shows a plot of .FOUR command output from the circuit of figure 3.2 whose SPICE input file appears in figure 3.3. The circuit is a single-BJT amplifier, driven by a pair of signals: 10mV at each of 8kHz and 11.5kHz. The spectral analysis is carried out using .FOUR at 250Hz, with 48 points. With these inputs a TDFD analysis[9] is being performed. The details of this test are not important here, save to know that the important information is found from the magnitudes of the two stimulus frequencies,

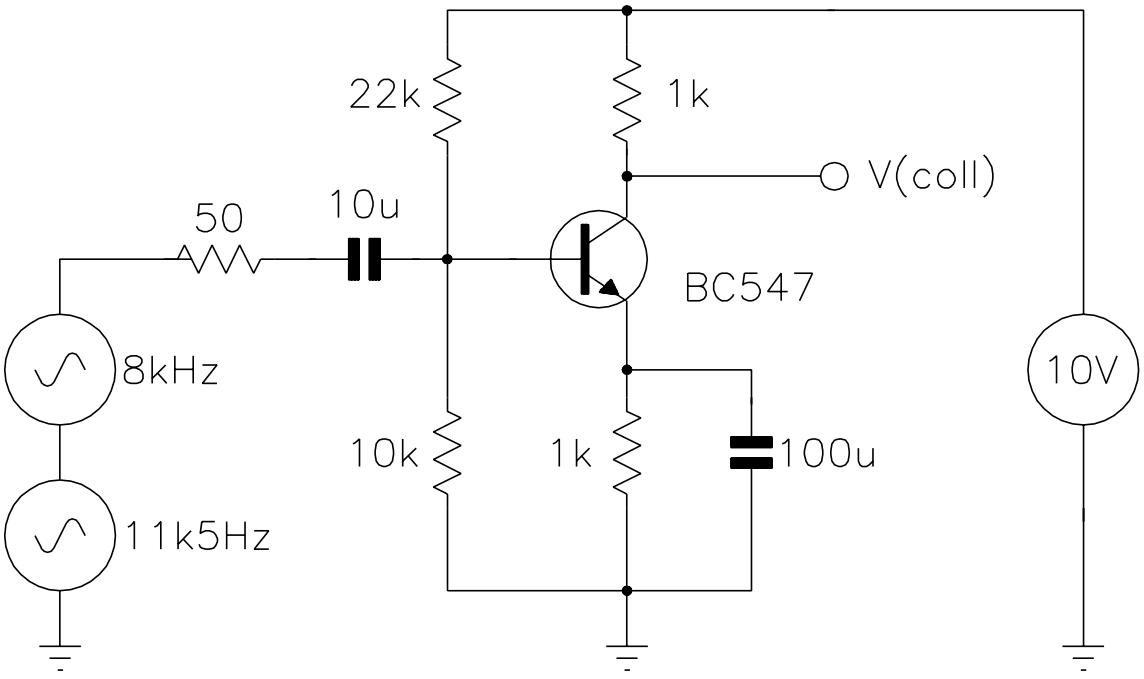

Figure 3.2: A simple BJT circuit used to illustrate the use of .FOUR.

$2\omega_0 = 2\pi 8\text{kHz}$ and $3\omega_0 - \delta = 2\pi 11.5\text{kHz}$, and the components at $\omega_0 + \delta = 2\pi 4.5\text{kHz}$ and $\omega_0 - \delta = 2\pi 3.5\text{kHz}$.

Two different stimulus frequencies are present. The .FOUR command is given a fundamental frequency that is at a sub-multiple of the stimulating frequencies, chosen so that all frequencies of interest (and in fact, all the frequencies that can be expected) are multiples of the "fundamental". Also, the sub-multiple chosen is not the largest that might fill the above requirements (500Hz, since 8kHz and 11.5kHz and all the sums and differences of their multiples will also be multiples of 500) but possesses numerous harmonics at which no significant signal is expected to appear—the odd multiples of 250.

The results are very interesting. The basic run with MicroSim's PSpice shows a noise floor at about -80dBV. The collection of points at that level, even where the ideal result would be zero (-∞dB), warns against trust in any result approaching that level. It is tempting to say that there is a signal at 7kHz, but it is getting close to the noise, and the result there should be treated with some suspicion.

The same circuit run through Berkeley's SPICE 3f2 compiled on a PC returns a different

```
Test of .FOUR analysis ------------------
Vin1 1 0 dc 0 ac 1 SIN(0 .01 8k)
Vin2 in 1 dc 0 SIN(0 .01 11.5k)
Rin in 2 50
Cin 2 base 10uF
Rbu base vcc 22k
Rbl base 0 10k
Q1 coll base emit BC547 1
Rc coll vcc 1k
Re emit 0 1k
Ce emit 0 100u
.model BC547 NPN(Is=10E-15 Vaf=100 Bf=300
+  Xtb=1.5 Br=1 Nc=2 Cjc=5p Cje=5p
+  Ikf=30m Tf=100p RB=100 RE=.2)
Vcc vcc 0 dc 10
.tran 1u 10m 0 10u
.four 250 48 V([coll])
.end
* --------------------------------------
```

Figure 3.3: PSPICE source file for BJT circuit.

result. There is excellent agreement on the signals of interest in the TDFD test, lending confidence to both simulators, and the results. However, the Berkeley graph shows a lower noise floor—about -100dBV—and an erratic spectral structure. It seems to confirm the presence of a 7kHz signal, but it disagrees with PSpice as to the magnitude.[2]

Running PSpice with the total transient duration doubled and the density of points doubled (1 point per 5μs over 20ms now, the trace labelled "PSpice 2x") leads to a trace differing only in a reduction of the noise floor at lower frequencies and the quadruple duration required to obtain it.[3] Another "signal" pops up at 1kHz, but it does not look trustworthy in the light of the noise floor level! (The Berkeley version compiled and run on a Sun workstation gives a dynamic range of over 200dB when run with sufficient duration and density. This may be a result of greater numerical precision.)

The next trace results from the Berkeley simulator run with a forced maximum timestep of 9μs instead of 10μs. The *fourier* command (equivalent to .FOUR in PSpice and Berkeley

[2]In fact there is a possible fourth order intermodulation term there at 7kHz.

[3]The PSpice noise floor may arise from its numerical methods, known to be different from those in the Berkeley simulator, for which to source code is available.

SPICE 2) performs an interpolation, in this case of order one.[4] The result of interpolating from the inconvenient spacing of 9μs is a marked rise in the "noise floor", despite the increase in data available. This interpolation noise may often arise when the original data is not appropriately spaced. This will be discussed further in 3.3.1, and section 3.5. It is usual to have to force a maximum timestep in order to obtain a satisfactory point density, and the number is typically a suitably round submultiple of the period, so the problem often takes care of itself.

The inset in the graph expands the results about the wanted 4.5kHz point. Being the smallest of the peaks whose magnitude is actually important in this test, it is the best place to look for unacceptable signs of any problem. The inset clearly shows the error introduced by the interpolation of the data.

If PSpice is run with a duration of 5ms instead of 10ms, a similarly upsetting rise in the noise floor occurs. With this change only the first 1ms, not 6ms, of data is ignored. This time the error comes from the interference of the non-harmonic, transient part of the solution that has not been given long enough to decay. The noise floor has the characteristic spectral shape of the transform of a sawtooth, which is to be expected as the decaying exponentials sampled across the 4ms window (set by the use of 250Hz) will have this shape. (This long transient part of the solution is due mainly to the 100μF emitter bypass capacitor and is not related to the frequencies present.)

3.3 FFT Post-processing

Many versions of SPICE have associated post-processors: PSpice has PROBE, Berkeley SPICE has NUTMEG, Intusoft has IntuScope, etc. Most post-processors offer Fast Fourier Transform (FFT) analysis for time-domain SPICE output. This approach has two important numerical differences which separate it from the inbuilt .FOUR analysis. Firstly, it usually does *not* automatically truncate the time-domain data and use only the last

[4]The Berkeley simulator permits the order to be increased, which would improve the results in this case.

period's worth for any particular frequency. Secondly, it uses the FFT algorithm which generally computes 2^n points equispaced in the frequency domain, using 2^n points equispaced in the time domain. This algorithm is used in the interests of speed—it is much faster, especially for large data sets where all frequency points are wanted.

Using an FFT in a post-processor, the user must set the transient analysis limits, including TSTART and TSTEP, the maximum timestep value, to obtain a suitable interval of data after a suitable delay, with suitable point density.[5] Since SPICE will virtually never return 2^n equispaced points, an interpolation must be used, and the resulting noise accepted. Also the user must settle for 2^n frequency-domain results at multiples of the reciprocal of the duration of the time data. These frequencies can only be changed by changing the time window, which usually requires a re-run. These minor concerns must be weighed against the convenience and speed offered by post-processing programs.

No common SPICE textbooks address the problems noted here. Further, graphs presented in [13] and [11] using PROBE are presented with a *linear* Y-axis scaling, rather than a logarithmic or dB scale. This has the effect of concealing the workings of the transform even more effectively than the usual, sparse, non-graphical .FOUR command. Any shortcomings arising from problems of data density, reactive settling of the output and interpolation will likely be as invisible as if traditional .FOUR analysis output alone was available.

3.3.1 Example

Figure 3.4 shows PROBE Fourier analysis results for the same circuit as before, with user-defined and PSpice-defined timesteps. The small black dots indicate points which are produced by the FFT algorithm, while the other symbols identify the traces. PROBE interpolates linearly onto the screen between points returned from the FFT. The x-axis stretches out to far beyond 20kHz, but the part which is less interesting has been deleted from the diagram for clarity.

[5] Note that changes made to the displayed X-axis range may *not* be reflected in the data used when a Fourier analysis is carried out: The Fourier command may use the full available data set irrespective of axis setting.

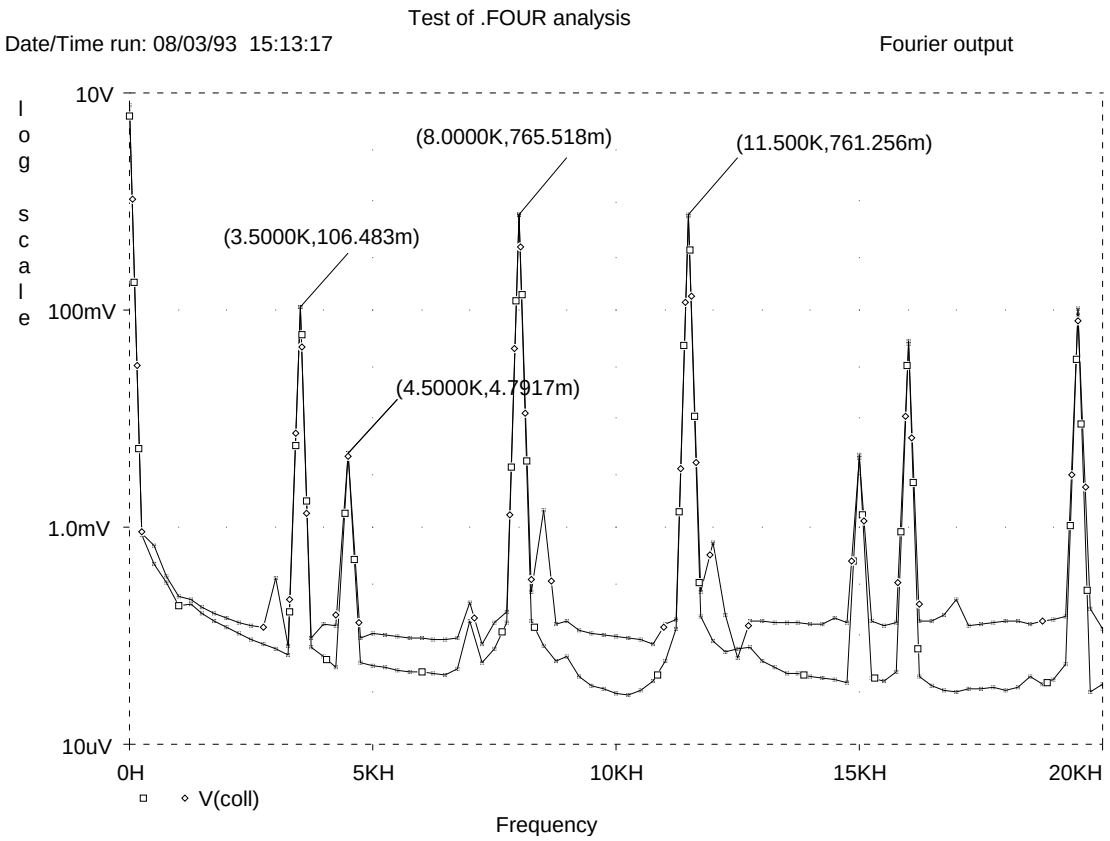

Figure 3.4: Fourier output from PSpice's PROBE for the circuit of figure 3.3. The traces are both of collector voltage, V(coll), one using timesteps limited to a maximum of 10μs, and the other using PSpice-selected timesteps.

The trace identified by the small diamonds, with the higher "floor", is the trace resulting from allowing PSpice to choose timesteps. By allowing PSpice to adjust its transient analysis timestep, as is usual when no Fourier analysis is intended, the data which is passed to PROBE requires interpolation onto a grid of 2^n points from highly irregular spacings. This interpolation can introduce a significant amount of noise. (The interpolation probably happens anyway. The noise introduced is less when the interpolation is from a grid which has substantially the same spacings as the grid onto which the interpolation is going, and when the source data is more dense.) In fact, PSpice chooses its timesteps very well, and PROBE performs its interpolation very well. PSpice has set a noise floor of about 80dB, adequate for most situations, and the rise in noise with interpolation is typically less than 10dB. The SNR is thus very similar to that obtained from .FOUR.

3.4 Windowing Non-harmonic Signals

Both SPICE's inbuilt DFT routines and the FFT of post-processors can be used with waveforms which contain a frequency that is not known or not constant, or with compound signals with components that are not simply, or not at all harmonically related. One example is a free-running oscillator. Such situations are uncommon. Intermodulation tests can mostly be performed in SPICE with more conveniently selected frequencies without the change materially affecting the result (8k and 11k5 have been used above, instead of 8k and 11k96 tha might be used in a real measuring instrument). Frequency domain results are not often required in the case of circuits with instabilities which depend upon large signal effects. Nevertheless, the need can arise.

The solution is to window the data: a well-used process in practical DSP systems, but one not in common use with CAD signals. Windowing means to multiply the data by an envelope, commonly a raised-cosine or other bell-shaped function. The purpose of this is to flatten the ends of the span of data, removing possible discontinuity produced by circular replication of the data. (The discrete transform assumes this circular replication, which would give rise to a step if the starting and finishing levels are not identical, and a point of non-differentiability if the derivatives at the ends of the data set are not all equal. This jump injects signals at all frequencies, which pollute other values and raise the apparent "noise" floor.) The window is otherwise selected to disturb the data as little as possible. It is not our intention to embark here upon a DSP-style description of windowing effects with the discrete Fourier transform. An excellent and comprehensive reference for the interested reader is [19], while [20] gives an interesting implementation example.

Windowing is usually easy in post-processors: the waveform to be analysed can be multiplied by an arbitrary mathematical function before it is transformed. When using the .FOUR function, or when the post-processor does not easily allow multiplication of the waveform by a mathematical function, the windowing can be achieved in the SPICE circuit. This can be done with an independent and a dependent source. For instance, the four lines

```
Vwindow win 0 dc 0 SIN(1 -1 250)
R1placebo win 0 1k
Eprod prod 0 VALUE = {V(coll)*V(win)}
R2placebo prod 0 1k
```

can be added to the example discussed above, and the analysis carried out on V(prod) instead of V(coll). (These lines are in PSpice format. Note the necessity for "placebo" resistors in parallel with the voltage sources in order to avoid the error generated by having less than 2 connections at a node. This error is not generated by some versions of SPICE, e.g., Berkeley's)

The window has two effects: the desired elimination of the noise arising from the step jump, and the introduction of its own distortion. The art of window selection involves getting the least distortion, or the type of distortion which is least objectionable for a given application. A good all-round window is the raised-cosine, which involves multiplying the time domain data by the function $1 - \cos(2\pi ft)$, where f is the fundamental frequency whose period is the length of data you have. This window, used with a search for f and its harmonics only, distorts in a simple way: the result at each frequency point affects adjacent points. For example, if there is a peak at $7f$, half its value gets added to the results at $6f$ and $8f$. This problem is overcome if the frequency points are twice as close together as needed, or conversely, the time domain data is twice as long as necessary.

3.4.1 Example

The oscillator circuit shown in figure 3.5 can be simulated using PSpice, and its output analysed with Probe. Figure 3.6 depicts the startup transient that occurs as a result of the marginally stable amplitude control loop. Figure 3.7 shows an FFT spectral analysis carried out using Probe, on a suitably-delayed, 100ms-long portion of the output voltage, both with and without windowing. The improvement in the signal-to-noise ratio is clearly visible: The third harmonic, some 70–80dB down from the fundamental, becomes discernible. The fundamental is at (roughly) 1.05kHz.

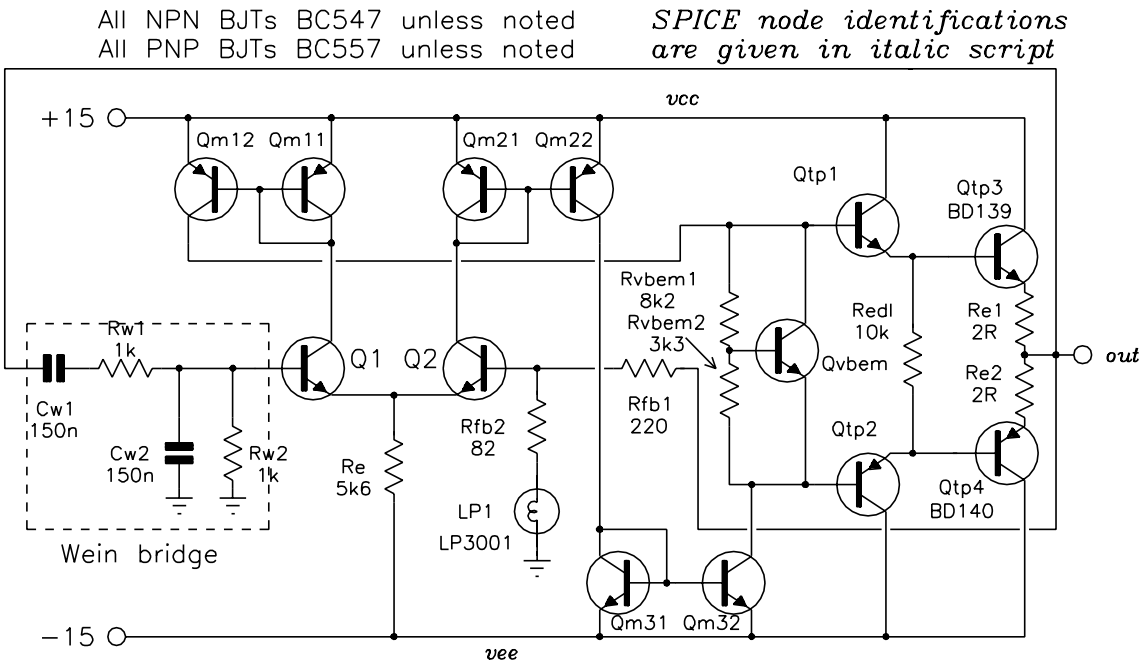

Figure 3.5: Circuit of a simple audio oscillator.

3.5 A Flexible Post-processor using DFTs

All the spectral analysis capability of the .FOUR command and of FFT routines can be provided by a simple program operating on the time-domain data provided by SPICE. What is required is a program which may be applied to *any* length of data with *any* spacing, and which will return a good spectral estimate at *any* requested frequency, or an error message if this is not possible for some reason. A program called "DFT" has been written to fulfill the need. This program has been used extensively over the last few years by university students and other researchers, and in a commercial design situation.

DFT accepts input in the form of time-voltage pairs of values from an ASCII file. Such a file is easily obtained by editing SPICE .PRINT output, from post-processors such as Berkeley's NUTMEG, or indirectly using special ASCII output facilities such as PSpice's .PROBE/CSDF command. Output is in the form of ASCII frequency-magnitude data, at any given frequency and requested multiples of this fundamental frequency. DFT normally applies a raised-cosine window of appropriate width to the data, spline interpolates[21] to a regular grid if point spacing is excessively erratic, and looks after details associated with discrete Fourier transformation. The output can be read just like SPICE output, or can

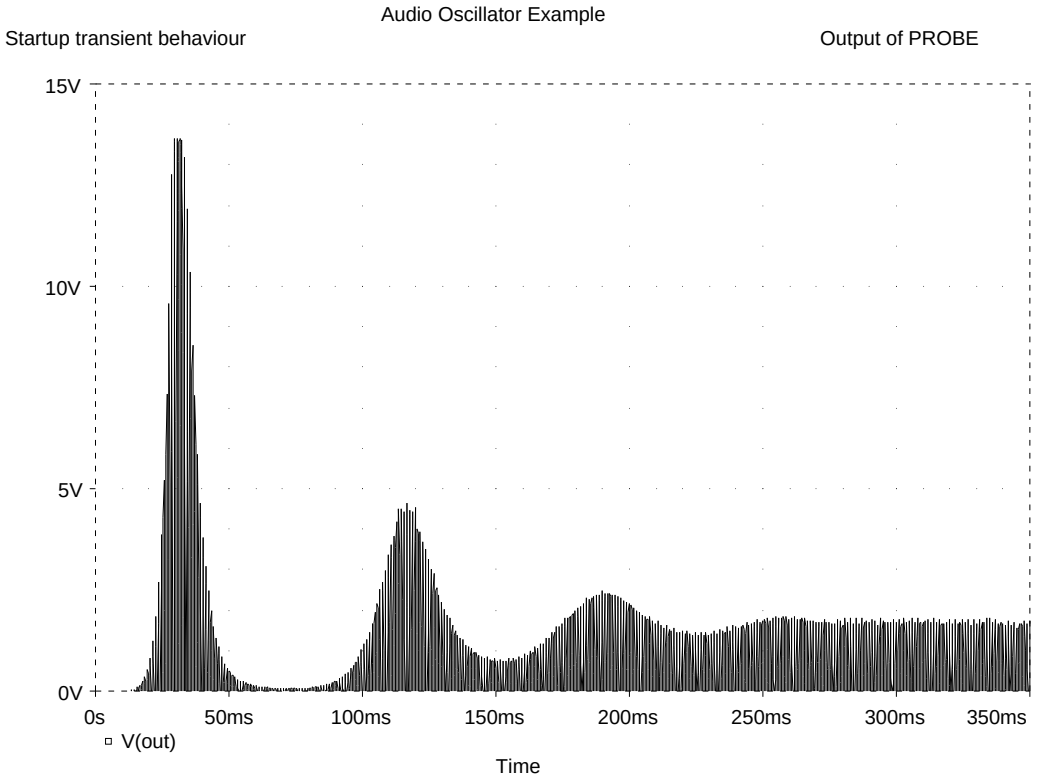

Figure 3.6: Time-domain display of the startup transients that occur in the audio oscillator circuit of figure 3.5. Note that only the (symmetrical) positive half of the waveform is shown for clarity.

be displayed in graphical form with any commercial data presentation packages such as Graftool, Stanford Graphics, xgraph, gnuplot, etc.

An optional parameter can be used to inhibit windowing, or to force or prevent interpolation. While DFT makes a sensible decision on whether to interpolate, there will be cases where the opposite is better, and the option allows this to be tested. Furthermore, a method better than trapezoidal integration for use with erratically spaced data, has been developed.

3.5.1 The Discrete Fourier Transform of Non-equispaced Data

The spectral component is obtained from the continuous Fourier integral

$$F(\omega_i) = \int_{-\infty}^{+\infty} f(t) e^{-\jmath 2\pi \omega_i t} dt \tag{3.1}$$

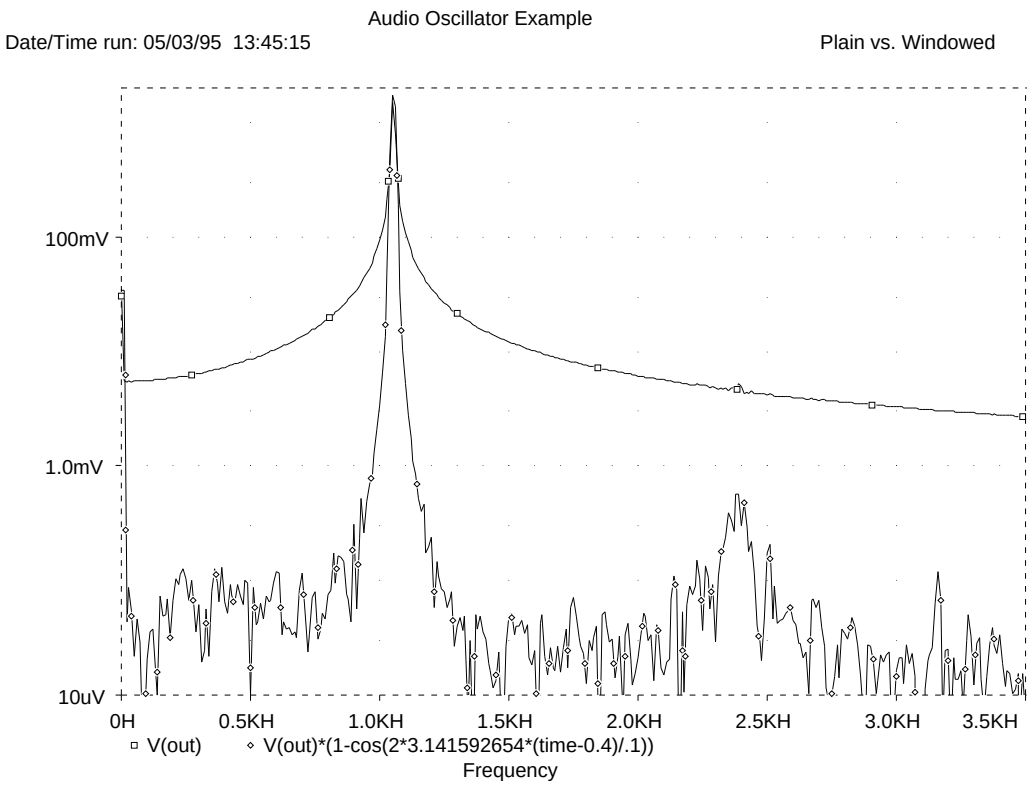

Figure 3.7: FFT spectrum analysis of audio oscillator carried out using Probe, with and without raised-cosine (Hanning) windowing.

With n finite, discrete time data, this may be rewritten without loss of generality as

$$F(\omega_i) = \sum_{k=0}^{k=N-1} \int_{t_k}^{t_{k+1}} f(t) e^{-\jmath 2\pi\omega_i t} dt \tag{3.2}$$

With the data equispaced, and performing the integral by the trapezoidal method, this becomes the familiar equation

$$F(\omega_i) = \sum_{k=0}^{k=N-1} f(k\Delta t) e^{-\jmath 2\pi\omega_i k\Delta t} \tag{3.3}$$

which is the sum of the areas of the trapezia whose two vertical heights are the products of the two functions. This is trapezoidal integration, or integration of the linear interpolation of the functions between the known data points. It is possible to perform the integration by other methods.

SPICE is apt to produce few data points per unit time where the circuit voltages are moving slowly, and many at other places. Linear approximation of $f(t)$, and of the exponential in equation (3.1) for larger ω_i, can be a bad approximation over these larger steps.

The calculation implied by equation (3.2) may be carried out, even with unequally-spaced data, by the trapezoidal method, but also by other methods.

Suppose we wish to integrate across one irregular interval (the k^{th} one, say) presuming that the function $f(t)$ is linearly interpolated, but without making the same presumption for $e^{-\jmath 2\pi\omega_i t}$. This would involve completing the integral

$$\int_{t_k}^{t_{k+1}} y(t)e^{-\jmath 2\pi\omega_i t}dt \tag{3.4}$$

where $y(t)$ is the linear interpolation $y(t) = at + b$. For, say, the imaginary part this becomes

$$I_k = \int_{t_k}^{t_{k+1}} y(t)\sin(2\pi\omega_i t)dt \tag{3.5}$$

Integrating,

$$I_k = [\frac{a\sin(2\pi\omega_i t)}{(2\pi\omega_i)^2} - \frac{y(t)\cos(2\pi\omega_i t)}{2\pi\omega_i}]_{t[k]}^{t[k+1]} \tag{3.6}$$

which is readily evaluated. Similarly for the real part

$$R_k = [\frac{a\cos(2\pi\omega_i t)}{(2\pi\omega_i)^2} + \frac{y(t)\sin(2\pi\omega_i t)}{2\pi\omega_i}]_{t[k]}^{t[k+1]} \tag{3.7}$$

leading to

$$F(\omega_i) = \sum_{k=0}^{k=N-1} (R_k + \jmath I_k) \tag{3.8}$$

3.5.2 Example

The circuit in figure 3.8 is a simple BJT, LC oscillator, tuned very roughly to 1MHz. Suppose it is desired to investigate the spectral output of this circuit, particularly with some variation of the supply to the circuit. Performing a spectrum analysis on this can prove difficult, as the frequency is not precisely known and may vary with power supply level. The circuit is set to operate with 20μs time to stabilise, 200μs to collect data, and a 20kHz sinewave variation in the power supply. This data window—200μs—will allow 5kHz resolution, or 10kHz with the raised-cosine window. The expected answer includes the signal at the fundamental, circa 1MHz; one might also expect some sidebands spaced 20kHz from the fundamental; there will also be harmonics of the fundamental. Will the harmonics have sidebands? (Some thought suggests the answer will be yes.)

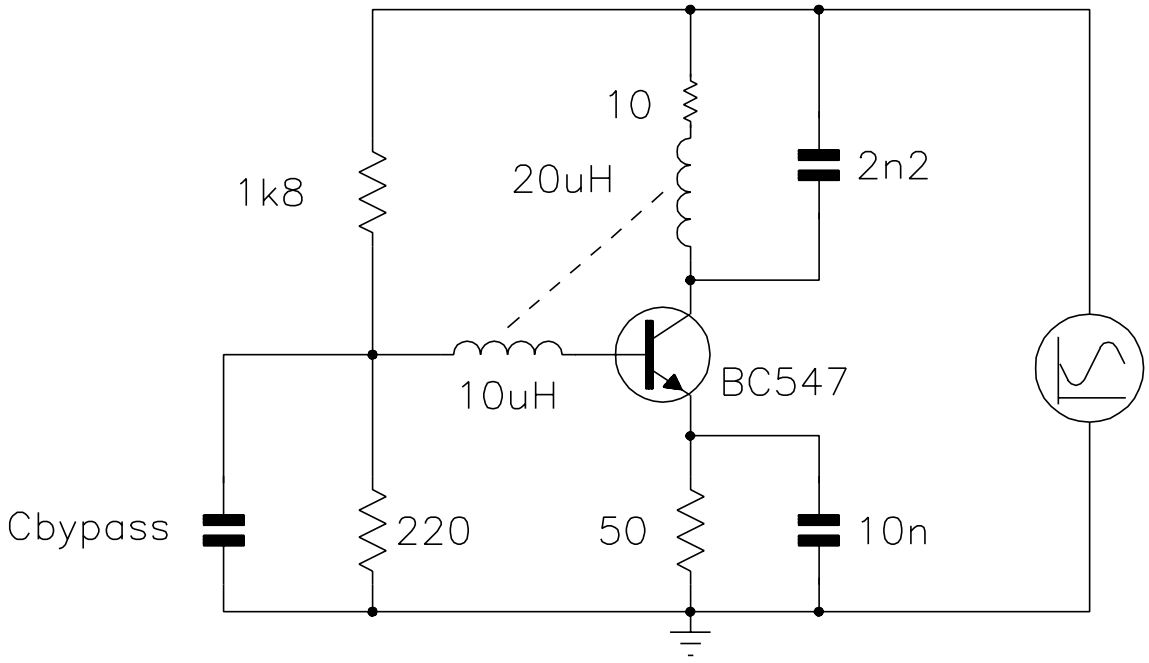

Figure 3.8: Circuit of a simple RF, BJT oscillator.

Figure 3.9 shows the results of the simulation, obtained using the program described in section 3.5, applied to the results obtained from PSpice. A program, "CSDF2TXT" was written to convert PSpice ASCII transient analysis output (optained using the /CSDF option of the .PROBE command) format to ASCII time-voltage pairs. The following batch file shows how DFT was used to provide the data for figure 3.9:

```
rem Get spectrum of V(coll) from OSC.CIR
rem Perform SPICE simulation
call pspice osc
rem Extract time-voltage data
csdf2txt coll osc.txt >osc.dat
rem Group of points about fundamental
dft.exe osc.dat 5000 40 120 >osc3f.dft
rem Groups of points about harmonics
dft.exe osc.dat 5000 40 260 >>osc3f.dft
dft.exe osc.dat 5000 40 400 >>osc3f.dft
rem extra points to fill out peaks
dft.exe osc.dat 696250 3 0 >>osc3f.dft
dft.exe osc.dat 716250 3 0 >>osc3f.dft
dft.exe osc.dat 676250 3 0 >>osc3f.dft
rem
dft.exe osc.dat 697500 3 0 >>osc3f.dft
dft.exe osc.dat 717500 3 0 >>osc3f.dft
dft.exe osc.dat 677500 3 0 >>osc3f.dft
rem
dft.exe osc.dat 698750 3 0 >>osc3f.dft
dft.exe osc.dat 718750 3 0 >>osc3f.dft
dft.exe osc.dat 678750 3 0 >>osc3f.dft
rem
dft.exe osc.dat 2112500 3 0 >>osc3f.dft
dft.exe osc.dat 2072500 3 0 >>osc3f.dft
```

The sidebands are clearly visible on fundamental and harmonics. Note that the fundamental frequency has fallen between multiples of 5kHz—it is between 695 and 700kHz. The points at these frequencies are at approximately the same level, suggesting that the frequency is about 697.5kHz. The second harmonic (1395kHz) is thus very near a datum, and the peak at that frequency looks symmetrical.

Certain "extra" points have been computed by additional passes of the DFT routine. Note the extra points about the fundamental frequency, its harmonics, and the sidebands of the harmonics. This confirms the position of the frequency peak—a facility readily possible with the DFT-based (rather than FFT-based) approach.[6]

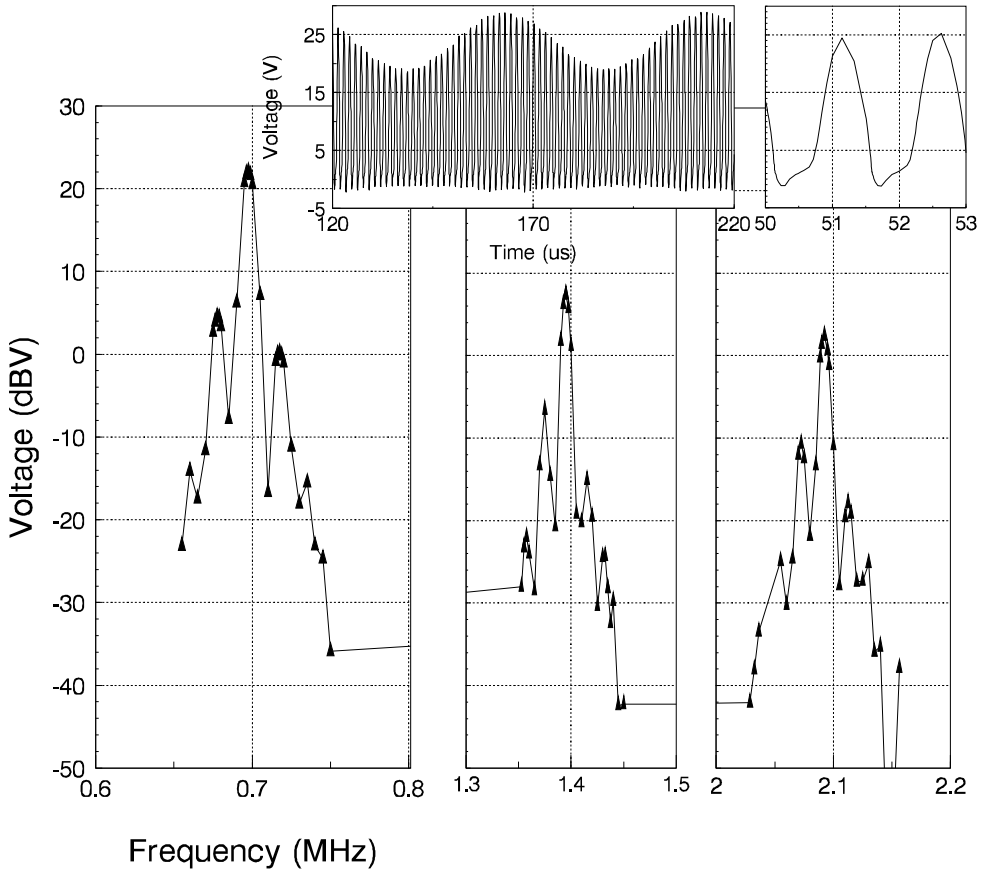

Figure 3.9: Spectrum obtained with DFT post-processor, including windowing, of the collector voltage of the circuit of figure 3.8. Insets show expansions of the time-domain data, including an expansion to show the shape of a single cycle.

[6]It should be noted that the data provided by an FFT, even limited solely to a number of points that is a power of 2, is sufficient to allow peaks in the spectrum to be located, by an appropriate interpolation process. However this is not so simple a process as is afforded by the use of a DFT whose frequency spacing may be arbitrarily chosen.

3.6 Application of DFT to Measured Time-voltage Data

Unsurprisingly, measured data resembles SPICE output. Note, however, that it is usually, rather than rarely, of an uncertain, inconvenient frequency. In many situations it is desirable to make a comparison between the measured and simulated results. A truly general spectrum analysis program should be equally applicable to both data sets. By way of demonstration, two spectrum analyses of measured signals are presented: One obtained using a modest DSO with 8-bit resolution and 4k sample length, the other using a sophisticated ADC subsystem boasting 110dB linearity, virtually unlimited data length, and output digitally decimated from a 10Ms/s, 23-bit ADC.

3.6.1 Measurement 1

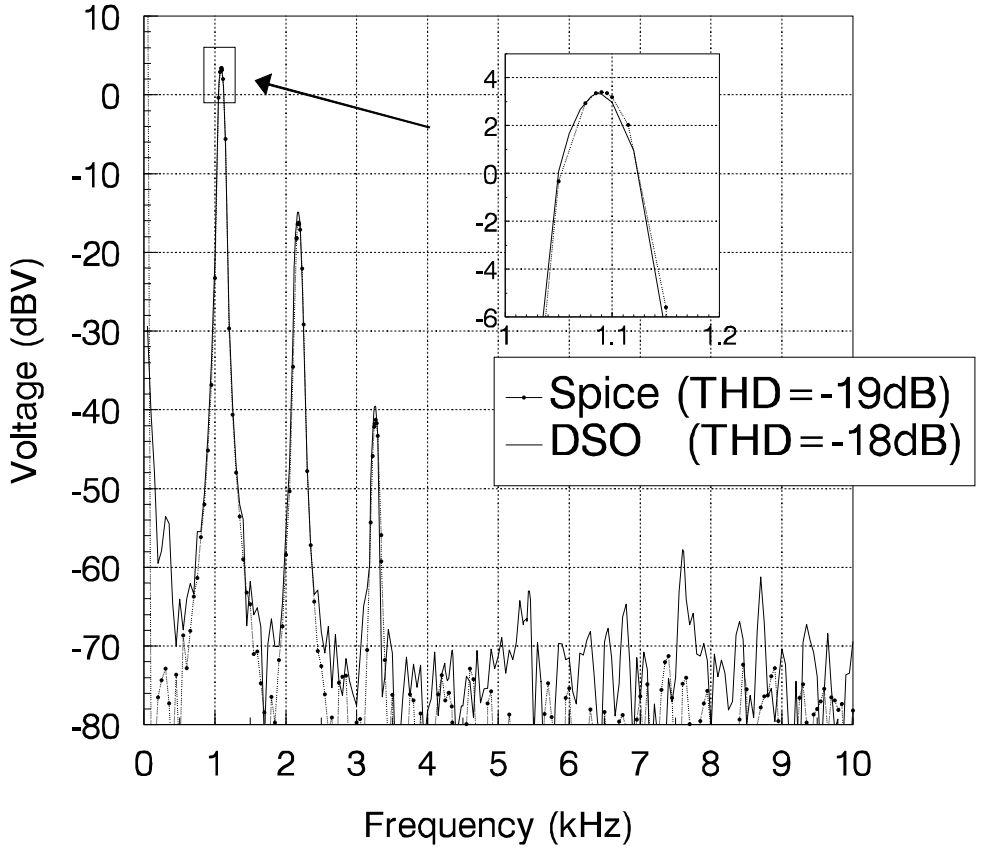

Figure 3.10: Spectrum of the output of a BJT amplifier computed with DFT from a DSO trace dump, and compared with similar SPICE output. The 'scope was set to 2ms/div, with about 5 divisions of deflection, and returned 4096 8-bit data points. The inset shows the fundamental peak, whose shape is set by the window function and the ≈20ms time record length.

Figure 3.10 shows data used to make a THD prediction and measurement on an a simple BJT CE amplifier. By way of comparison, an HP3561 Dynamic Signal Analyser measured the THD as -18.4dB, and an AWA F242A Noise and Distortion Meter read −18.6dB on the same signal, while SPICE predicts slightly below -19dB. The reliable dynamic range is of the order of 60dB, set by the linearity of the scope ADC. (i.e., distortions of the order of -60dBc are visible when analysing a known low-distortion signal.)

3.6.2 Measurement 2

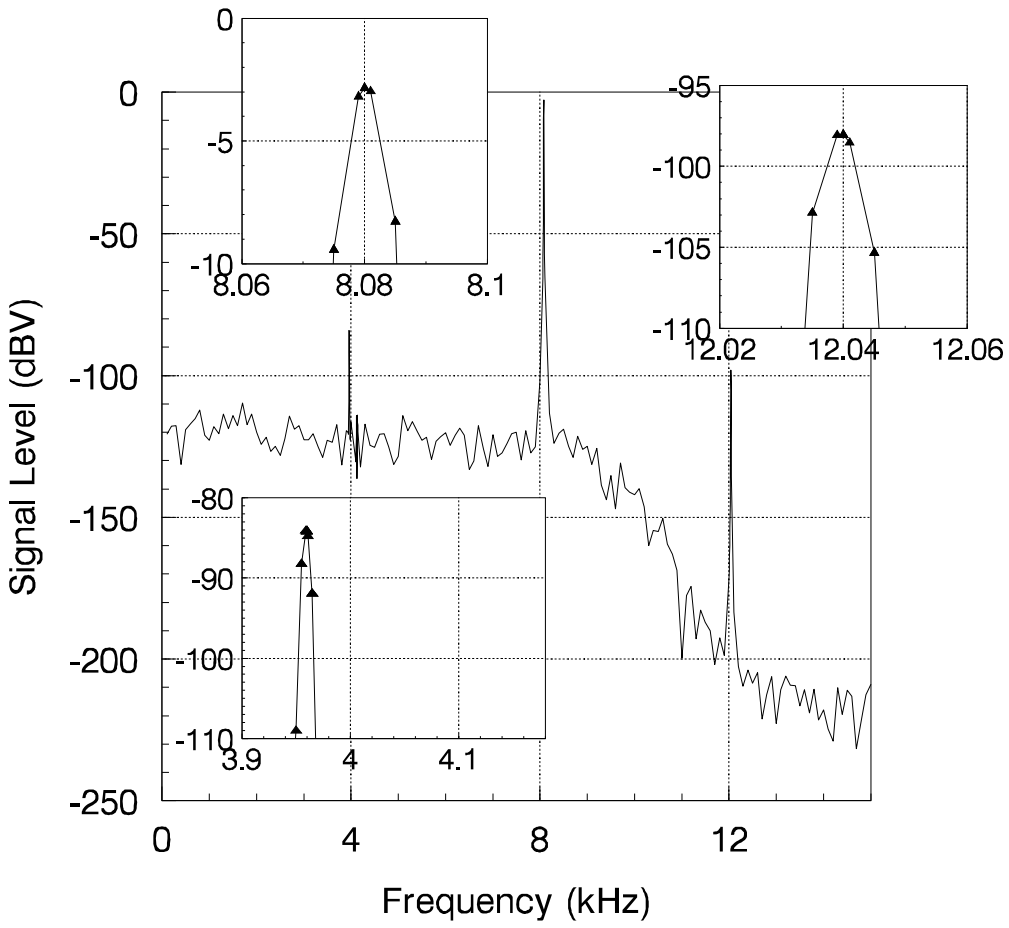

Figure 3.11: Analysis of 8192 samples of a $2V_{p\text{-}p}$ TDFD test signal. Data was originally sampled at 10Msps and decimated to 39ksps before analysis. Insets show relevant TDFD signal peaks at ≈3960, ≈4120, ≈8080 and ≈12040Hz.

Figure 3.11 shows a plot of the output from DFT used to perform a spectrum analysis of a TDFD signal acquired with a high-performance digitiser. The data window is slightly longer than 200ms, and has been windowed with a raised cosine. The broadening of the peaks resulting from this window is again clearly visible in the insets. Note the Y-axis scale, spanning 250dB, well beyond the capability of most spectrum analysers.

Since the residual components of the source near 3960Hz and 4120Hz are well below the -110dBc rating of the digitiser[22], the even-order component visible in the lower inset is attributed to the amplifier under test, giving a TDFD of -86dB. The overall dynamic range is slightly in excess of the 110dB claimed for the digitiser.[23] The noise floor has been

obviously shaped by the response of the decimation filters, and the 12040Hz component attenuated by over 90dB.

3.7 Complex Circuit Simulation Example

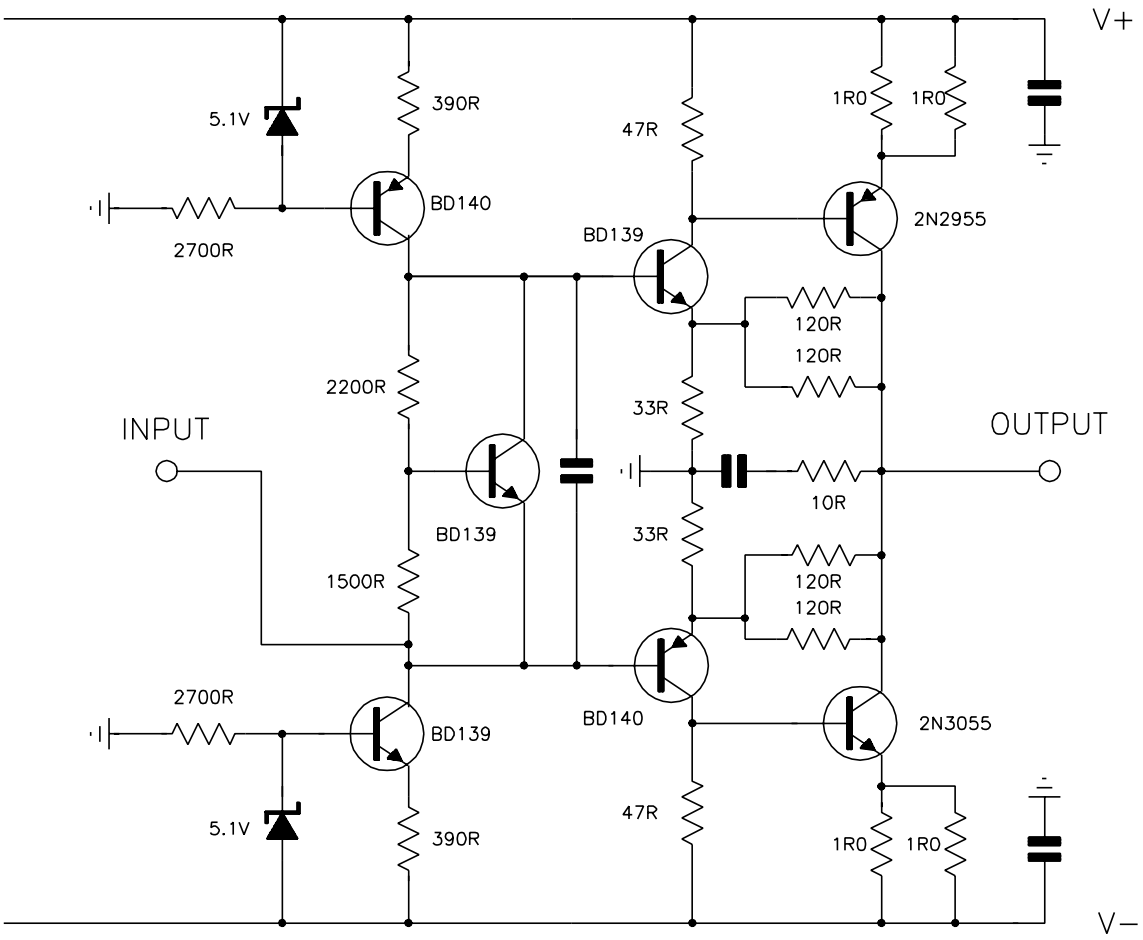

Figure 3.12: Circuit of an example amplifier using bipolar junction transistors (BJTs).

Consider the circuit of the amplifier shown in figure 3.12. This circuit is that of a "uncomplicated" amplifier designed to boost the output of an operational amplifier sufficiently to deliver 50 Watts into a nominal 8Ω load with ±35V supplies. Despite its relative simplicity, no reasonable means of estimating the distortion that might be encountered in operation can be developed from the literature. (Indeed, even the calculations of small-signal impedances are not trivial.) However, a few minutes work using SPICE and DFT yields the results shown in figures 3.13, 3.14 and 3.15.

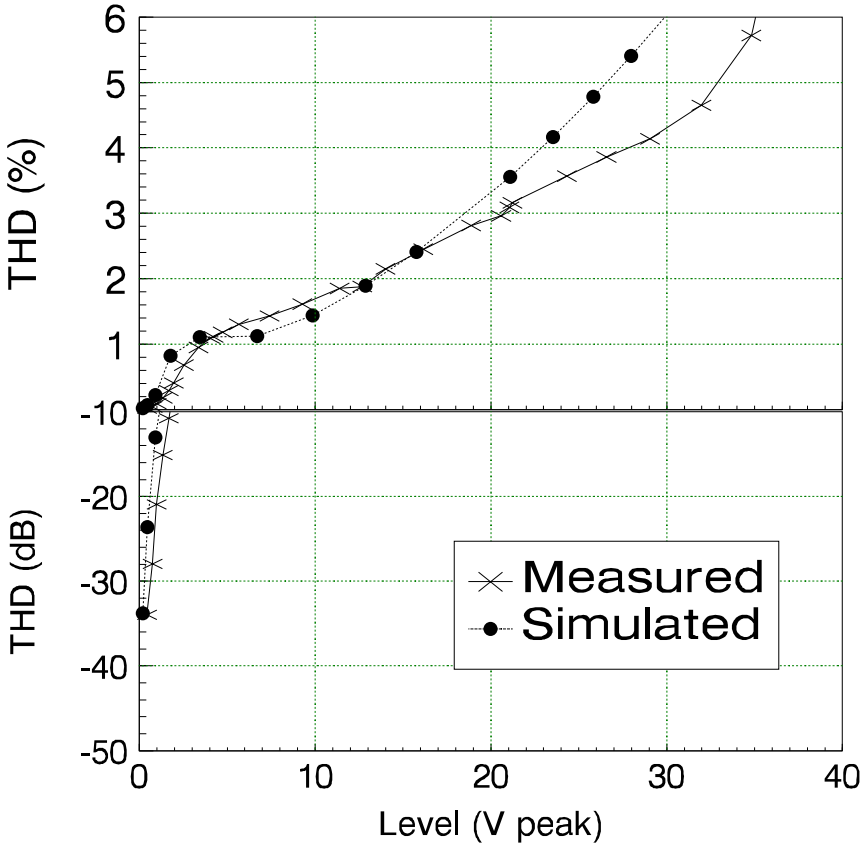

Figure 3.13: Comparison of simulated and measured THD as a function of output level for the circuit of figure 3.12. Note that the vertical scale is linear (%) in the top half of the graph and log (dB) in the lower for clarity.

Agreement between simulated and measured results may be regarded as excellent. In the case of total harmonic distortion (THD) for varying output level, both the unusual trends and the absolute values are predicted with greater certainty than might be expected with some measurement techniques and more repeatability than might be expected with construction technique. (The increasing deviation for levels in excess of 20 volts peak is a consequence of modelling the unregulated supplies as perfect sources, betrayed by mains-frequency sidebands appearing on tones.) The spectrum of the amplifier driven with a single sinewave, as shown in figures 3.14 and 3.15, is predicted with admirable fidelity: discrepancies are typically only significant for components more than two orders of magnitude below the fundamental. The only special care taken to obtain these results

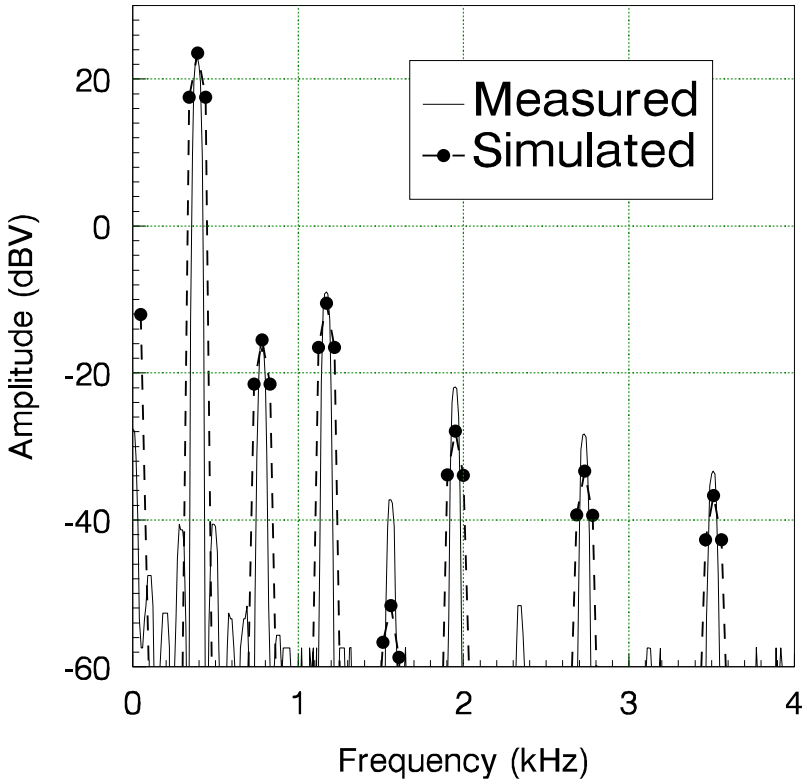

Figure 3.14: Comparison of a simulated and measured output spectrum for the circuit of figure 3.12 delivering ≈20V peak into 8Ω.

was the measurement of transistor current gain (β or h_{FE}), which was set approximately in the four device model cards by means of the BETA and IKF model parameters.

3.8 Conclusion

The DFT program offers several advantages:

1. It is typically much faster that using an FFT, since far fewer points are needed for distortion analysis than are yielded from an FFT;
2. It can deliver frequency data points at any frequency;
3. It optimally handles non-equispaced time domain data.

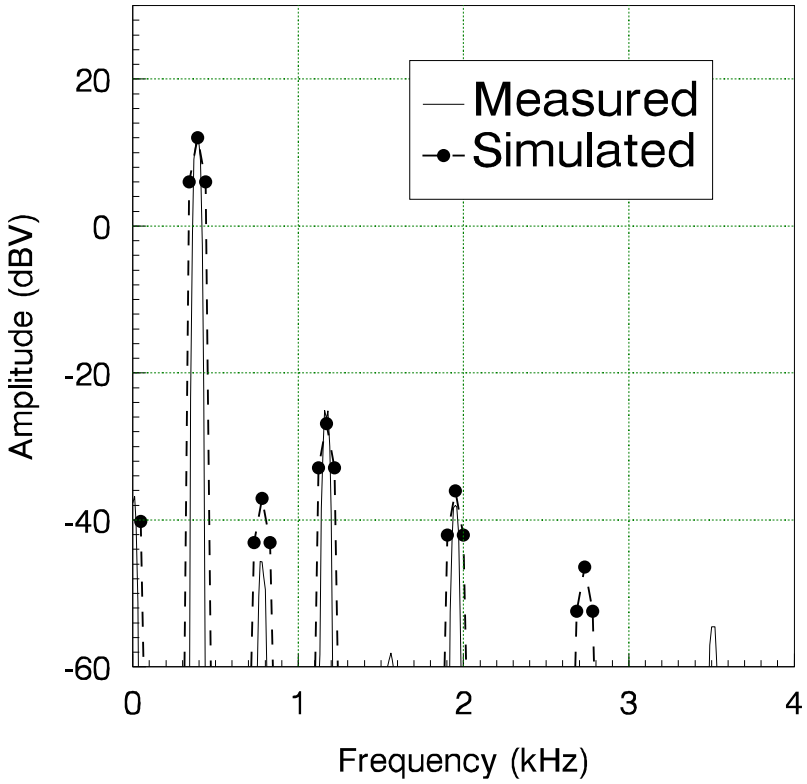

Figure 3.15: Comparison of a simulated and measured output spectrum for the circuit of figure 3.12 delivering ≈5V peak into 8Ω.

Examples of both distortion prediction and measurement have been given, using SPICE and digitised time series. In the next chapter, DFT will be used as a critical component of a new method for development and verification of a low-signal loudspeaker driver model.

Chapter 4

Characterising Loudspeaker Driver Nonlinearity

Considerable interest has been shown in the literature recently toward the problems of linearity in moving-coil loudspeaker drivers.[24]–[30] Some models address only nonlinearity in the Bl[1] product.[28] In [24], Kaizer presents an overall, Volterra model fitted to harmonic and intermodulation measurements (in addition to straightforward linear measurements). This model produces "reasonable agreement" between calculated and measured distortion in the output signal of an accelerometer added to the driver cone. In [26], Klippel extends Kaizer's model, and identifies parameter measurement as the principal obstacle. Again acoustic harmonic and intermodulation measurements are used to determine parameters. Most work has concentrated upon the determination of nonlinearity in the transfer function[2] although the fact that the impedance must be non-linear is implicit in much of the work, particularly with respect to current drive.[25]

In these instances, parameters are extracted by a numerical error-minimisation process

[1]The current-force conversion factor is universally referred to as "Bl product", pronounced "bee-ell product".

[2]For an electro-acoustic transducer this is sound pressure level (SPL) divided by input voltage or input current.

designed to select the parameter vector that best permits the model to predict the observed (non-linear) behaviour. Beyond any concerns associated with pitfalls in achieving optimisation with many dimensions, the authors are presented with a difficulty: checking the validity of a model whose parameters are selected using the very behaviour the model is attempting to predict. Almost any model, no matter how uncorrelated with the physical reality of the system it models, can achieve the feat of predicting that behaviour upon which it is optimised. This is not a problem unique to loudspeaker drivers![3]

The addition of the accelerometer by Kaizer provides a different observable phenomenon by which to check the model's prediction with measurements not used in its fitting. While truly dynamic position sensing may be applied to the cone of a driver, complex equipment is required. The simple accelerometer may disturb the driver and in any case provides a signal removed from position by a second derivative. This signal is not far removed from the spectra to which the model is fitted.

Most recently, Olsen has approached the measurement of nonlinearity in the mechanical components by making mechanical measurements.[30] This approach avoids pitfalls of previous methods. Apparently carried out simultaneously with this work, it resembles the work reported here in several ways. However, it requires special equipment that must be carefully mounted to the driver under test using glue. Further, no verification of the extracted functions has been reported so far, and it is thus unclear how effective is the method.

Results obtained to date, often by the cited authors' own admissions, have not been spectacular. Shortcomings are typically excused because the "model breaks down at higher frequencies due to the nature of hysteresis effects, eddy currents and cone breakup effects".

This chapter presents:

[3] An example is the wide variety of models of III-V semiconductor devices, whose physical workings are not well enough understood to permit a single, meaningful, physically realistic model to be developed.[31]

1. A physically-based, behavioural, driver model incorporating Bl, inductance and compliance nonlinearity effects, and which could readily be expanded to incorporate heating effects in the coil ohmic resistance, R_e.[29]

2. A procedure to fit the model to a driver employing solely linear parameters[4] and simple, mechanical measurements, involving neither nonlinear acoustic nor nonlinear impedance measurements;

3. Verification of the equivalent, nonlinear, circuit model and curve-fitting procedure by demonstrating good agreement between prediction and measurement of the acoustic and current distortion of the driver.

Because the model requires only linear parameters and simple mechanical measurements, it may be obtained for a driver using methods already at the disposal of most laboratories. Since it is physically-based, it provides direct insight into the causes of distortion within the driver. It can predict nonlinear distortion both in the acoustic output and in the current drawn by the driver.

4.1 Method

The basic idea is to determine the linear equivalent circuit of a driver at each position in a range of cone displacements, and to deduce the nonlinear behaviour from the variation in linear behaviour. The varying cone displacements are achieved by passing a direct current through the voice coil. The linear parameters at each position may be determined from measurements of small-signal impedance. Once the variation of the linear circuit elements with instantaneous cone position is known, the driver's dynamic, nonlinear performance may be simulated. The effectiveness of the characterisation can be checked by comparing predicted distortion (in either the acoustic output or the current into the driver terminals or both) with measurements made using dynamic distortion measurement equipment such as a spectrum analyser. Because there are no measurements common to the processes

[4]The linear parameters are related to, but are not, Thiele-Small (T-S) parameters. They will be referred to as "T-S" parameters to honour their inspiration.

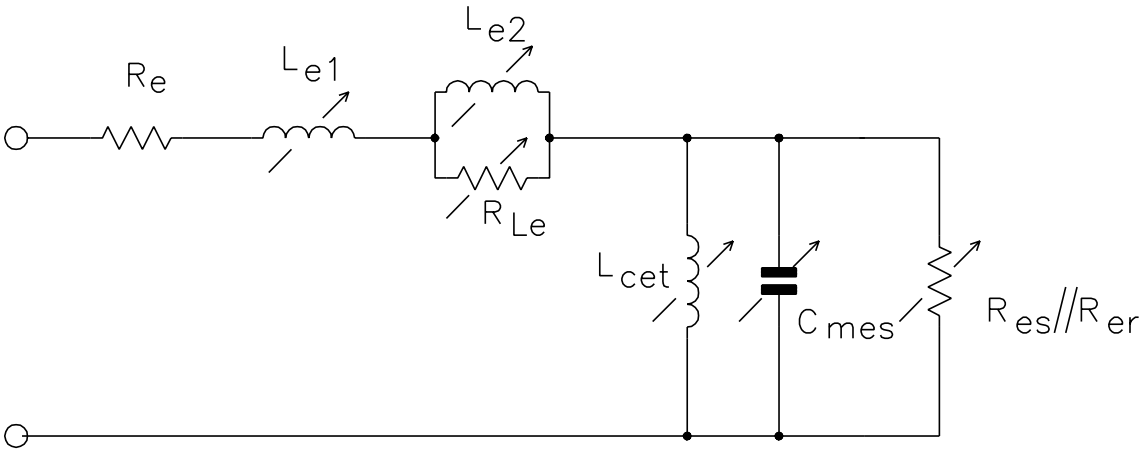

Figure 4.1: Equivalent nonlinear circuit of the driver.

of fitting and verifying the model, the check becomes a very powerful measure of the effectiveness of the method.

4.2 The Nonlinear Model

The familiar series-parallel resonant circuit topology is extended into a nonlinear model, retaining the twenty five years of intuition it has cultivated in this field. Figure 4.1 shows the equivalent circuit used. It strongly resembles the traditional model—a parallel resonant part, a series resistance R_e, and a series inductance—but includes a minor extension suggested by Wright.[32]. This is the shunted/unshunted series pair for voice-coil self-inductance. However, in this case, all but R_e become functions of cone displacement, x. In figure 4.1, the non-linearity of the elements may be expressed by reporting the element values as functions of variables (voltages and currents) in the circuit: this form is particularly suitable to behavioural modelling using modern circuit simulators such as Berkeley's SPICE 3F2 and MicroSim's PSPICE.[15, 14] This will be seen in section 4.5.

Kaizer [24] has correctly stated that if analysis is restricted to adequately low frequencies, then the dominant sources of non-linearity are functions of cone displacement: namely force factor, Bl, coil self inductance, L_e, compliance, C_{ms}, and suspension losses, R_{ms}.

These four elements are increased by two in number, as the Wright extension splits L_e into L_{e1}, L_{e2}, and R_{Le}.

Small [34] gives the definitions of the linear mechanical parameters. Making them non-linear functions of x:

$$R_{et}(x) = R_{es}(x) \parallel R_{er}(x) = \frac{1}{R_{ms}(x)Bl(x)^2} \tag{4.1}$$

where $R_{ms}(x)$ is a non-linear function representing the suspension's mechanical resistance,

$$L_{cet}(x) = Bl(x)^2 C_{mt}(x) \tag{4.2}$$

where $C_{mt}(x)$ represents the mechanical compliance of the driver suspension and box,

$$C_{mes}(x) = \frac{M_{ms}}{Bl(x)^2} \tag{4.3}$$

where M_{ms} is the constant mechanical mass of the driver diaphragm assembly, including air load. With the components represented as non-linear functions of x the model embodies nonlinear compliance and suspension losses, with the nonlinear Bl factor underlying all of these components. Cone displacement itself is given by

$$x = \int (V_{\text{vel}}/Bl(x))\delta t \tag{4.4}$$

where x is the cone position, V_{vel} is the voltage across the parallel tuned circuit in the driver equivalent circuit, and $Bl(x)$ is the instantaneous value of the Bl product.

The problem is now to obtain the six nonlinear functions.[5] Values of them, for any single displacement, can be obtained experimentally by measuring the driver's impedance versus frequency and numerically solving for component values by the least-squares method. The displacement x is varied by applying a dc current, and the non-linear trend of each component with dc current obtained. By the additional measurement of cone position versus applied dc current, the nonlinearities may be expressed as functions of x. Also, by measuring the force factor at cone equilibrium position, $Bl(0)$, $Bl(x)$ can be derived from $C_{mes}(x)$, and the driver parameters separated and expressed in mechanical or electrical terms.

[5]The Bl function is readily found from equation 4.3, and is not itself an independent function since M_{ms} is a constant found from small-signal measurements.

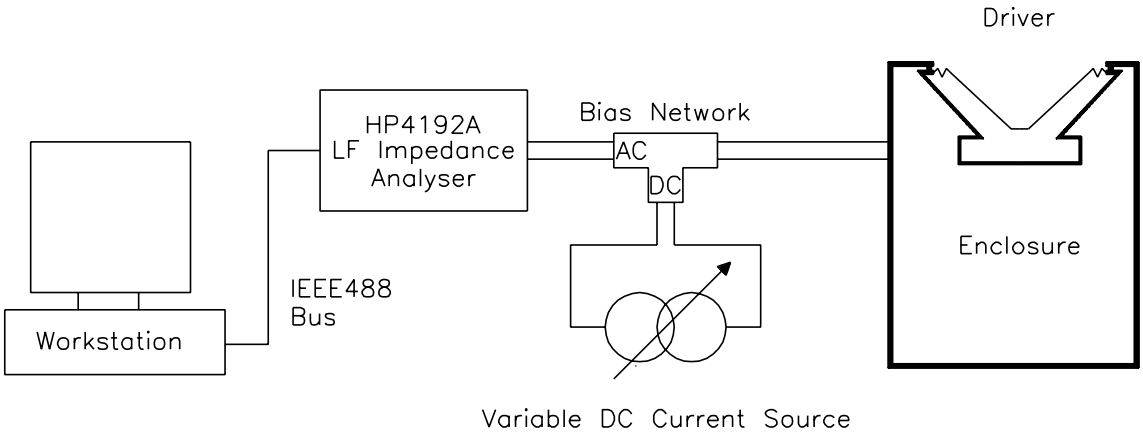

Figure 4.2: Diagram of equipment used to measure driver impedance for various dc coil currents.

The six functions are approximated using polynomials. These may be fitted to the tables of values for various displacements by straightforward mathematical methods.[35]

The validity of this method rests on the assumption that the electrical properties do not vary significantly with drive current (i.e., the Bl product is not significantly affected by the dc displacement-control current nor any signal current). The approach also relies on the assumption that the mechanical properties, including suspension-related ones, are not significantly frequency-dependent. With some drivers, especially those employing cloth and/or paper spiders[6], this assumption appears not to hold. Attempts to fit the model to such drivers did not yield sound predictions of nonlinear behaviour.

4.3 Driver Characterisation

The characterisation of a driver involves two stages: measurement of the physical properties of the driver, and extraction of the parameters from the data. The measurements are detailed in sections 4.3.1–4.3.3, and the extraction in section 4.3.4.

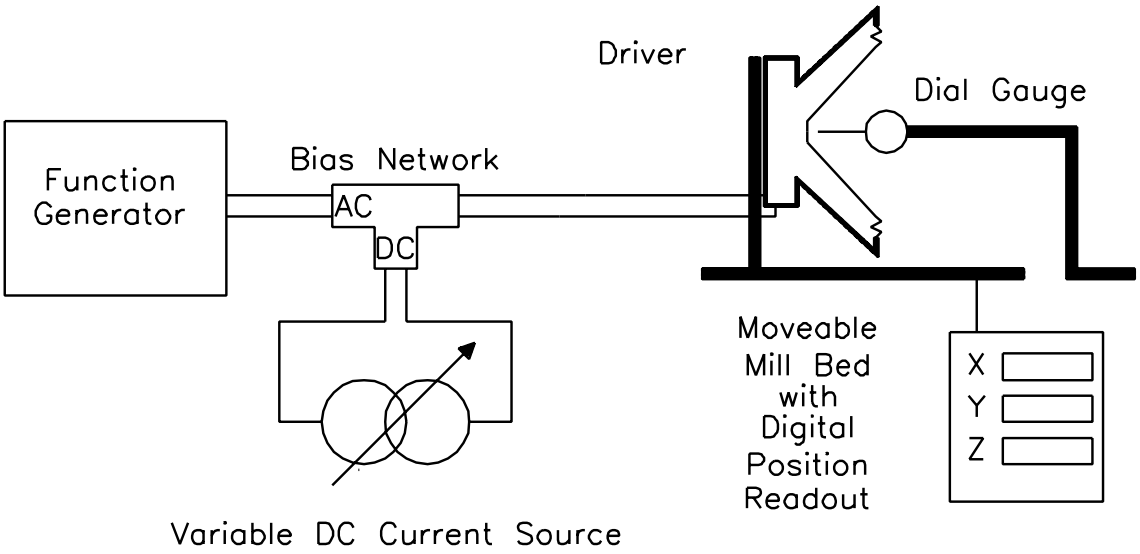

Figure 4.3: Diagram of setup for measuring cone displacement against dc coil current.

4.3.1 Impedance Sweeps

Figure 4.2 shows the equipment used to extract the impedance curves. It consists of an HP4192A LF Impedance Analyser, controlled via IEEE488 bus by a computer. The HP4192A utilises a four terminal method to measure vector impedance, and can be made to interface with an external dc current source easily. A current source capable of sourcing or sinking up to 1.5A is straightforwardly made. These signals were used to drive the test driver in a non-ported enclosure.[7]

Care may have to be taken with some drivers in making these measurements, in order to prevent suspension hysteresis effects interfering with results. The cone must be settled into its displaced equilibrium position prior to impedance measurement.

4.3.2 Current-displacement Curve

The cone's displacement due to dc current was measured using an numerically-controlled (NC) milling machine with digital position readout, as shown in figure 4.3. After setting the dc current to the required value, the bed of the milling machine was extended until

[6]The spider is the main mechanism attaching, supporting and positioning the cone.
[7]This was done to simplify driver mounting and to facilitate acoustic distortion measurements.

contact was made with the cone. By superimposing a small ac signal onto the dc current injected into the driver, the cone gives an audible buzz on contact with a whisker or the feeler of a dial gauge attached to the chassis, facilitating the measurement of displacement. Again hysteresis effects may require that care is taken with the measurements, especially with larger drivers.

4.3.3 Zero-displacement Bl

The measurement of the driver's force factor at zero cone displacement was achieved by a simple "balance of forces" method. Firstly, with the driver's axis horizontal to prevent the cone sagging under gravity, a whisker attached to the driver chassis is positioned so as to just touch the cone. Then the driver is positioned vertically, and the cone weighted with a series of known masses. The dc current required to return the cone to its equilibrium position is noted. By plotting the added mass versus applied dc current, $Bl(0)$ can easily be found from the gradient.

4.3.4 Computation

The first step in calculation is the extraction of box-modified T-S parameters for each impedance sweep. This may be accomplished by straightforward numerical optimisation (minimisation of mean square error), using a general-purpose implementation of the simplex algorithm such as that embodied in Microsoft Excel's Solver. Values of M_{ms} and R_e are determined.

R_e is treated as constant; dc measurements suggested that its variation was negligibly small in the case of the drivers that have been examined. In the event that a driver exhibits significant variation of R_e with voice coil temperature, the method can still be expected to work, provided the time constant is long in comparison to the signal period. Only with slowly varying music program content would changes occur on the timescale of measurement.

Weighting of the data was achieved not by varying the weight of any individual data point but by varying the point density in different regions, chiefly increasing the number of measurements near the parallel resonance. After this extraction, one constructs data files containing the value of various circuit components tabulated against dc displacement current.

The next step involves fitting polynomials of appropriate order to various data sets. First a fit is made to the current-displacement data. This is used to replace the value of dc current associated with each sweep with its correct cone displacement. Subsequently an approximation polynomial is fitted to each component value, and to $Bl(x)$. Thereafter, each is associated with a set of coefficients relating its value to displacement. (These are the model parameters.) The order of the polynomial used in each regression was selected by fitting a range of orders and finally selecting the highest order above which the coefficients became negligible, or above which ripples became visible in the range of interest. Experience suggests that seventh order is generally quite adequate, and hence measurements should be carried out for about twenty values of displacement in order to provide an adequate number of samples for the regression. For the polynomial fitting and plotting 3D-Vision's Graftool was used.

Once the parameters have been determined, a suitable behavioural modelling tool (such as PSPICE) can be used to determine predictions of non-linear cone displacement, sound pressure level (SPL), and driver terminal current for any arbitrary drive signal. Acoustic distortion is found by using DFT on the predicted SPL data.

4.4 Example

An example characterisation was carried out on a Peerless 831921 200mm driver. This driver has a modest power rating and is made with a polypropylene cone with foam surround.

4.4.1 Measurements

Figure 4.4 contains three plots associated with the measurements described in sections 4.3.1–4.3.3. It provides a visual summary of the data used in fitting the model to the driver.

4.4.2 Computed Functions

Figure 4.5 contains plots of the nonlinear functions computed from the data presented in figure 4.4. This figure provides a visual summary of the driver component nonlinearities. If the object of the modelling is to determine the nonlinearity of one or more driver parts, this figure represents the desired output; i.e., this figure displays the functions representing the nonlinear contribution of the various components of the model.

4.5 Verification of Nonlinear Parameters

The drive-current and acoustic distortion can be predicted using the nonlinear model implemented as a behavioral model in SPICE. Figure 4.6 gives the equivalent circuit as implemented in SPICE. The effective formulae for current in two of the inductors are shown for example. Note that values for Bl and cone position are computed by the subsidiary loops at the bottom of the circuit using nonlinear voltage-controlled voltage sources: Position is computed from velocity and Bl-product, and Bl-product from position.

Predictions made with this model may be compared with measurements. It is this comparison that provides independent verification of the validity of the process and the success of the measurements.

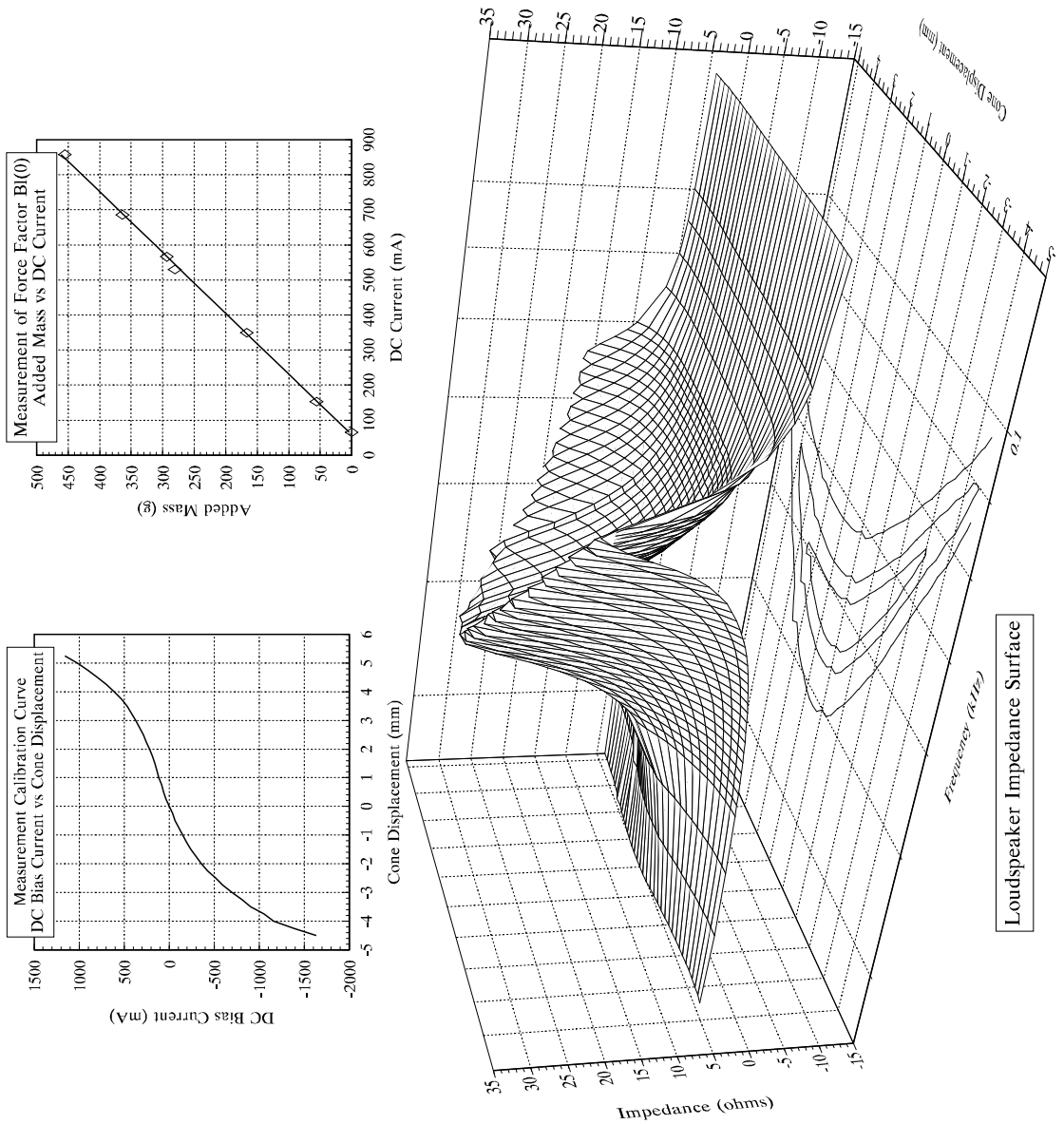

Figure 4.4: Plot of data used for parameter extraction of the Peerless 831921 driver. (Impedance data has been decimated for clarity.)

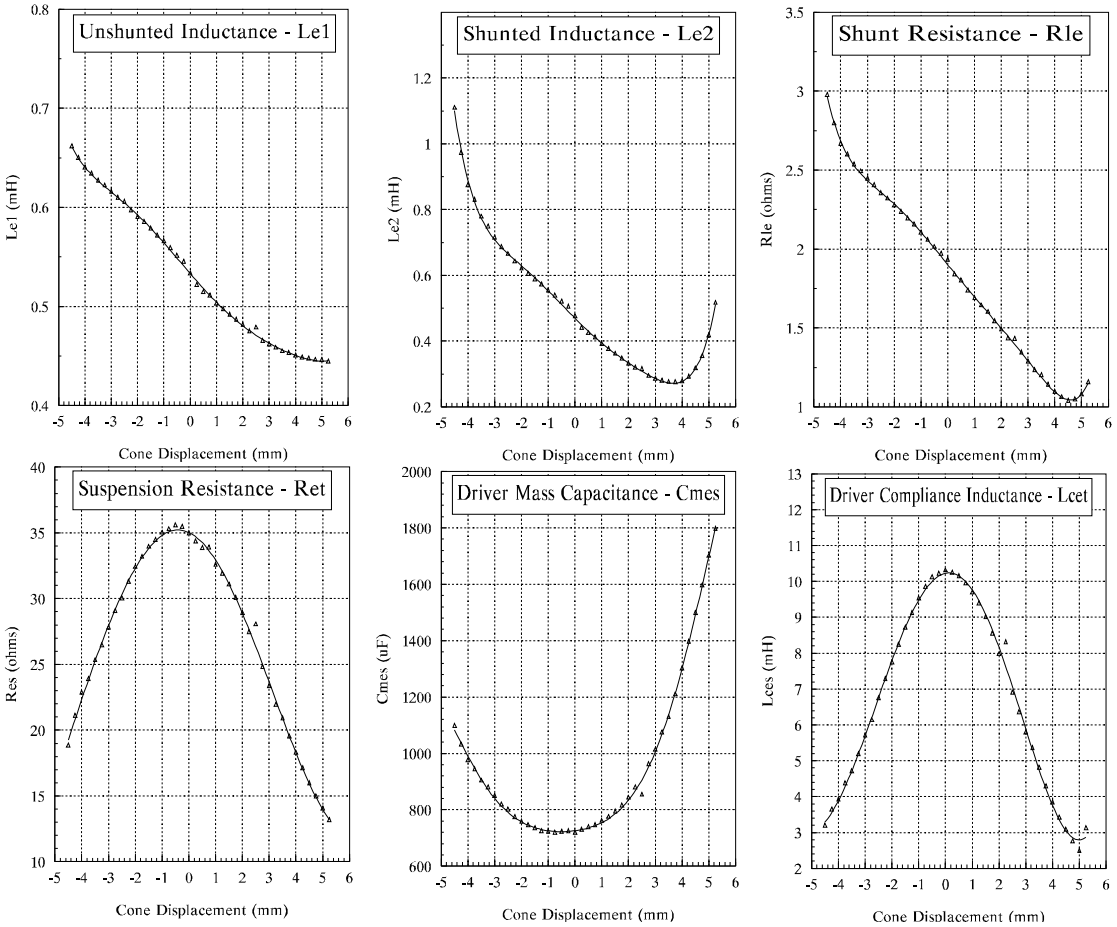

Figure 4.5: Plots of the nonlinear parameters extracted for the 831921 driver: L_{e1}, L_{e2}, R_{Le}, R_{et}, C_{mes} and L_{cet} against cone displacement. Note that R_{et} and L_{cet} are box-modified equivalent parameters.

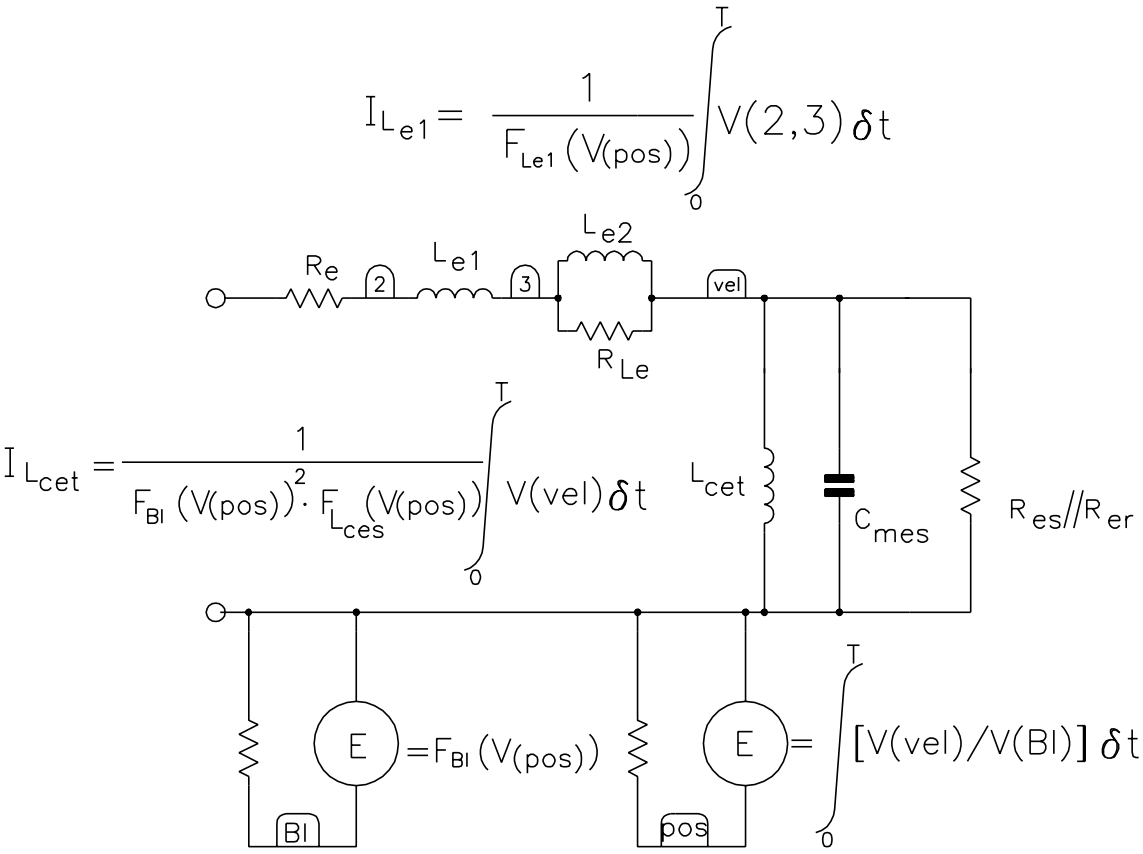

Figure 4.6: Behavioural model of the nonlinear driver as implemented in SPICE.

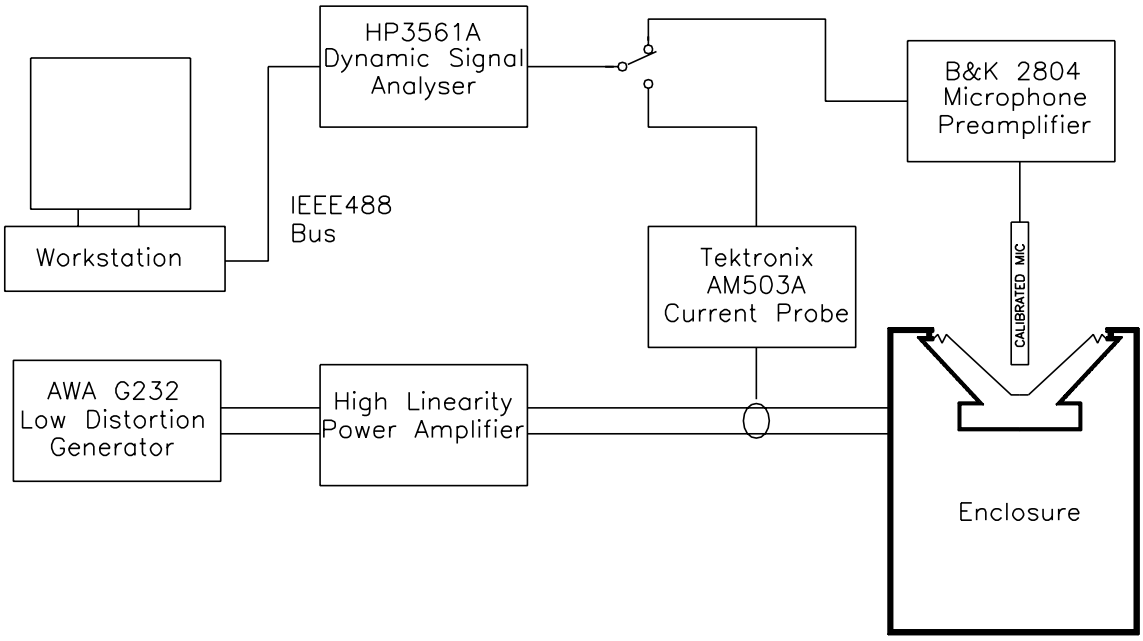

Figure 4.7: Block diagram of the equipment setup for measuring driver acoustic and current distortion.

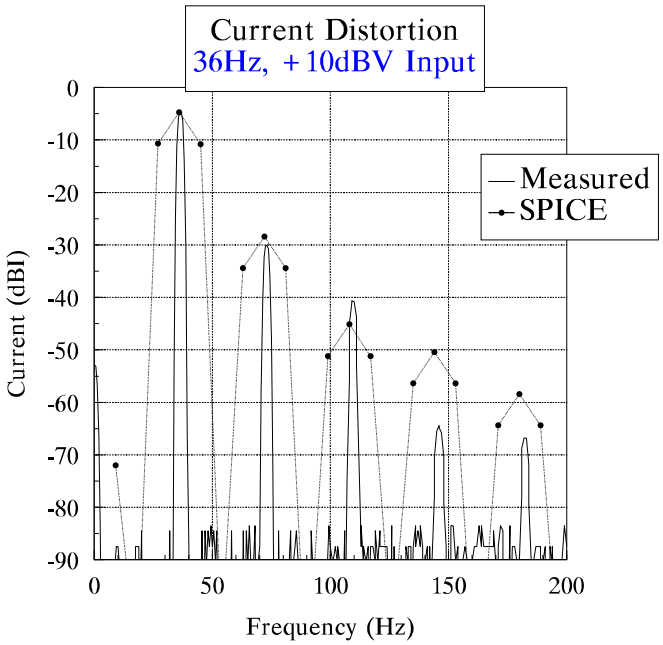

Figure 4.8: Comparison of predicted and measured driver current distortion at 36Hz.

Frequency	Drive Level	Predicted SPL THD	Measured SPL THD
36Hz	-5dBV	-23dB	-26dB
54Hz	+10dBV	-29dB	-31dB

Table 4.1: Predicted and measured SPL THD.

4.5.1 Distortion Measurements

Figure 4.7 depicts the equipment setup used to measure distortion in both driver current and the driver acoustic near field.[36] Figures 4.8 and 4.9 display typical comparisons of predicted and measured current spectra. Table 4.1 tabulates values of predicted and measured SPL THD, and table 4.2 values of drive current THD. The larger discrepancies at the higher frequency of 238Hz are attributed to the relatively small cone excursion involved.

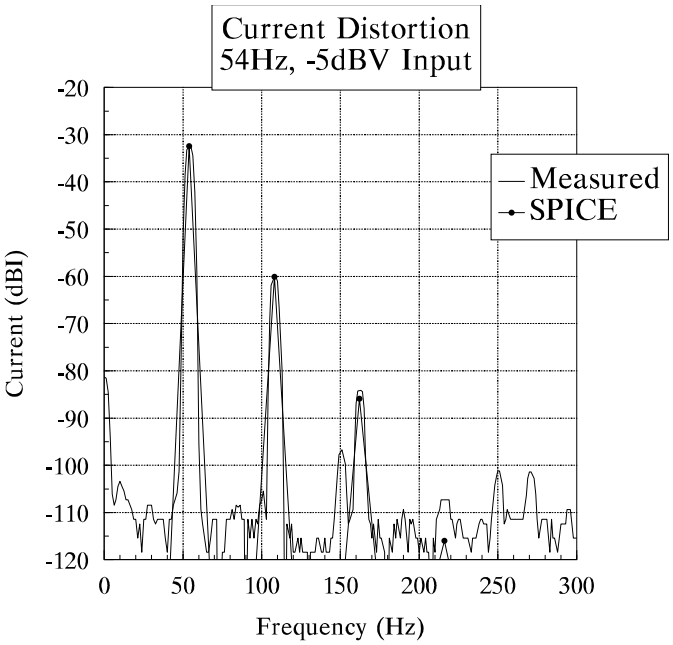

Figure 4.9: Comparison of predicted and measured driver current distortion at 54Hz.

Frequency	Drive Level	Predicted SPL THD	Measured SPL THD
36Hz	+10dBV	-24dB	-25dB
36Hz	-5dBV	-38dB	-37dB
36Hz	-20dBV	-53dB	-52dB
54Hz	+10dBV	-13dB	-13dB
54Hz	-5dBV	-28dB	-28dB
238Hz	+10dBV	-55dB	-41dB
238Hz	-5dBV	-63dB	-44dB

Table 4.2: Predicted and measured driver current THD.

4.6 Conclusion

A new nonlinear driver model has been presented. A novel method of obtaining model parameter values has been devised. This method is more readily accessible to industry. Agreement, superior to that previously reported, was obtained between measured acoustic and impedance distortion. The work reported in this chapter can be expected to provide researchers with a powerful tool for the design of more linear loudspeaker drivers, in the form of straightforward assessment of driver components.

The method is eminently practical, in contradistinction to the conventional analysis methods reported in texts such as [8]. The technique is being used by Lorantz Audio Services, an Australian manufacturer and exporter of high-end loudspeaker drivers.

This chapter hs used the low-signal analysis too "DFT" in conjuntion with a low-signal model and a new fitting method to demonstrate successful prediction of low-signal distortion. In the next chapter, the measurement of small nonlinearities is addressed.

Chapter 5

A Total Difference-Frequency Distortion Meter

Non-linear distortion of a system, typically an amplifier, is customarily measured with the total harmonic distortion (THD) method. As noted in chapter 2, although doubtful that distortion can be meaningfully characterised with any single real number, THD is widely used for this purpose. Total difference-frequency distortion (TDFD) has been proposed as a better measure.[37] A design for a TDFD meter has appeared in the literature.[38] Results obtained with this meter are quite impressive, and TDFD instruments of this design are in use.

Although TDFD promises improvement both by virtue of its test signals being placed at the high end of the system passband, and by virtue of its measurement bandwidth being much smaller, this existing design is demanding. Little more has been reported in the literature, and there is no commercialisation of the design, despite at least one attempt. Evidently the advantages of TDFD as a measure have not to date outweighed the difficulties associated with the development of a meter, as no commercial meter has appeared in the marketplace.

This chapter describes a new and superior TDFD meter. It boasts an unequalled 135dB

dynamic range, requires no critical filters as with THD and previous TDFD meters, overcomes limitations associated with previous TDFD meters, and is fast and easy to use.

5.1 Motivation

The question arises as to what motivated the design of a meter this sensitive. Engineers are presently managing with systems that are at least 20dB less sensitive, while most commercial systems have performance 40dB below what this meter offers.

Design to any performance level requires the ability to measure even better performance. The greatest need for design of high-linearity circuits is a meter that can measure very small distortion levels. Further, understanding low-signal operation of circuits follows from understanding the curvature of the characteristics of semiconductor devices. Information about small characteristic nonlinearities can be inferred from sensitive distortion measurements.

Another use is the measurement of non-linearity in current-mode circuits, which are presently finding some popularity. Current is typically sensed by looking at voltage dropped across a (small) series resistance. The series resistance is kept small so as not to disturb the circuit significantly. When non-linearity is present in a current signal, the magnitude of the non-linear components in the voltage dropped across the current sense resistor can be very small indeed. The sensitivity of the TDFD meter can be used to allow such measurements while keeping the sense resistor small, and the circuit undisturbed.

It was probably once asked "what use is a DMM with 4-plus digit resolution when power supplies are only required to be accurate to 5%?". The scepticism of that question has been denied by the proliferation of such DMMs, and the presence of 8 or more digit ones, such as the HP3458A. Engineers discovered that the resolution of a DMM could be used to see small voltage drops along a supply track on a PCB, leading to locating, say, a logic IC with an output pin shorted to ground. Other applications have added to that list with

time. It seems likely that uses for meters with this sensitivity will be found in time, if such meters are available.

5.2 Potential of TDFD

To measure TDFD, a system is excited with two signals at $2f_0$ and $3f_0 + \delta$. TDFD is defined in terms of the amplitudes of the two exciting signals, and the amplitudes of signals which appear at $f_0 + \delta$ and $f_0 - \delta$. Typically $f_0 \approx 4$kHz and $\delta \approx 40$Hz; thus the test frequencies will be near 8kHz and 12kHz, and the results of interest near 4kHz. Now, TDFD offers two main advantages over THD, and each of these advantages has an appeal both to the person using the TDFD figure, and to the engineer designing an instrument to determine it.

The first advantage is that a TDFD test subjects the Device Under Test (DUT) to standard signals at the high end of the passband region.[1] THD measurements which allow for a reasonable number of harmonics (three might be an absolute minimum and ten is often desirable) are forced to apply the stimulus at a frequency which is a submultiple of the top frequency in the band. THD is often measured at 1kHz. Thus the slew rate demanded of the DUT in such a test does not approach what may be demanded in actual operation with music. If the THD measurement is carried out at a higher frequency, fewer harmonics are within the passband. Results may be different at different frequencies.

To the buyer of products (whose distortion has been measured and stated), TDFD can be expected to betray problems across the band; further he does not need to check on or interpret results in the light of a test frequency figure. To the instrument designer, TDFD does not need variable test frequencies, and so the meter will be simpler to design and build.

This second advantage is that the signals of interest are collected in a narrow region. A TDFD system must look only for signals close to 4kHz, whereas a THD measure must

[1]The frequencies chosen allow TDFD standards useful with 15kHz broadcast systems as well as high fidelity systems with 20kHz bandwidths.

collect information in a bandwidth many times the test signal frequency. TDFD distortion product signals can be separated electronically from the copies of the stimulus signals and measured with relative ease, and in a situation where the cumulative noise can be much lower.

To both the buyer and instrument designer, this means that TDFD meters will typically be able to see smaller imperfections for a given complexity and cost of measuring apparatus, and smaller imperfections in the ultimate limit.

5.3 Conventional TDFD Meters

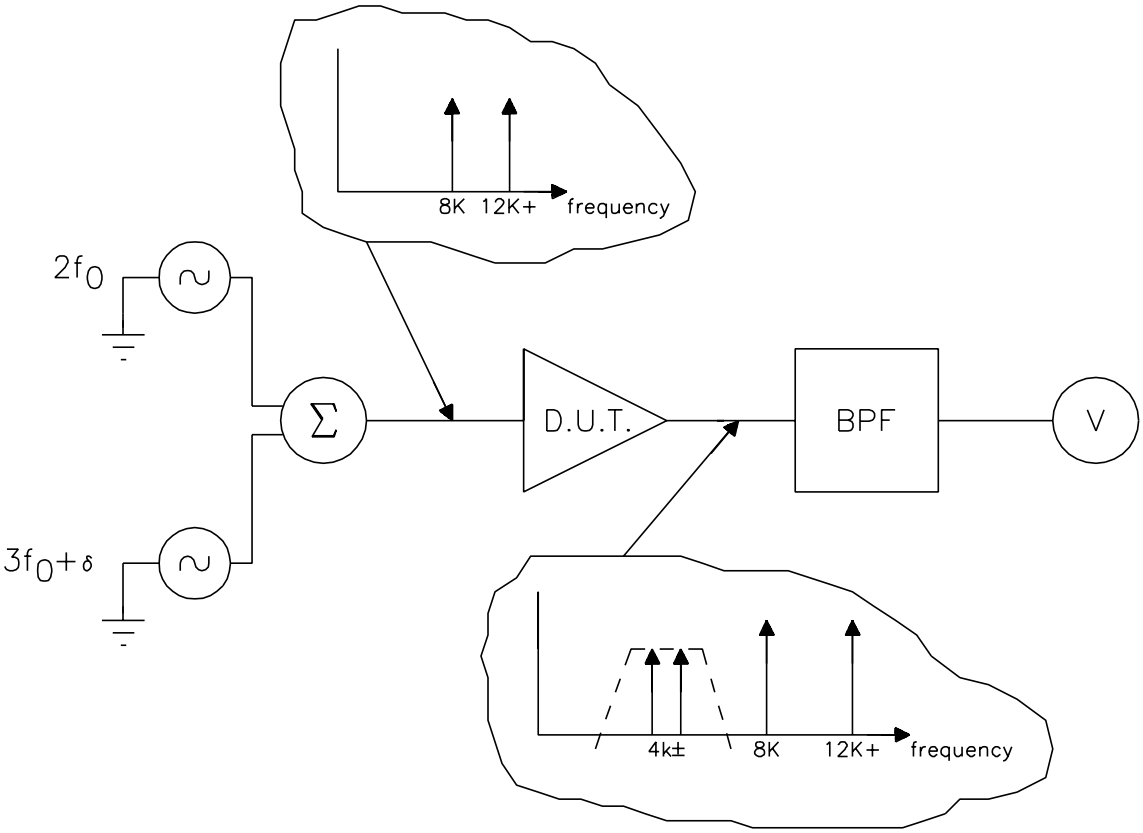

Figure 5.1: Block diagram of a conventional TDFD meter showing the signals present.

Figure 5.1 shows the block schematic layout of a conventional TDFD meter. Two signals are added together, passed through the DUT and then this output signal is filtered to isolate the components of interest about f_0, typically 4kHz. This design requires a tight bandpass filter (BPF). To achieve a given sensitivity the BPF must attenuate test signals

one octave above the signals of interest by the required dynamic sensitivity, without disturbing these wanted frequencies. In the published prototype -120dB of dynamic range was achieved using high-Q filters and crystal-referenced signals with a small value of δ. The measurement bandwidth cannot be easily made arbitrarily small, and so a noise limit arises from the value of δ and the amplifiers used. The meter made provision for the attachment of a spectrum analyser to allow odd and even components (the $f_0 + \delta$ and $f_0 - \delta$ components) to be subsequently resolved.

5.4 Ultimate THD and TDFD Meter Performance

Of course, measurement bandwidth of a meter may be made almost arbitrarily small by the use of Spectrum Analysis (SA) in place of the RMS voltmeter depicted in the block diagram. This applies both to THD and TDFD systems. If the filters (BPF for TDFD and notch for THD) are suitably clean and free of distortion and are left in, the dynamic range may be considerably extended. In this case the limitation of a THD measurement system becomes the signal source. It is possible to achieve residual distortion levels 120dB below fundamental.[62] In the case of TDFD the harmonics of the test signals are not a problem since they do not contribute energy at the same frequencies as the fundamentals after the distortion process. The limitation becomes the residual TDFD components from the sources. The new meter uses a "lock-in amplifier", or homodyme mixing system and low-pass filter, to effectively achieve exactly the functionality of a spectrum analyser.

5.5 A New TDFD Design

Figure 5.2 shows the block schematic layout of the new meter. Note that the amplitudes of the required components are found by synchronously detecting them using a lock-in amplifier IC.[2] The synchronous detector is given reference frequencies derived from the

[2]This process may be likened to spectral analysis—synchronous detection, sometimes called lock-in amplification, is equivalent to mixing a required component down to dc. The meter is in fact a spectrum analysing instrument with limited and specific frequency range.

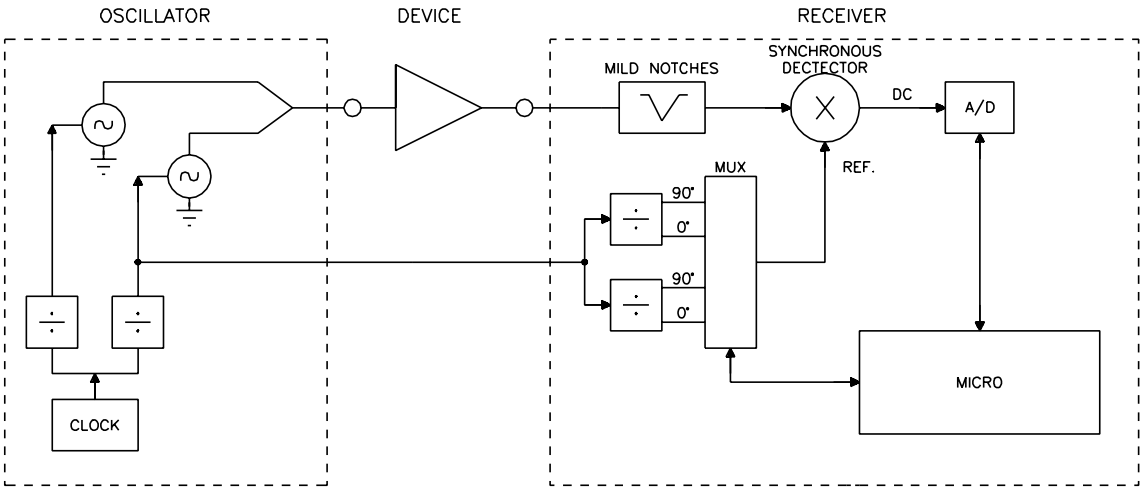

Figure 5.2: Simplified block diagram of the new TDFD meter.

same divider chain as is used to develop the test tones. Quadrature components are used to allow for the unknown phase shift in the DUT and filters. Filters are used in front of the detector to extend its dynamic range, but these are not demanding filters, and simple, fixed-value, passive, twin-T circuits are adequate.[40]

The advantages of this approach over the conventional are:

- *very* low noise bandwidth (under 10Hz, set in conjunction with reading speed);
- simple filters (with only modest tolerance components);
- odd and even order distortions can be separated;
- equivalent THD can be displayed.

5.5.1 Detail

Figure 5.3 shows a more detailed block diagram. The limiting factor on performance for decreasing magnitude test signals is the receiver noise. This is set by the amplifiers used (NE5534 or equivalent opamps) and the bandwidth of the system. The lock-in amplifier incorporates a 10Hz LPF which allows readings to be taken every couple of seconds, with up to three seconds required when re-ranging. Averaging can be used to reduce noise a

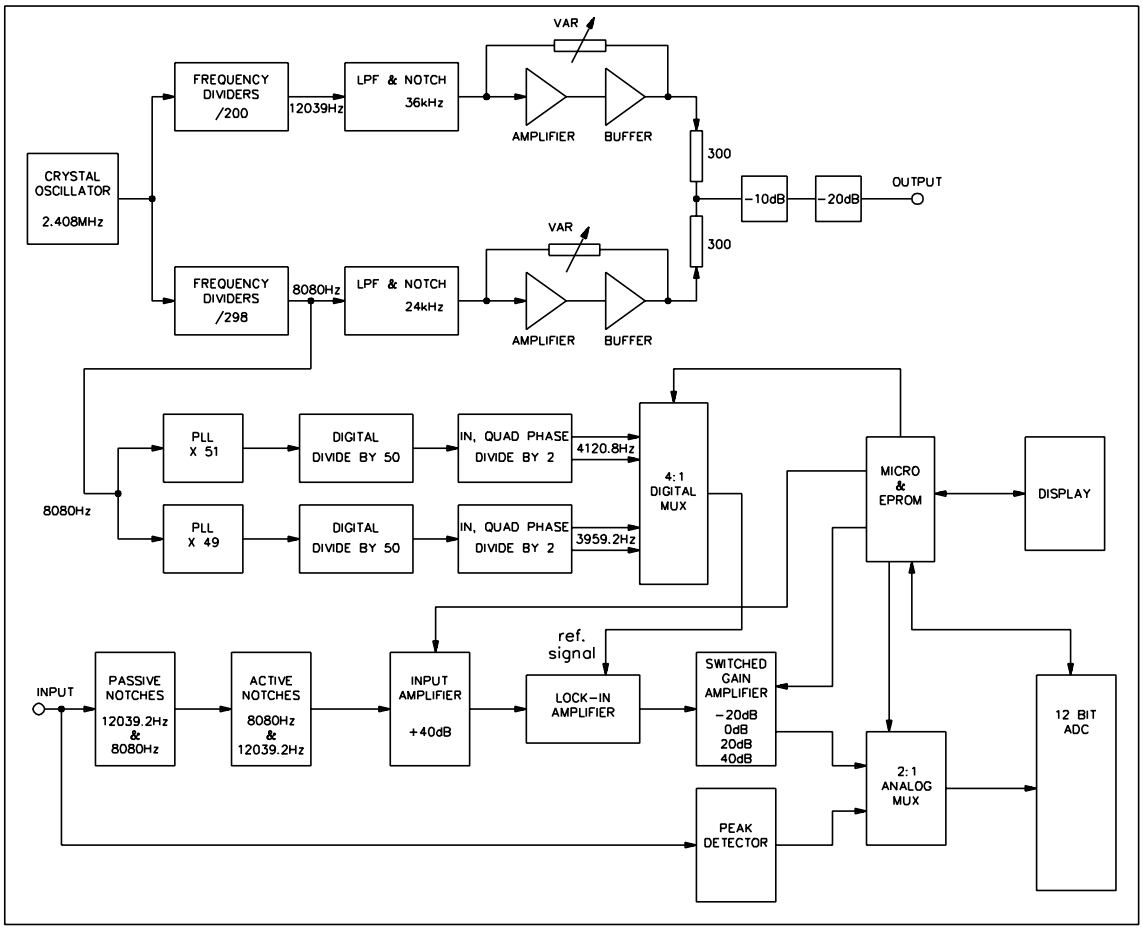

Figure 5.3: Detailed block diagram of TDFD meter.

little at the cost of reduced reading rate, but the system is close to the noise floor set by earlier amplifiers.

When larger signal levels are required out of the signal sources, the limitation becomes the residual TDFD components in the source. Care must be exercised in the development of the two test tones, as intermodulation levels less than 120dB below the fundamental can occur on the bypassed supply of digital divider ICs and can thence propagate through (saturated) gates and out into the shaping filters. With careful layout the residual TDFD becomes significant only when the buffered output amplifiers are delivering close to their maximum levels. In this case, some signal can propagate through the resistive summer from one amplifier into the output of the other channel, intermodulate, and then return to the system output. It is possible to limit this by increasing the attenuation through the summing network, with consequent loss in maximum possible signal and/or increase

in output impedance. Unity-gain buffer ICs working inside the feedback loop of an opamp reduce the buffer stage output impedance to an absolute minimum, and then with 300Ω resistors a maximum signal level of $20V_{p\text{-}p}$ is obtained with residual distortion as shown in figure 5.5. The fine control of output level is effected by adjusting the feedback around these final buffers, so that TDFD falls away as the output level is reduced below maximum. Coarse control is effected by switched attenuators after the summing network. This appears to be the best trade-off.

The analogue input chain starts with a pair of passive notches, and then leads to active ones. The use of passive notches at the input results in lower input impedance, about 5kΩ, but allows the system to use amplifiers in the chain whose non-linearity does not then compromise the sensitivity of the meter as a whole. After these, coarse ranging is performed with a 40dB gain amplifier which may be bypassed for large inputs. After the lock-in amplifier more (non-critical) switched gain stages are available to condition the signals for the ADC. Finally a peak detector is used to measure the amplitude of the test signals emanating from the DUT.

The reference signals for the synchronous detector are developed using PLL frequency multipliers and digital frequency dividers. The prototype frequencies were 8080 and 12039Hz. Quadrature versions of these are sequentially multiplexed by the microprocessor to the single synchronous detector. As indicated earlier, having the synchronous detector's reference signals derived from the same master clock as is used to provide the test tones ensures that they track precisely. Although a crystal reference is used in the prototype, the sole need for frequency accuracy lies in the fixed nature of the passive filters which protect the amplifiers and lock-in amplifier IC from the stimulus signals. It is possible to vary the reference a little without adverse effect. It would be possible to make a version of the instrument with the capacity to phase-lock the master reference to an external signal in the form of a TDFD stimulus signal: The separated $2f_0$ component may be used as the source for the master clock. The value of this approach is to allow a CD player to play a disk carrying the two tones at 8080 and 12039Hz and to use the receiver section of the instrument to measure the player's TDFD, effectively ignoring the test tones generated

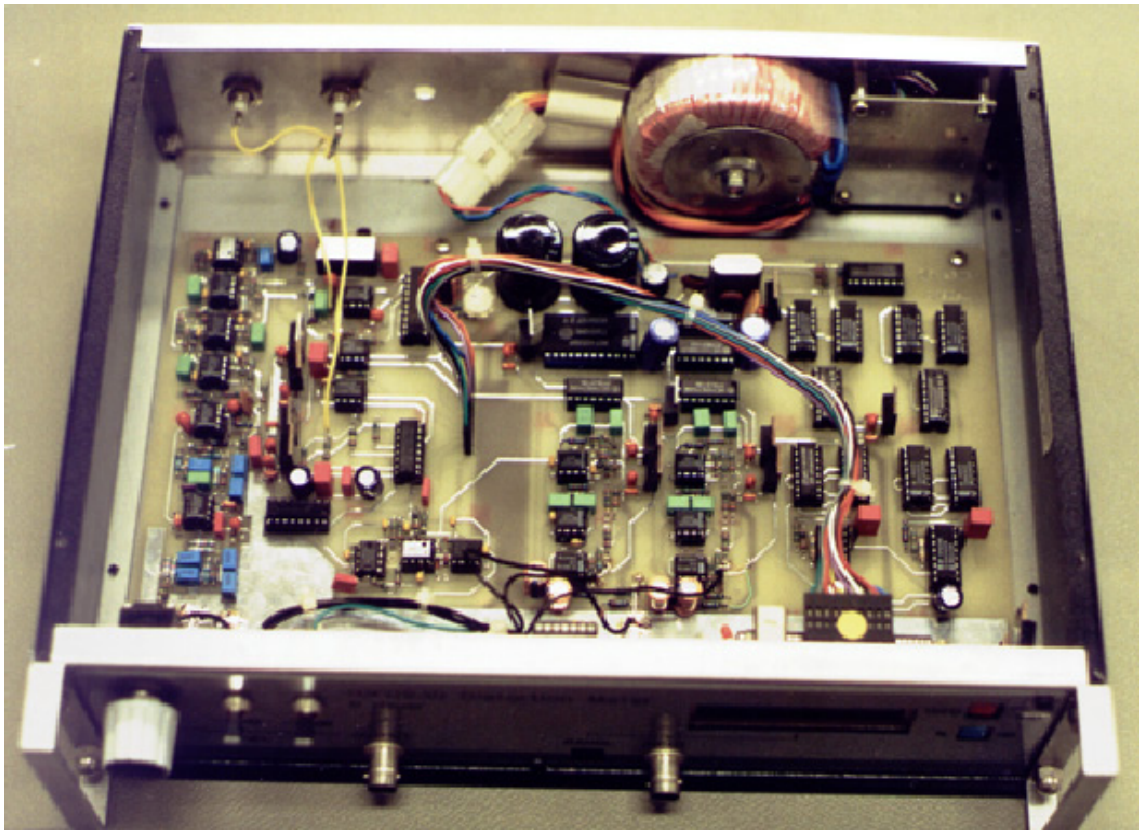

Figure 5.4: Photograph of the prototype TDFD meter. The various analog sections are visible on the main PCB. A commercial PCB carrying the display and microprocessor is hidden behind the front panel. The toroidal power transformer is visible on the rear panel.

internally. There also exists the possibility of providing the instrument with switchable or optionally different filters in both the receiver and transmitter sections, and a switchable or different master reference clock, so that it may have its frequencies scaled down by a factor of five, say, for telephone systems, or twenty for woofer tests. The principle can also be used scaling up in frequency, to make similar tests, say, at RF frequencies.

Figure 5.4 is a photograph of the prototype instrument.

5.6 Verification of Performance

At levels where a spectrum analyser can directly identify the TDFD components (75–80dB below fundamental with an HP3561A Dynamic Signal Analyser) the meter may be compared directly to results determined by observing the DUT while the meter is

connected. Checking that the meter is operating correctly at lower readings is more difficult, and two confirmatory techniques have been used.

In order to both test the meter, and to exemplify circumstances in which TDFD outshines THD as a scalar measure of circuit quality, some simple circuits have been constructed. Their THD performance, and TDFD performances as measured both by spectrum analyser and by the meter, are compared with results simulated using PSPICE. A slew rate limited amplifier (SLA), a class-B circuit with consequent severe crossover distortion (CBA), and a 15kHz low pass filter circuit (LPF) have been used. Measurements are presented for each circuit, and for the series connections of the two deliberately bad amplifiers and the LPF circuit. Table 5.1 shows the results of the comparison.

The minor discrepancies in the simulated results are attributed to the fact that simulations did not use precise component values. (It was the intention only to weakly verify the trends.) The spectrum analyser used was an HP3561A, which has a dynamic range of 80dB, resulting in readings that are noisy when distortion components are about 75dB below the test tones.

5.6.1 Calibration for Very Low Level Non-linearities

Evidently, the region of most interest and concern is that where this instrument alone can make measurements. Two methods have been used to check its calibration at lower levels.

1. The gains of the notch filters can be accurately measured at the frequencies near 4kHz using a calibrated, synthesized generator. Subsequently, a spectrum analyser is connected after the notch filters. The anticipated result is then easily calculated by hand: The calculation consists of adjusting the magnitudes of components displayed on the spectrum analyser by the known attenuation in the filters. The hand calculation is checked against the value with that displayed by the microprocessor.

2. Alternately, low-level calibration can be carried out by using a synthesized generator to inject known amplitude components at the exact TDFD component frequencies

Meter/SA(sim)	V_o	THD(1k)	THD(10k)	TDFD	comments
SLA	4	-75/-74(-77)	-47/-47(-42)	-39/-38(-41)	• Note that only higher THD frequencies expose slew problem; • TDFD signal slews ≡≈THD(10k) for same $V_{\text{p-p}}$.
LPF	4	-78/-76(-70)	-79/-78(-74)	-77/-77(-67)	• THD(10k)<THD(1k) as expected
SLA+LPF	4	-73/-73(-68)	-58/-58(-54)	-40/-40(-40)	• Slew limiting implies THD(10k)>THD(1k)... • LPF effect implies THD(10k)<THD(1k)... • TDFD not fooled!
CBA	4	-45/-45(-45)	-26/-28(-27)	-43/-43(-46)	• Crossover gives many harmonics (≫ 10), so 100kHz THD/SA meters may not be accurate! • $2V_{\text{be}}$ slew causes increase in distortion with frequency.
CBA+LPF	4	-45/-45(-48)	-43/-45(-47)	-44/-44(-43)	• LPF roughly cancels CBA effect as THD frequency changes. • TDFD unaffected by LPF. • TDFD not susceptible to LPF "improvement".

Table 5.1: Table comparing TDFD and THD measured and predicted results for some simple circuits. Numerical values are expressed in dB. Simulated results are in parentheses, and are included as confirmation only.

(3959.2Hz and 4120.8Hz for this machine) directly into the input, while leaving the peak detector connected to a high-level signal. The output of the lock-in amplifier is observed on an oscilloscope to check that the beat frequency is adequately low (below 0.1Hz), and the meter readout compared to hand calculations using the known amplitudes of the high-level and synthesized low-level signals.

Results of these methods typically agree to better than $\pm\frac{1}{2}$dB from -70 to -130dB, on the prototypes. The meter is accurate to ±1dB.

5.7 Examples of Use

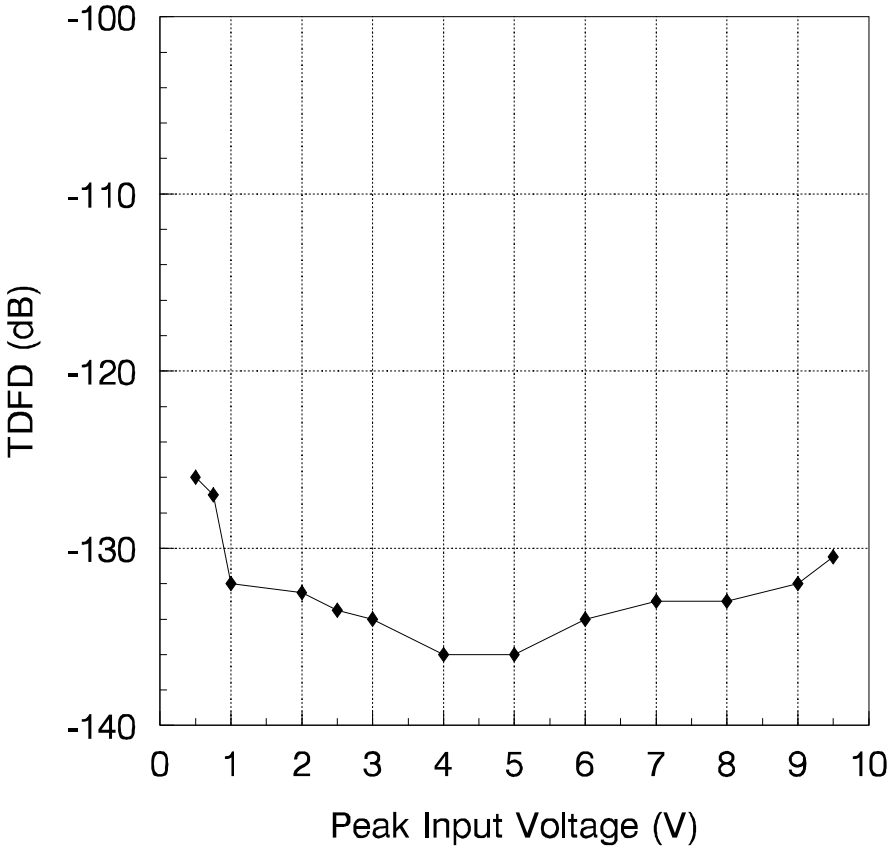

Figure 5.5: Residual TDFD in prototype meter.

Some measurements are included here to demonstrate the meter. Figure 5.5 shows the residual distortion of the prototype, as noted previously. The graph shows an optimum result between 3 and 6$V_{p\text{-}p}$, with the floor rising for the reasons noted above. Figure 5.6 shows the TDFD of a laboratory buffer amplifier, constructed by the author. Labels on the graph indicate typical limits for commercial THD equipment and the best reported THD of an oscillator. Figure 5.7 shows the TDFD of a '5534 opamp at various signal levels with various source impedances. The meter is able to readily see the onset of effects as source impedance rises; these would not be visible with a less sensitive meter. Although rather uninteresting a result, figure 5.8 shows measurements of distortion against time after application of a stimulus for another laboratory amplifier. The result serves to

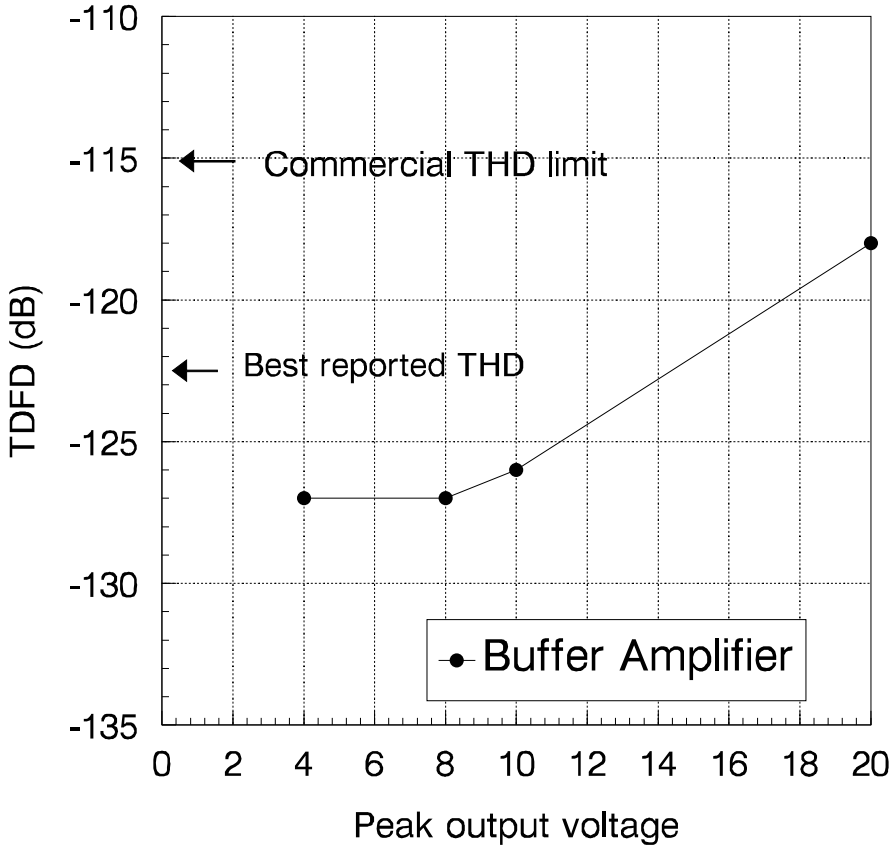

Figure 5.6: Distortion of an example amplifier measured with the prototype meter.

demonstrate the speed with which the instrument obtains results, and consequently the rapidity with which it is possible to see the effect of adjustments, etc., using this system.

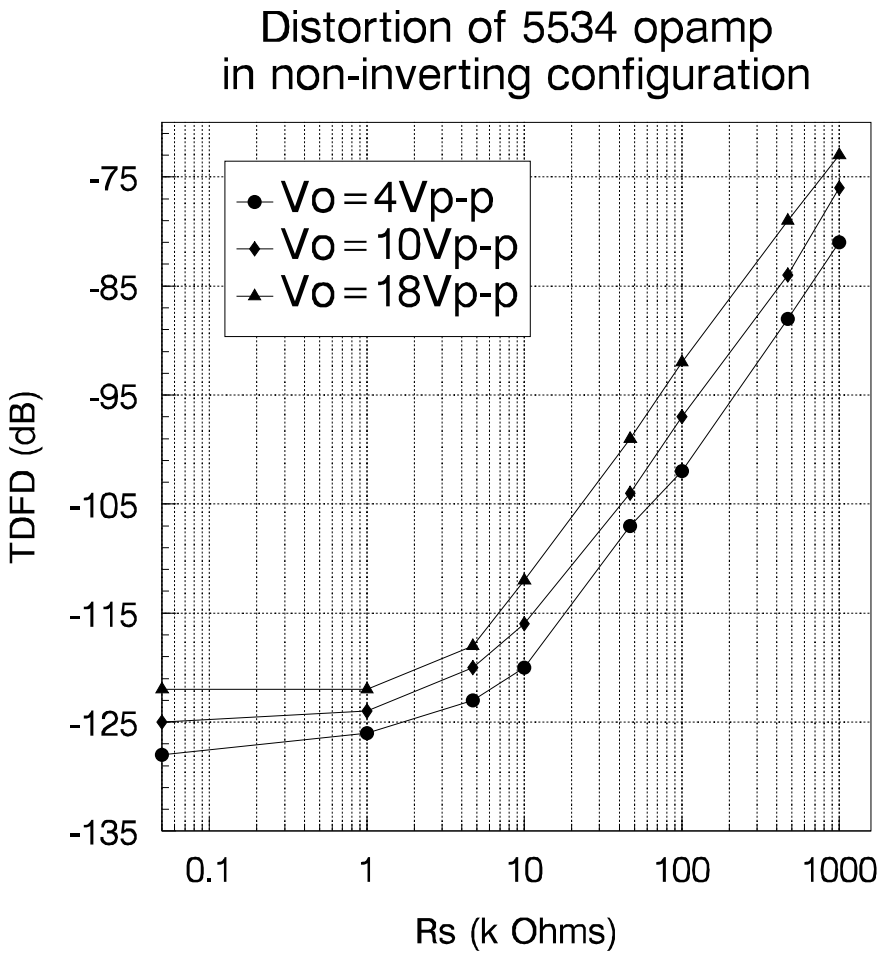

Figure 5.7: Comprehensive set of measurements made with the prototype. Note that the levels of this standard circuit are below those that may be optained with conventional THD systems.

5.8 Conclusion

A meter that uses synchronous detection and that achieves uniquely-high performance total difference-frequency distortion measurement has been demonstrated. The instrument

- offers the best sensitivity to non-linearity reported to date (-135dB);
- is convenient to use (no calibration or adjustment);
- obtains results very swiftly (< 5 seconds to range and read);
- offers discrimination of even and odd order distortion products;

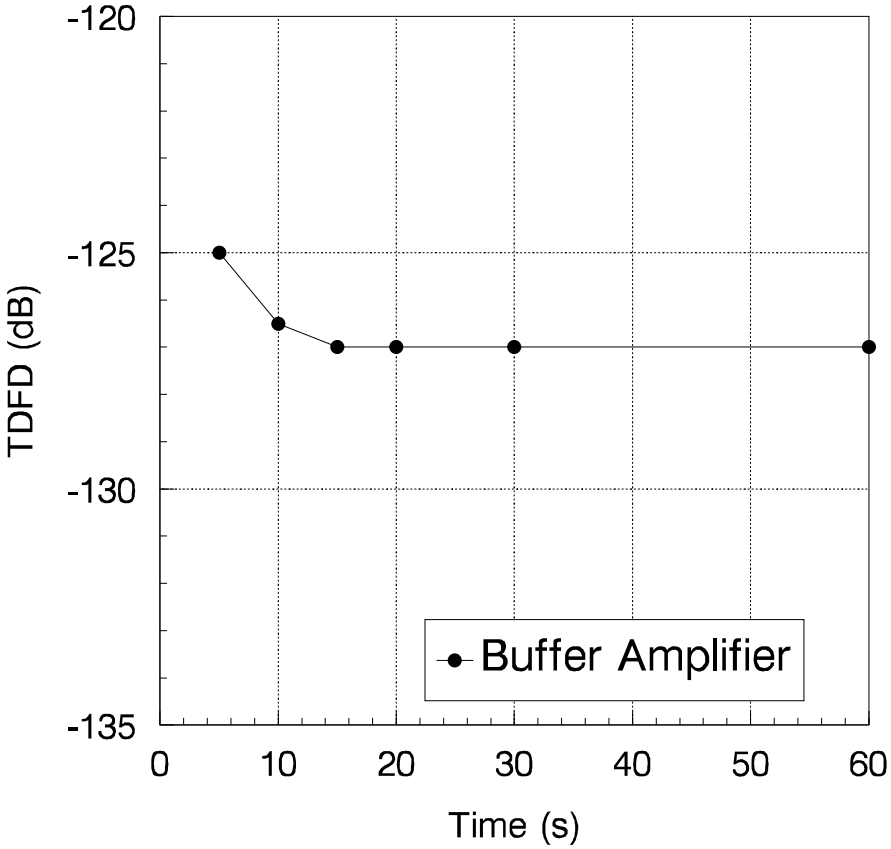

Figure 5.8: TDFD versus time for the same buffer amplifier as measured in figure 5.6.

This is a powerful demonstration of what is possible in the field of linearity measurement.

Chapter 6

Arbitrary Pulsed Semiconductor Parameter Analyser

This chapter first describes the need for, then the design of, and finally various novel applications of, an instrument vital to low-signal modelling and simulation of III-V semiconductor FETs. Much of the system design detail has been relegated to several appendices. At the time of writing, the system described here is being taken to market by Hewlett-Packard's Santa Rosa Systems Division.

6.1 History

Transistors fabricated in III-V semiconductor material exhibit relatively large variation in their characteristics with channel power dissipation (temperature) and/or applied field strength (trapped charge). Increasing the power dissipation (and hence the channel temperature) of a III-V FET typically reduces the drain current by a several percent. When the characteristics of a device are measured in the conventional way, typically with an instrument such as the industry-standard HP4145, so-called "thermal dispersion" causes the drain-characteristic curves to "droop" at higher drain voltages and currents. This is a straightforward consequence of well-researched effects.

Since thermal dispersion occurs on a timescale that is long compared to the period of signals typically processed by RF circuits (μs or ns), but short compared to the period typically required for high-precision characteristic measurements (tens of ms or more), the system cannot be regarded as *time-invariant*. This prevents the successful use of data from one timeframe being used to predict performance in the other. In other words, any estimates of transconductance, g_m, or output conductance, g_{ds}, for example, based on low-frequency ("dc") measurements, will lead to erroneous predictions of high-frequency (RF) performance. This applies to any parameter of course, not solely linear ones such as g_m and g_{ds}.

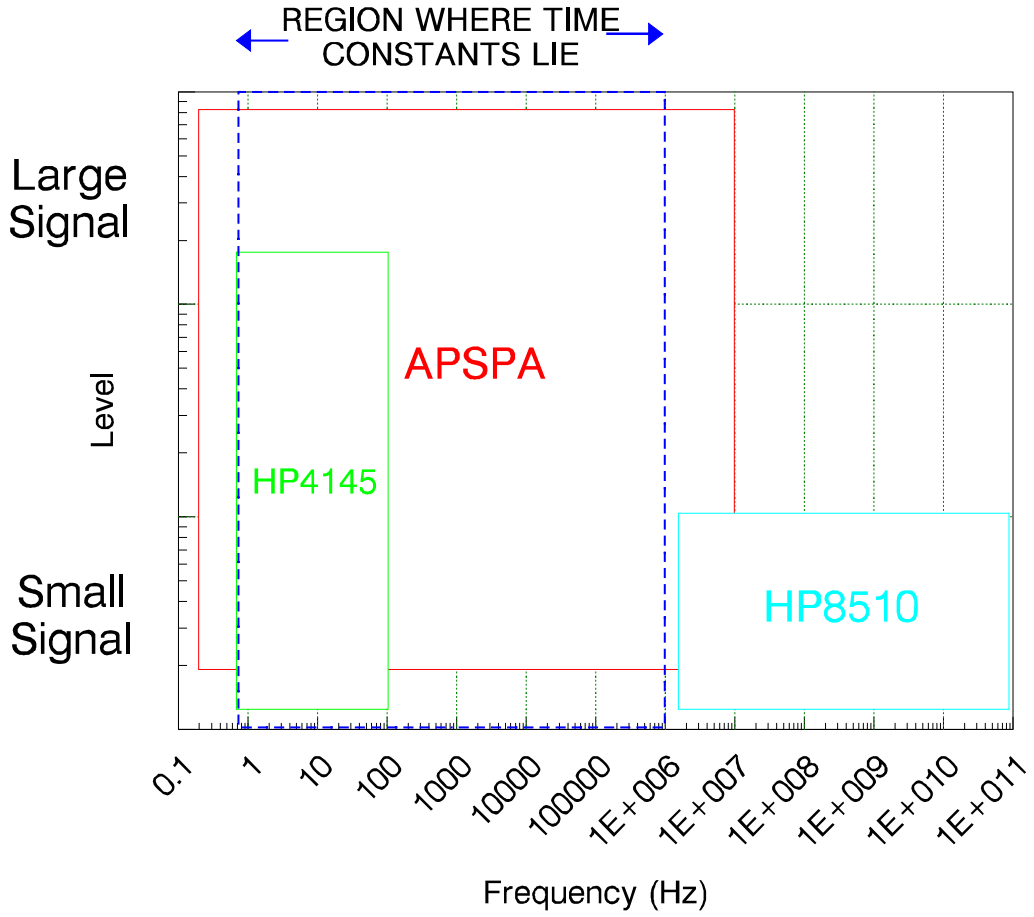

Figure 6.1: Typical location of time constants associated with GaAs devices and regions covered by conventional measurement systems are shown graphically on a "speed/level" plane.

Figure 6.1 presents the situation graphically. Most time constants, associated with thermal and charge migration effects, are believed to fall in the region between a few seconds and a

few microseconds. Established measurement systems, epitomised by the industry-standard HP4145B and the HP8510C, cannot characterise a device all the way from small-signal to large-signal, from below to above these limits.[31] Knowledge of nonlinearity, especially when level-dependent, clearly demands large-signal and small-signal data gathered from both sides of the region where behaviour changes.[1] The span of the APSPA system to be described is included in the graph.[2]

The last decade has seen the GaAs device community realise the need for pulsed (i.e., fast) measurement of devices. A number of reports of systems making these measurements have appeared in the literature.[50]–[55]

There is now evidence to support the idea that thermal effects are not solely to blame for the dispersion of characteristics, but that charge storage effects may also be significant, if not dominant.[47, 45] In fact it has been acknowledged for some time that charge storage can contribute to time dependence of characteristics,[95] but the extent is still open to debate. The author has demonstrated charge-storage phenomena in a device elsewhere.[61]

A more recent phenomenon is long-term or permanent change in device characteristics, reported first in 1995. This is thought to be brought on by EPROM-style charge migration and storage, induced by extreme or sustained electrical "stressing" of devices through large-signal operation.[60, 59] The short-timescale characteristics, and hence any change in them, are only visible with pulsed measurement.

6.1.1 Split-model Response

A brief digression on the response of industry to the fundamental nonlinear-modelling problem posed by device time-variance is appropriate.

[1] Knowledge of the behaviour in the transitional region will mean measurements through it, but this is less critical since GaAs devices do not normally process signals in the transitional frequency range.

[2] For clarity of the argument the graph does not show the extensions to the HP4145 region afforded by modular systems such as the HP4142, nor proposed large-signal RF systems such as the Microwave Transition Analyser-based systems under development at Belfast University.[48, 49] These do not compromise the argument, but would complicate the picture.

Because of the difference between dc and RF characteristics, early designers used one (simple) model, fitted to the low-frequency measurements, for bias simulation, and another (simple) model fitted to RF (S-parameter) data for small-signal simulation. This split-model approach persists in industry for several reasons:

1. It avoided the need to devise a full model, incorporating the history-dependent effects, and to solve problems of fitting such a complex model;
2. Because it relied on elementary equivalent circuit models, it could be implemented with existing simulators;
3. Bias and small-signal simulation results were sufficient for the first generation of III-V semiconductor circuits.

This thesis concentrates on low-signal operation, where such a two-model approach cannot succeed.

MESFET models

The MESFET is the most common III-V device in use. The need for an adequate full model, even for this most common III-V FET, has spawned a number of improved models, including Statz[80], Rodriguez[81], TOM[82, 83], PS[84], Root[85, 86], Scheinberg[87], various orders of Curtice[88, 89, 90], Materka[91], Khatibzadeh[92] and others[93, 94, 95].

The Parker-Skellern model appears to be gaining somewhat wider acceptance than others.[84, 97] It is available in SPICE from Berkeley, from MicroSim's industry-standard PSPICE,[14] and from IntuSoft's IntuSpice.

6.2 The APSPA Measurement System

Comprehensive device measurements are needed to investigate device effects that must be incorporated in the models, and to which to fit the models once they are developed.[31]

These fast, pulsed, device measurements are referred to a "pulsed-bias" or "pulsed-I/V" measurements. In order to explore the characteristics of GaAs devices, a pulsed-I/V system has been made, with the support of Hewlett-Packard Systems Division. It has been made to be extremely versatile. It offers high accuracy, unequalled freedom in the format of the stimulus-responses it can provide, high power, and is software integrated with a network analyser and external high-power pulsers. (Specifications are given in appendix A.) Because of this, it is called an "Arbitrary Pulsed Semiconductor Parameter Analyser" (APSPA), after the classic, industry-standard, HP4145 Semiconductor Parameter Analyser (SPA).

The APSPA system was unveiled at the exhibition of the 1996 IEEE Microwave Theory and Techniques Symposium (MTT-S) in the Hewlett-Packard booth, and presented in a paper at the subsequent Automated RF Techniques Group (ARFTG) meeting.[61] The system was well received by the international community, and at the time of writing it is being adopted as an HP product. Appendix B describes the demonstrations prepared for HP and run at the symposium. Appendix C is a brief user's manual for the APSPA measurement system.

Figure 6.2 shows the block diagram of the APSPA system, figure 6.3 gives the physical dispositions of the VXI instruments in the rack, and figure 6.4 is a photograph of the prototype.

The basic pulsed-I/V system is constituted by the VXI modules and the current probe(s). This system consists of four basic sections.

1. An embedded controller in the form of a Radi-Sys EPC-7 80486-based PC.
2. A hardware timing controller, an HP E1450A[3] 160MHz Timing Module, and a custom patch-panel for distribution of the control signals.
3. A gate Source-Measurement Unit (SMU), comprised of an E1340A Arbitrary Function Generator (AFG) and an E1446A DAC/Amplifier making up a pulse source,

[3]The E1450A does not usually appear separately in catalogs, but is part of the D20 digital test system, the associated pattern modules being deleted in this application.

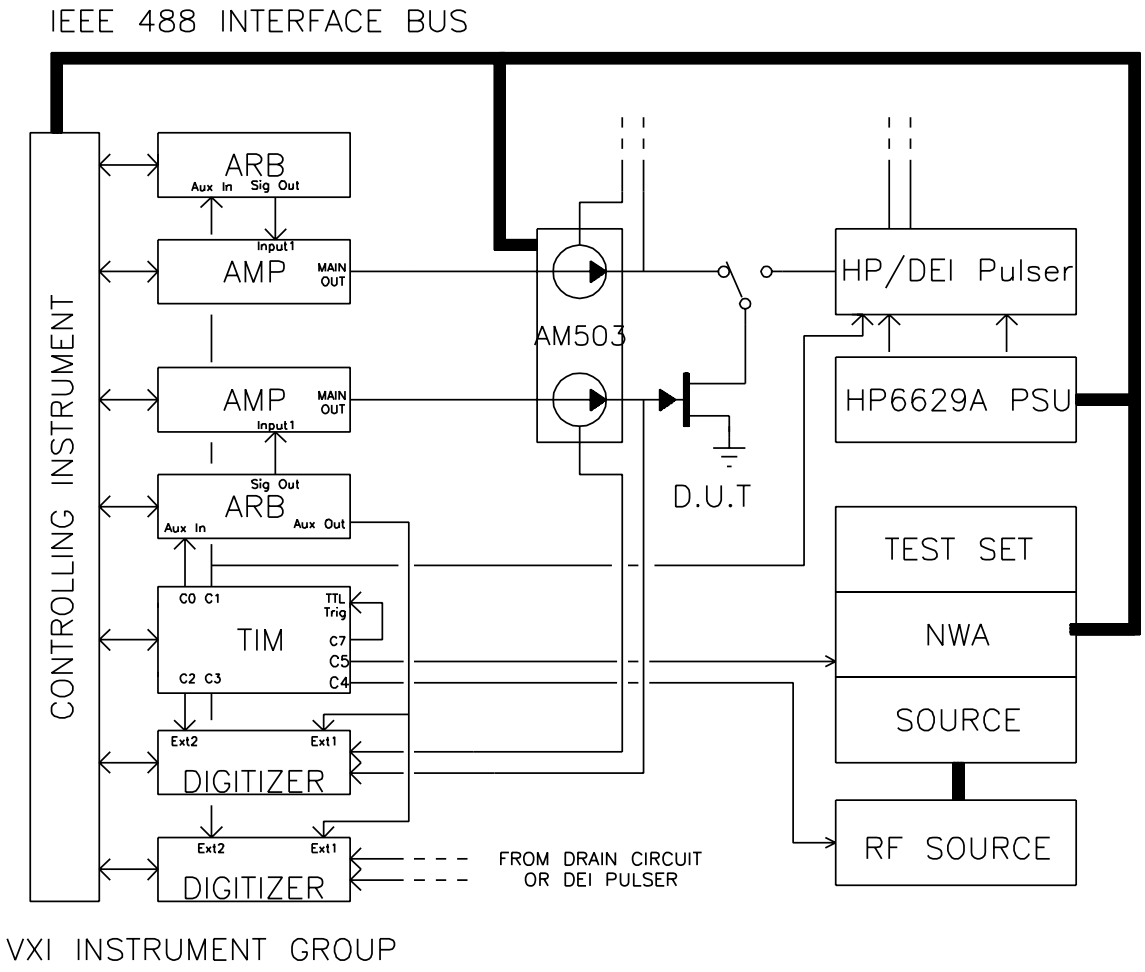

Figure 6.2: Block diagram of the APSPA measurement system.

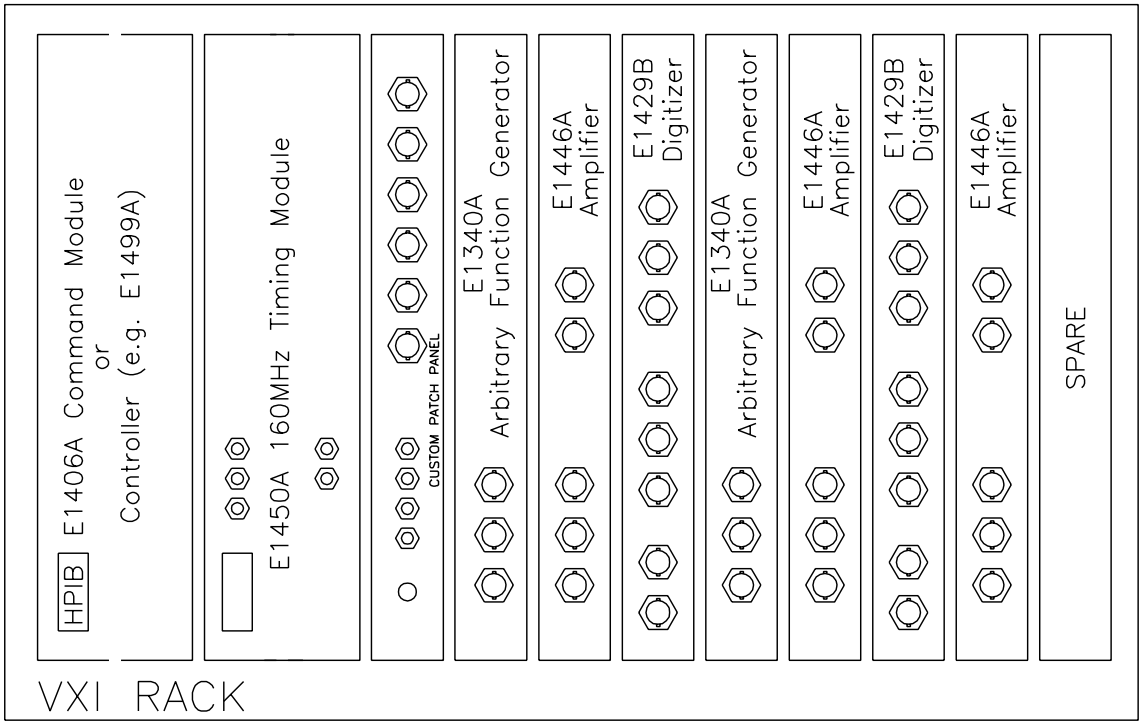

Figure 6.3: Layout of the instruments in the VXI rack.

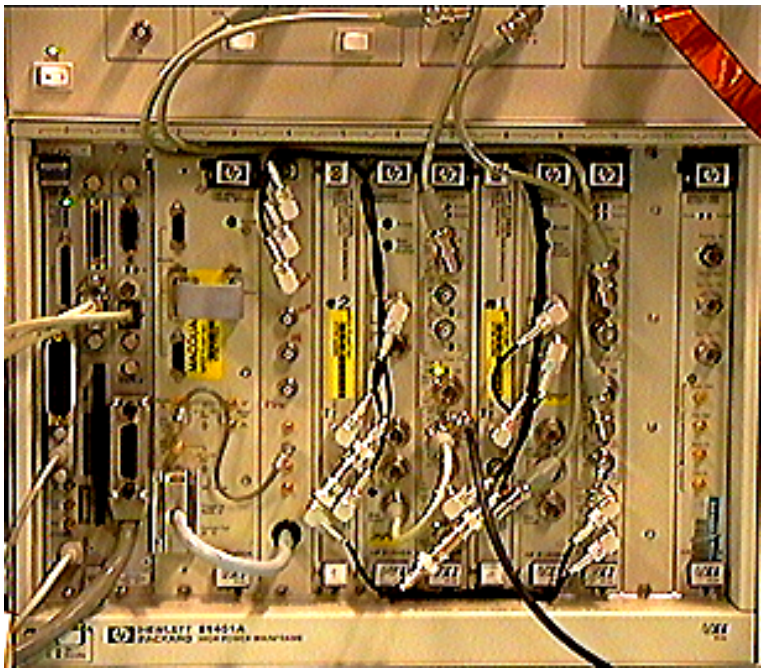

Figure 6.4: Video image of the assembled VXI APSPA system.

and an E1429B, two-channel, 12-bit digitiser for recording voltage and current via a Tektronix AM503, dc-50MHz, Hall-effect probe.

4. A drain SMU, identical to the gate SMU, but with the option of a second E1446A power amplifier to increase current capability.

In the basic mode of operation, the AFGs are loaded with 4096 data points, starting with a zero, and then with 2048 points to be measured loaded in alternate locations interspersed with zeroes. The data in the points-to-be-measured locations represent the required deviations away from the desired device quiescent point for gate and drain voltages. The DACs within the amplifiers are loaded with a value that produces the desired gate and drain quiescent points. (Considerable scaling is done on each run to set the amplifier gain and AFG output amplitude to ensure that the best possible dynamic range is obtained for the particular voltages to be delivered. Also, a calibration test is run to allow the cancellation of amplifier offsets via soft modification of the data to be loaded.) The amplifier outputs are enabled, and the test sequence applied to the device under test (DUT). The digitisers are then programmed to arm on the AFG end-of-sequence marker, and to digitise voltage and current values synchronised by the timing module.

The test data are normally generated from start, stop and step values for gate and drain voltage. They may also be read from a data file, permitting arbitrary trajectories in the V_g/V_{ds} plane. These can easily be generated, for instance using an AWK script or compiled C program.

The digitisers will take as many samples as required in a single pulse, and will cycle through the pulse sequence as often as is required to reach the requested averaging factor. The time window within the pulse during which data is considered "valid" may be arbitrarily set. The event time for application of RF, gate and drain stimuli is also arbitrarily defined. (Details of this are given in the software manual in Appendix C.) As data is downloaded from the digitisers, it is averaged in real time and stored. If more data is required than will fit in the half-megabyte memory, multiple downloads are used. If more points are requested than can be accomodated in the AFGs, multiple uploads are used.

The order of the data in the AFGs may be scrambled, and will be unscrambled after the averaging but before storage. This enables a powerful confirmation that the pulse time is adequately short and/or the pulse spacing sufficiently great, as will be described below.

If drain currents in excess of 1A or voltages in excess of 20V are needed, an HP85120A-K43 high power pulser may be added to the system. This gives 10A/50V capability, and is synchronised using the same timing hardware. However, the HP85120A can only switch between two preset levels quickly, and so multiple drain voltages force multiple download/digitise sequences. Scrambling is carried out only on the gate stimulus.

If S-parameters are to be measured using an HP85108C Pulsed S-parameter Network Analyser (NWA), only one point can be measured at a time. This is because the NWA requires one pulse to appear for each frequency and for each time it is to be averaged by the NWA. In this case, scrambling has no effect for the S-parameter measurements, but a separate run is used for the I/V data alone, where a scramble is effective.

6.3 Novel Applications

For examples of routine use of the system, the reader is referred to Appendix B. A number of capabilities unique to the APSPA system, among pulse measurement systems, are briefly presented here.

6.3.1 Selecting Pulse Timing

One must ask first, once pulsed testing is available, exactly how short a pulse must be, and how low must the duty cycle be, for the characteristics to be both isothermal and isodynamic.[4]

It is possible, using the arbitrary sequence capability of the APSPA system, to make

[4]Isodynamic in this application implies the same charge distribution, analogous to isothermal but for stored charge instead of temperature.

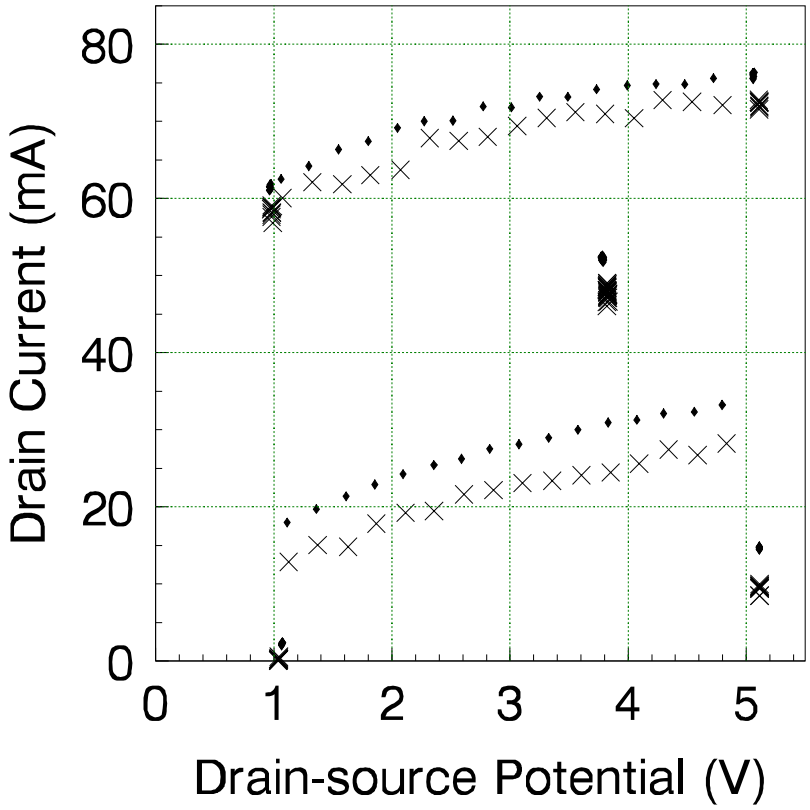

Figure 6.5: Graphical representation of the measurement sequence finding history-dependence error in a device.

measurements across the V_{gs}/V_{ds}-plane but to return periodically to a single point. Figure 6.5 shows such a sequence. (Figure 6.6 replots the data of figure 6.5 numbering points in sequential order. Although the multiple visits to the reference point are obscured, the ordering effect is neatly visible.) If there is some history-dependence affecting the measurement—that is, if the drain current at a measurement is affected in any way by the last measurement, the operating history of the device—the current will not be the same on subsequent returns to the reference point. If the values for a number of visits to the reference point are checked for deviation ("noise"), a measure of history-dependent error may be obtained for the timing (t_p and t_q) used. Figure 6.7 plots a surface representing the error as a function of pulse (t_p) and quiescent (t_q) intervals.

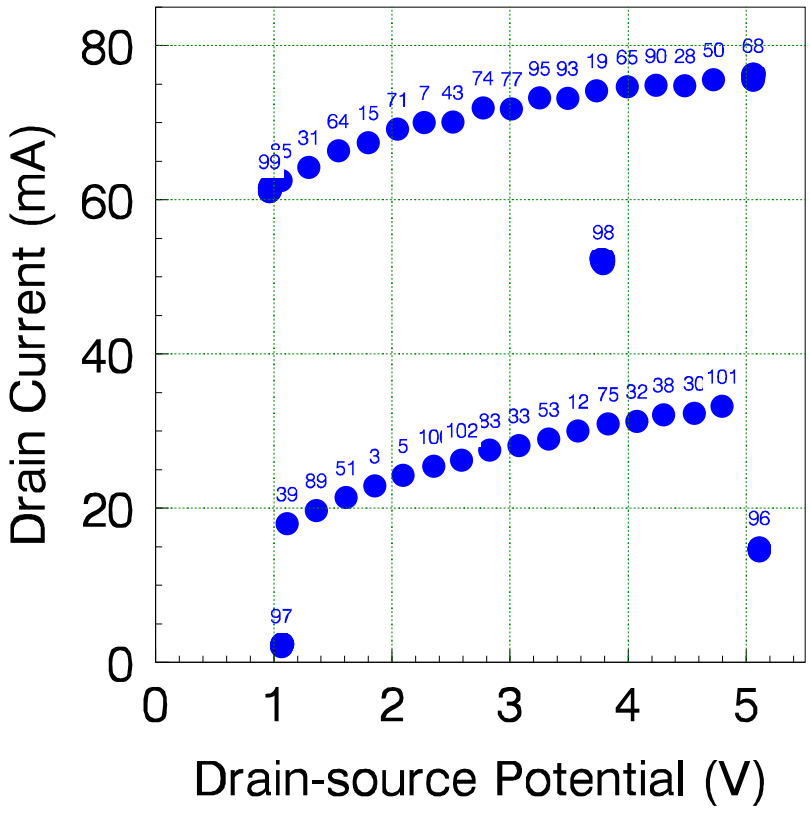

Figure 6.6: Repeat of some data from figure 6.5 with the points numbered in order of measurement to show the pseudo-random sequencing.

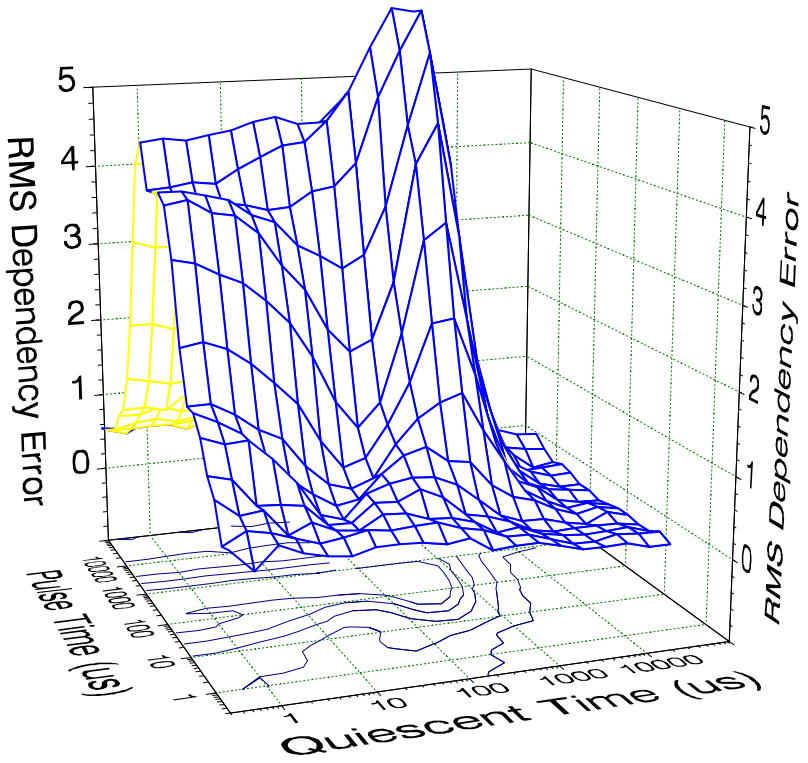

Figure 6.7: History-dependence error surface for a M/A-Com KF44 MESFET.

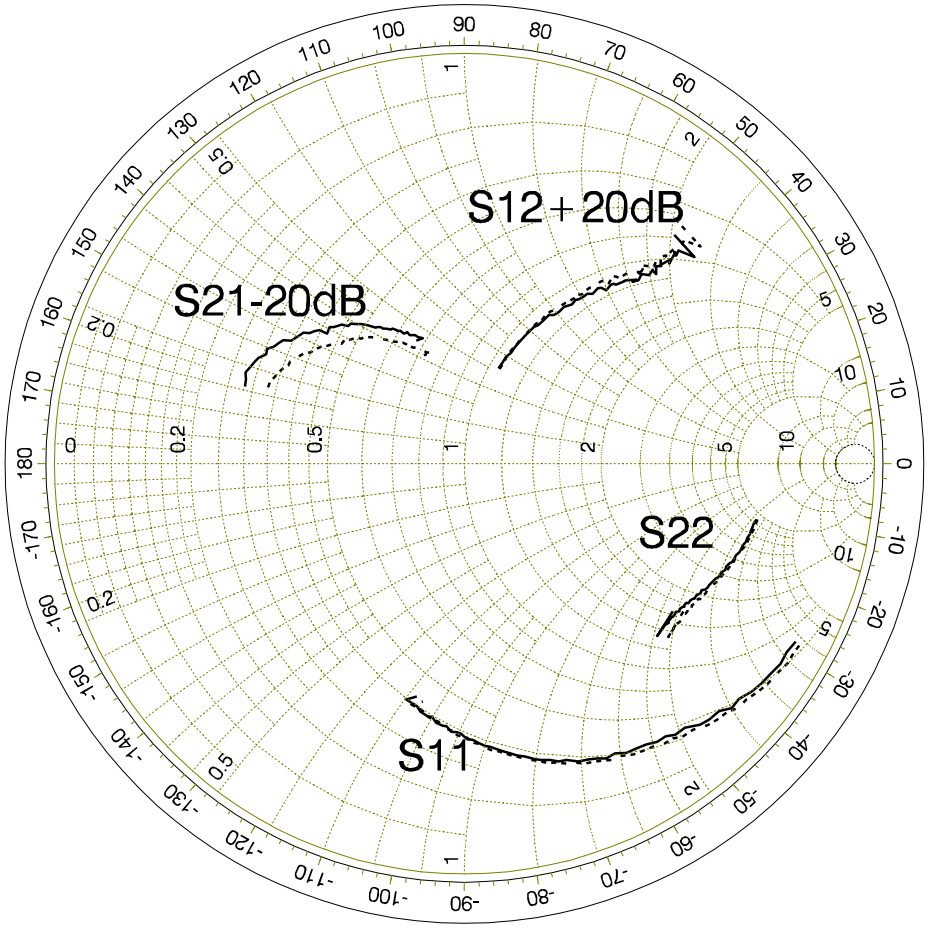

Figure 6.8: Comparison between S_{21} measured under pulsed conditions with parameters measured under conventional, dc conditions.

As noted elsewhere[57], true isodynamic, isothermal measurement of a characteristic requires that the pulse be so short that the surface is flat for that pulse time and simultaneously invariant with quiescent time. (i.e., the surface is flat all along the line with t_p equal to that value.) For these isothermal and isodynamic measurements to be characteristic of the quiescent bias point, rather than the average of the pulse conditions, requires that quiescent time be sufficiently long that the surface is flat and invariant with pulse time. (i.e., measurement is made on the corner of the surface where it seems flat looking along lines of constant t_p and t_q.)

6.3.2 S-parameters as a Function of Bias Point

Figure 6.8 shows the S-parameters for a GaAs device measured in 1μs pulses to a typical operating point (coming from a quiescent point of $V_{gs} = 0$ and $V_{ds} = 0$), plotted with S-parameters measured in the conventional way, with the device biassed at the same point at which the S-parameters are measured.

It is easy to understand that pulsed S-parameter measurements might be important in work where a device is to be used in a pulsed-power mode, such as is encountered in radar systems. This graph is included because it leads to a different, potentially important, point. The two sets of parameters in figure 6.8 differ only in their bias point: They demonstrate that S-parameters at a fixed bias point change with *average* operating conditions.[5] The implication is that it is potentially misleading to base linearity performance predictions on a set of S-parameters, perhaps measured across the portion of the device characteristics that it is expected will be traversed in operation, but measured each with the device biassed at the point of S-parameter measurement.

Figure 6.9 serves to clarify the point. Suppose it is intended to operate this device at $V_{ds} = 6$V and $I_{ds} = 300$mA. S-parameters are measured at, say, 9 points, as marked in the figure, centred about the bias point. Imagine that the S-parameters are to be used to predict nonlinearity arising because of changes in transconductance when the signal excursion is comparable in extent to the spread of the numbered, S-parameter measurement points. The message of figure 6.8 is that the estimates may be wrong if the parameters from points 1–4 and 6–9 are not measured with bias set at point 5.

6.3.3 Pulse Profiling

The pulse-profile capability of APSPA offers an efficient way of obtaining data from which the time constant positions described in figure 6.1 may be identified. Figure 6.10 shows

[5] This occurs, of course, because the g_m and g_{ds} at the point of measurement change with bias conditions; i.e., S-parameters strongly reflect simple properties measured at lower frequencies, as well as reactive parasitic effects.

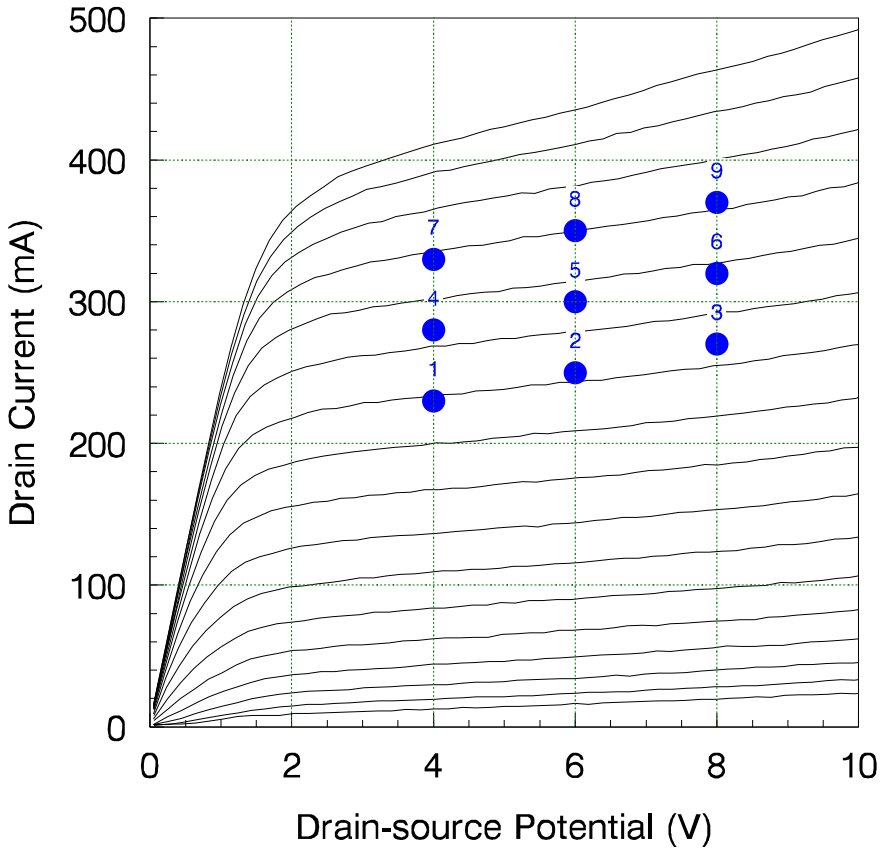

Figure 6.9: A device drain characteristic with points marking where S-parameters might be measured to assist in predicting low-signal linearity behaviour.

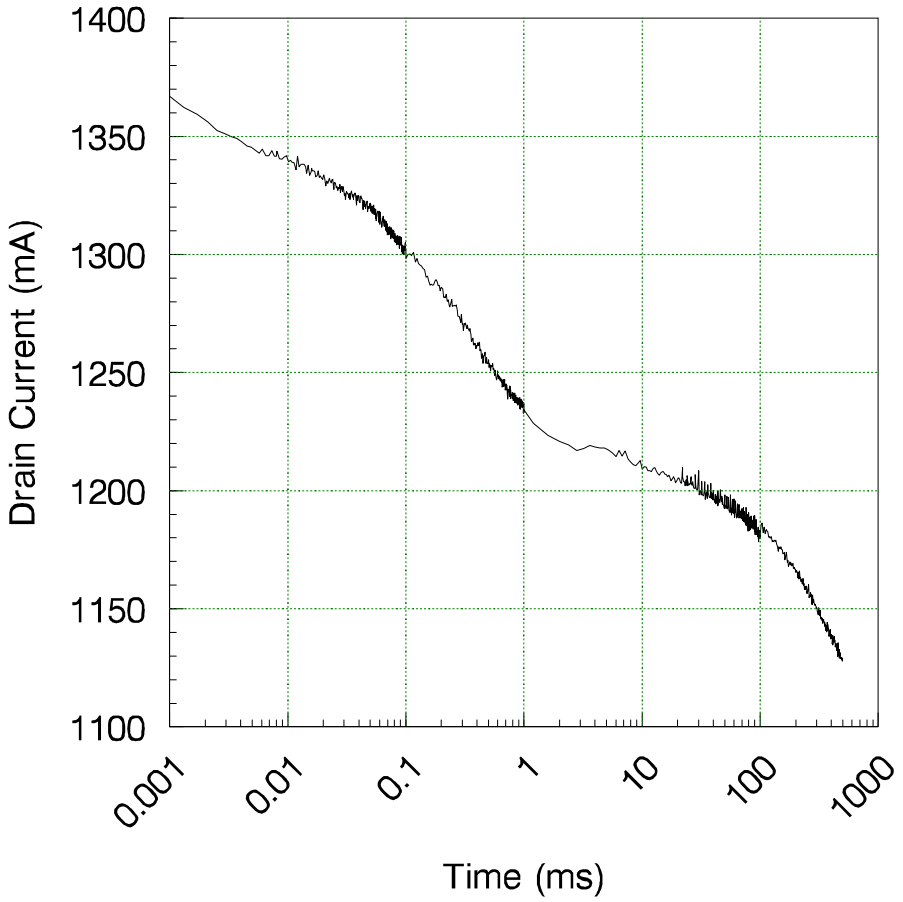

Figure 6.10: Drain-current as a function of time for an HP AT44101 FET, plotted on a logarithmic scale.

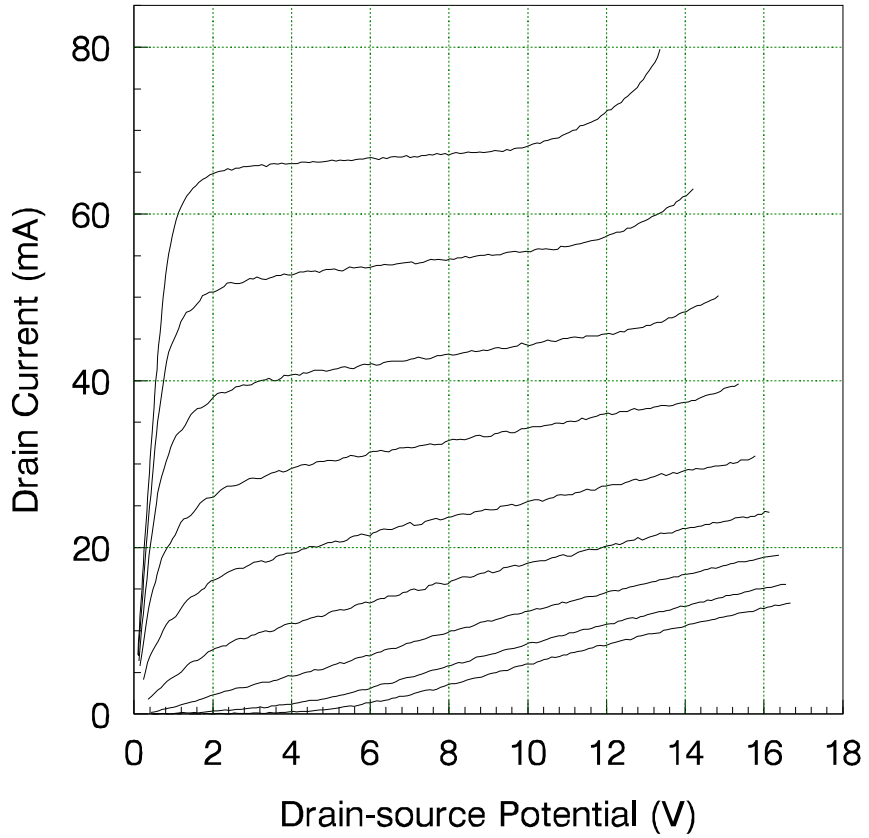

Figure 6.11: Characteristics of a GaAs FET measured *significantly* outside the device safe-operating-area (SOA).

data gathered for one MESFET. Numerical methods as developed in [58] may be applied to yield the wanted information for, say inclusion in models.[97] The system can even gather data in pseudo-logarithmic fashion which suits the approach in [58], for example, see figure B.9.

6.3.4 Breakdown Characteristics

It has been shown[65] that high-frequency breakdown in MESFETs occurs differently from dc breakdown. The speed of the APSPA system permits the RF breakdown region to be approached, both because it can use short pulses, but also because measurement need take only a few seconds, so it is possible to "wind up the knob" in almost real time,

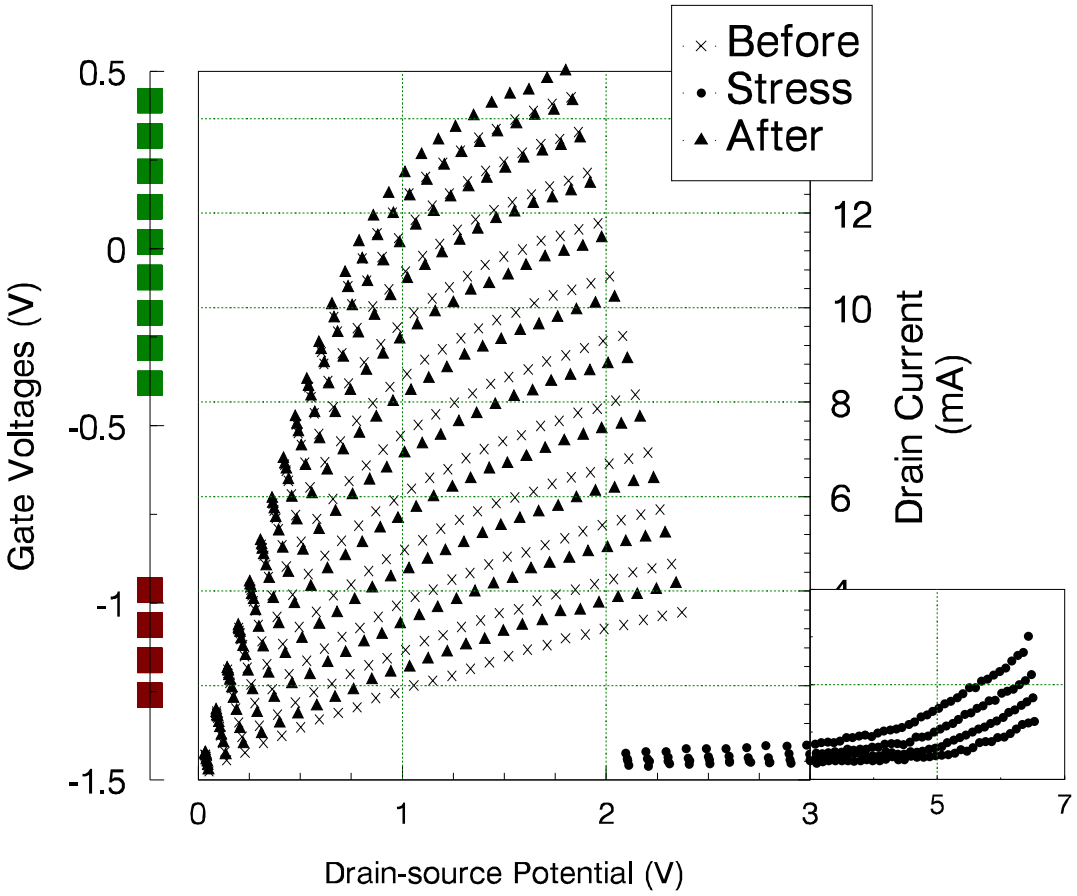

Figure 6.12: Characteristics of a Monofast FET before and a few seconds after exposure to field-induced stress. The graph shows also the response to the stress stimulus.

stopping once the breakdown is visible but not excessive. Figure 6.11 shows a set of such results. These either cannot be measured at low frequency, because positive feedback causes runaway and harm, or the results are not characteristic of what occurs in fast signal excursions.

6.3.5 Observing Charge-migration Effects

Charge-migration effects with long or infinite time constant, such as recently reported in GaAs FETs,[60, 59] affect the high-frequency characteristics. Not only can the APSPA system observe the change in characteristics, it can do so rapidly and repeatedly, to search for shifts. It can also *induce* charge movement. The system possesses the capacity to apply

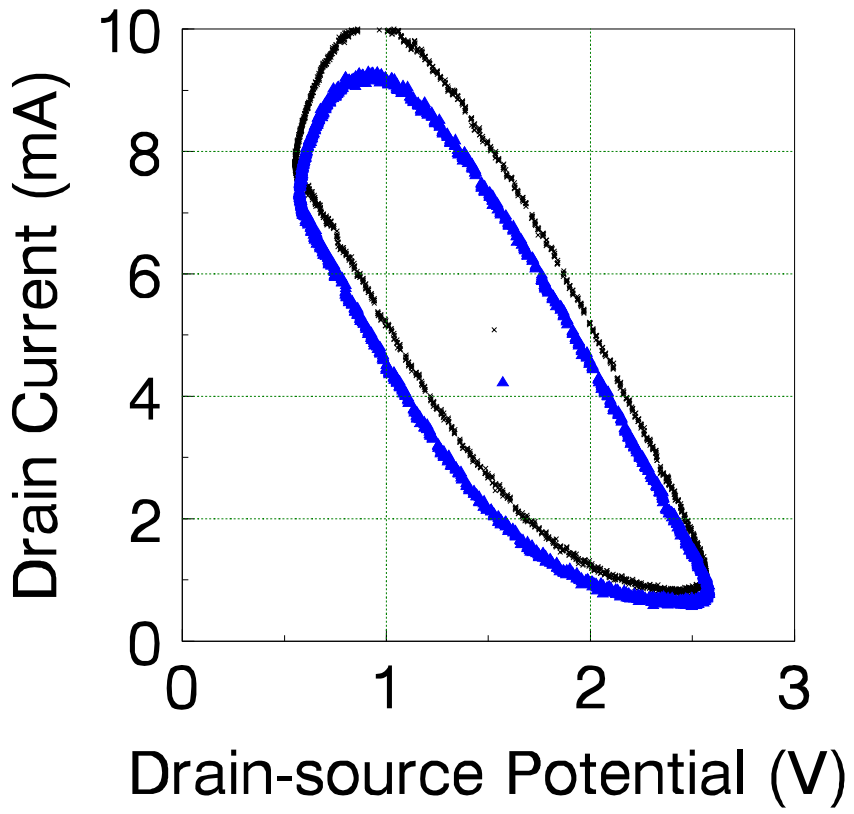

Figure 6.13: Drain trajectory of the Monofast FET in figure 6.12 before and after field-stressing with the stimulus simulating operation with a reactive load.

brief, controlled high-voltage pulses suited to charge injection, but of a duration and duty cycle such as to mimic full-power RF operation or to induce non-destructive breakdown. Figure 6.12 displays "before and after" characteristics of a FET whose operation has been altered after charge is forced onto an island deliberately left floating near the channel structure.

6.3.6 Arbitrary Trajectories

The arbitrary trajectory capability used first in the timing studies above can impose stimulus signals representative of an RF load-line. Figure 6.13 presents "before and after" simulated complex load line trajectories for a Monofast FET. Between the two (identical)

load-line style tests, it was exposed to short pulses sufficient to break down the gate-drain junction. The trajectories shown were made by establishing a series of $\{V_{gs}/V_{ds}\}$ voltages yielding the pseudo-elliptical voltage-current loops displayed, interspersed with returns to the quiescent point (shown by the single symbol near the centre of the loop). The three sequences are readily carried out by the controlling software, which has a script-like sequencing capability.

6.3.7 Channel Temperature Measurement

The system can readily see the effect of change in bias upon both the drain characteristic, and S-parameters. Figure 6.14 shows the drain characteristic of a Triquint FET for two quiescent conditions, overlayed with a conventional ("dc") measurement. The "warm" curves are measured for bias $V_{gs} = -0.25$V and $V_{ds} = 3.5$V, while the "cold" curves correspond to a bias point of $V_{gs} = -0.25$V and $V_{ds} = 0$V. Note that the warm and dc traces coincide at the bias point. The effect of apparent negative output conductance at higher I_{ds} values is also visible. What is most interesting is the difference between the two sets of pulsed data. The cold traces are more widely spaced about the operating point (designated by a letter "Q" on the graph) which is consistent with the transconductance dispersion effect commonly observed in these devices. The warm pulsed measurements are more indicative of the expected operation of this device about this operating point.

Figure 6.15 shows the magnitude of S_{21} measured under pulsed (isothermal) conditions at room temperature from a quiescent bias point with no power dissipation, compared with ones measured using conventional (static) bias conditions. There is a difference in the order of 1dB. This same difference can be reproduced by isothermal measurements carried out at various temperatures. The 70C trace in the figure (coincident with the static measurement) effectively identifies the channel temperature under static operation.

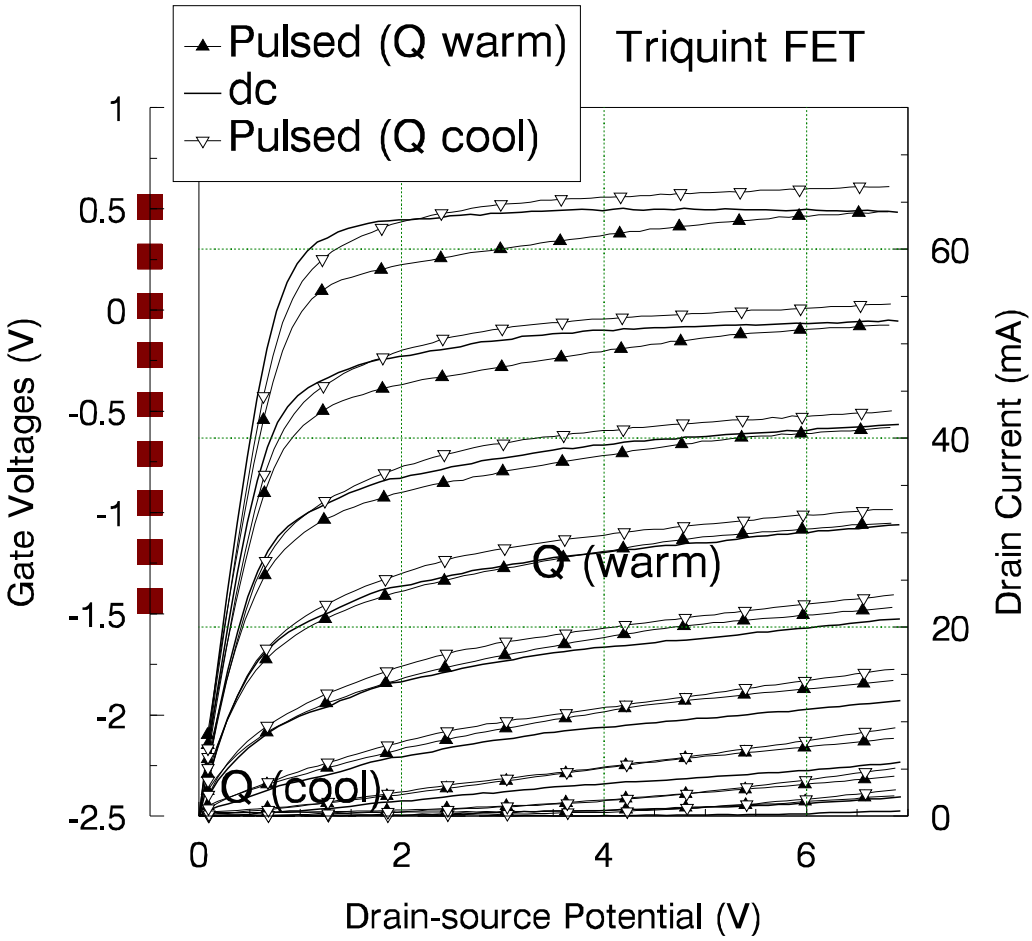

Figure 6.14: Characteristics of a Triquint FET measured under short pulsed conditions from two different bias points, compared to dc characteristics. Capital Qs mark the bias points.

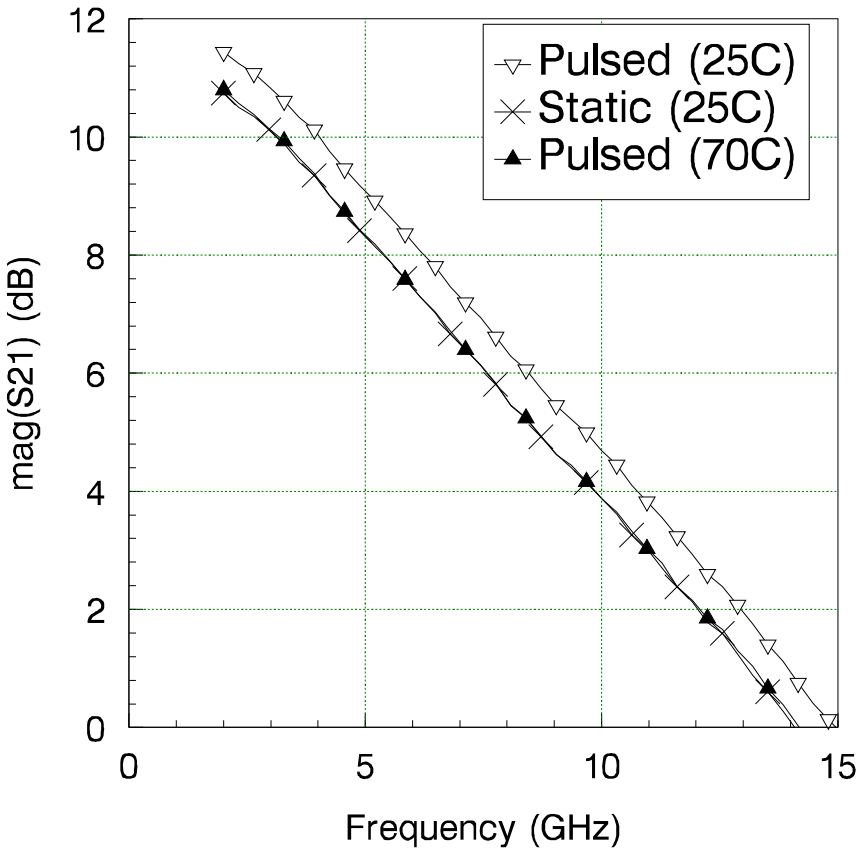

Figure 6.15: Magnitude of FET S_{21} measured under static conditions, and isothermal (pulsed) conditions with two substrate temperatures.

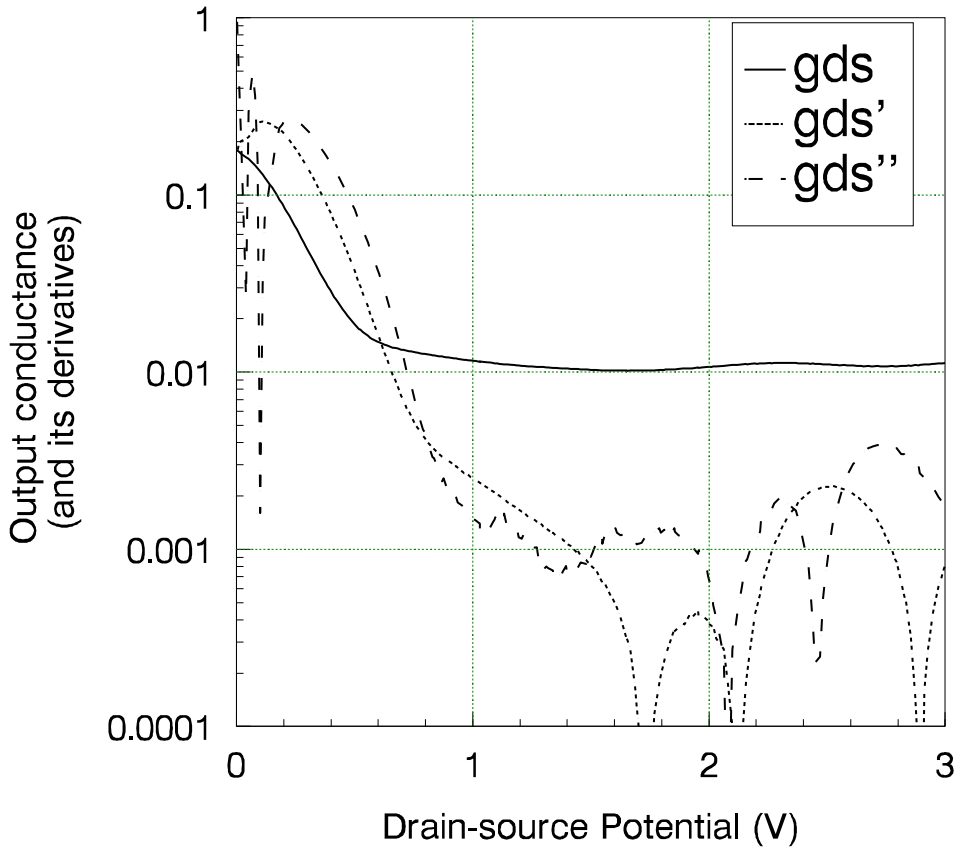

Figure 6.16: Plot of output conductance (g_{ds}) and its first (g'_{ds}) and second (g''_{ds}) derivatives, measured on a Monofast GaAs FET.

6.3.8 Derivative Extraction

Design and characterisation of circuits with attention to IMD performance involves characterisation of the derivative structure. (This will be pursued in greater depth in the next chapter.) Such analysis is directly possible using the APSPA system. Figure 6.16 shows the derivative structure for a Monofast FET, determined entirely from its pulsed-I/V characteristic.

Of interest is the behaviour of the derivatives for low drain bias. These are highly characteristic of GaAs FETs, and have been used as a tool to gauge the suitability of equations in device modelling.[97] The potential saving of effort offered by this extraction of derivatives from PIV data will be clearer after the work reported in section 7.3.

Chapter 7

Gain-derivative Surfaces

7.1 Introduction

The traditional measures of distortion—THD or TDFD for audio systems or IP3 for RF ones, for instance—are scalar, i.e., they are single real numbers. In no field is the inadequacy of this as a gauge of linearity performance so obvious as in audio: Amplifiers with inferior specification often do sound better than ones with superior specification. Perhaps this should not be surprising: Distortion is multivariate and multidimensional. It can depend upon frequency, load and source impedances, perhaps other factors, and it always depends on signal level. It can be predominantly even compared to odd, or low compared to high in its order. A multivariate encapsulation of distortion seems a logical progression.

A Gain-derivative Surface (GDS) is a 3-dimensional plot of a linearity-related measure against two independent variables. A set of GDS permit the disposition of various aspects of nonlinearity to be traced with respect to two independent factors. In this chapter it will be shown that GDS visualisations allow insightful conclusions to be drawn and useful designs to be achieved.

7.2 The Valve Amplifier Dilemma

Western Electric (AT&T) is reopening its vacuum tube plant, Westrex Tubes, in Kansas City, that closed in 1988.[105] Westrex state that "most high-end manufacturers offer a valve product". Recent interest in thermionic valves has grown to a level where factories are being reopened, while manufacturers offer equipment boasting the use of valves, and commanding a great price. Interest in the design of valve equipment has lead to the appearance of CAD models of valves.[10] One company offering a SPICE product reports a significant number of enquiries for valve models. The commercial response, and the publication by society fellows of papers concentrating on valve circuits, supports an old adage of audio enthusiasts: Valve equipment has something to offer.

It seems likely that the audible character of valve equipment results from a combination of its frequency response and its distortion habits. Groups of four GDS, showing fundamental, second- and third-order harmonic contributions, and THD, with frequency and level as the independent variables, will be used as a tool to search for an association between "valve" character and scientific measurements.

7.2.1 Measurement System

A large number of measurements are required to give reasonable resolution in a set of GDS plots: For instance, it is effectively necessary to perform 650 spectral measurements to achieve only 10 points per decade across the audio band, for 40dB of input range. A more comprehensive set might involve 1200 frequency/level points. Acquiring such extensive measurement sets is only feasible with bus-controlled instruments under computer control. Figure 7.1 is the block diagram of the system used to make the measurements. A programmable function generator with low distortion characteristics, an HP33120A, and an HP3561A Dynamic Signal Analyser (DSA) used as a digital spectrum analyser, were used.

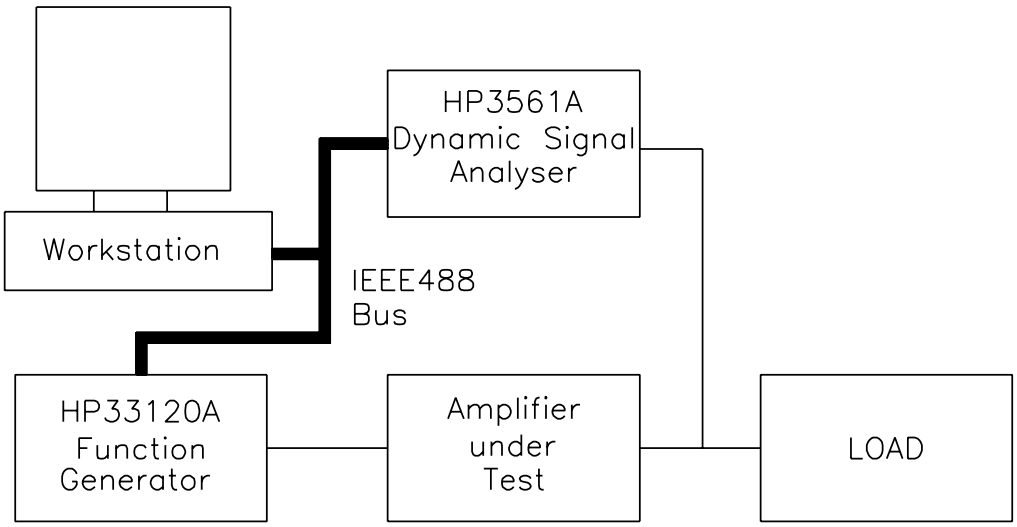

Figure 7.1: Block diagram of the gain-derivative automated measurement system.

The measurement residuals of the system are presented in figure 7.2. The data was gathered without averaging over and above the Fourier transform itself, and took approximately four hours of measurement and a further few minutes of numerical processing on a 486 PC. The dynamic range is in excess of 75dB (typically 78dB, limited almost simultaneously by generator residuals and the DSA dynamic range) for all levels and frequencies of interest.

Sets of GDS will be presented for several amplifiers, and conclusions drawn.

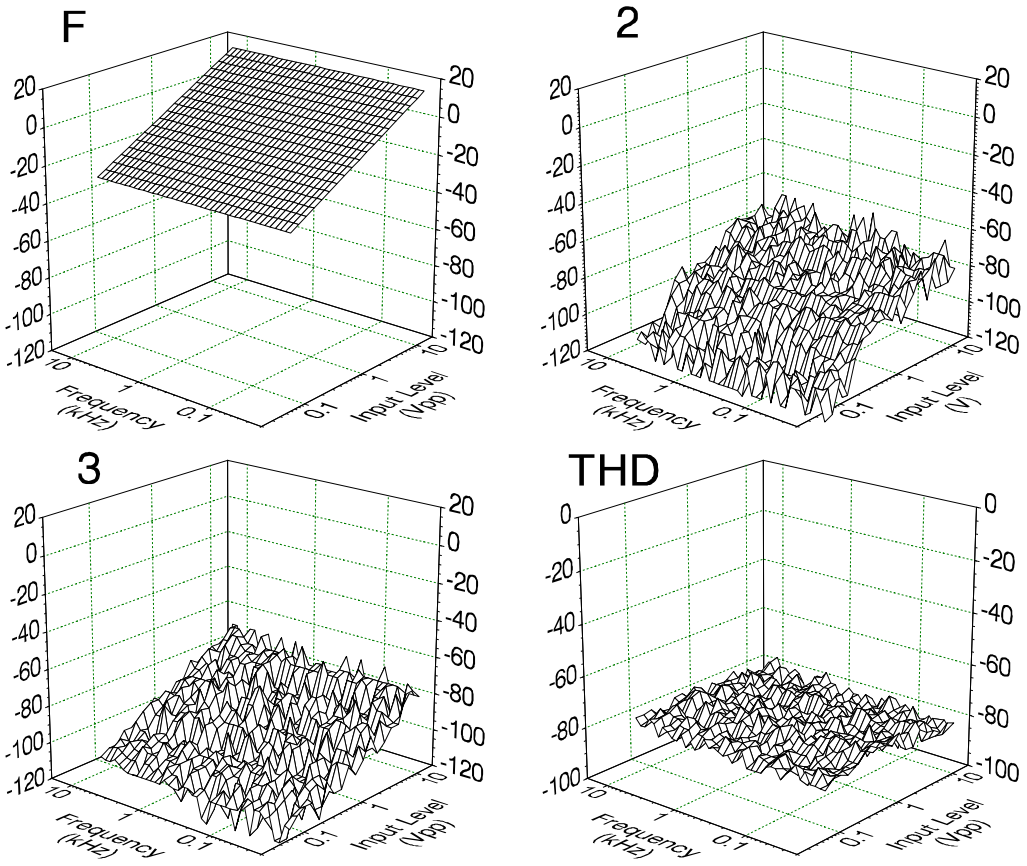

Figure 7.2: A set of gain-derivative surfaces with no DUT in place. The four surfaces depict the fundamental amplitude, the amplitude of second- and third-order components, and the THD.

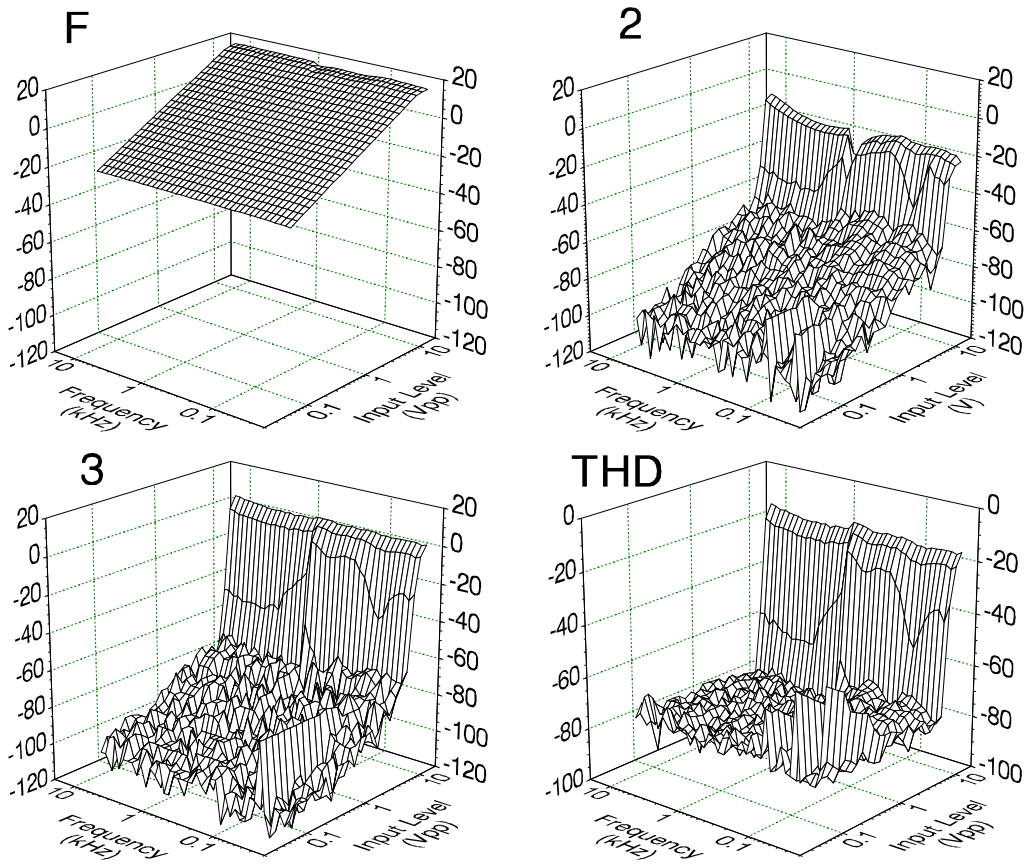

Figure 7.3: Fundamental, second-, third-order and THD surfaces for a Playmaster "Sixty-Sixty" with a loudspeaker load.

7.2.2 Amplifier Tests

A number of amplifiers have been measured, and sets of GDS plotted. Note that the scales have been maintained constant in all the figures for ease of interpretation. Some tests required 10dB of attenuation before and/or after the device under test (DUT), and this is not allowed for in the graphs: It is the shape, not the absolute value (gain or output level of the amplifier) that is important in the comparisons.

Figures 7.3 and 7.4 show GDS for a dc-coupled, all-solid-state, kit-form amplifier of relatively recent design, a Playmaster "Sixty-Sixty".[70] For the measurements of figure 7.3, a load more representative of a loudspeaker has been used, in comparison to the resistive

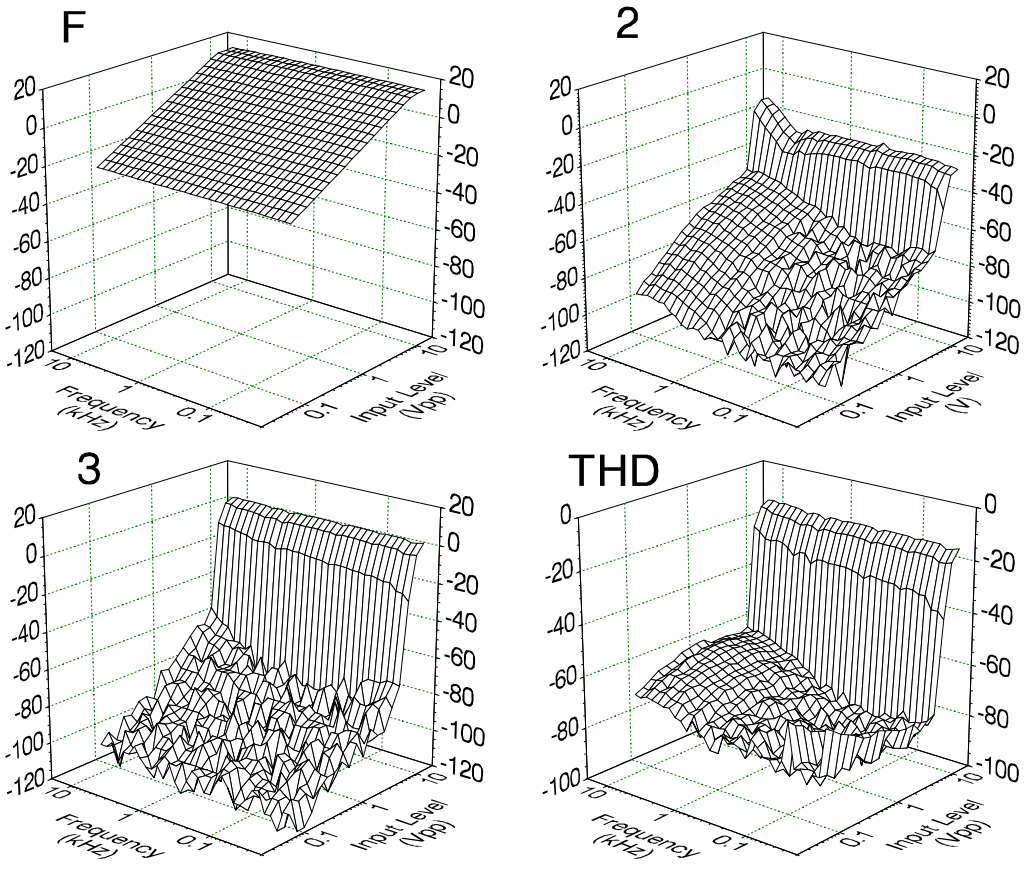

Figure 7.4: Fundamental, second-, third-order and THD surfaces for a Playmaster "Sixty-Sixty" with an 8Ω resistive load.

load more commonly employed. The simulated load was developed by the author[71] and is described in Appendix D. There is a great deal of information in each figure.

The thin "ridges" that appear for low levels and low frequencies are caused by mains-frequency hum entering the circuits from unregulated supply rails. The most obvious character of the surfaces is that they lie substantially at the noise floor up to the point of clipping, at which there is a sharp rise (a "waterfall" in appearance). Next note that the nonlinearity rises out of the noise with increasing frequency in the case of the resistive load, figure 7.4. This is less prominent in the case of the driver load, since the magnitude of the impedance of the driver load rises at higher frequencies. It is possible that this effect arises because this amplifier has an unusually low dominant compensation pole frequency, and thus declining desensivity with frequency, and also because it has an open loop gain that is dependent upon load impedance. (Both are common in modern solid-state designs.)

A most interesting observation is that nonlinearity does not fall away with level. The simple theory of chapter 2 says that second-order nonlinearity ought to fall at twice the rate that fundamental signal level reduces, and odd-order three times as fast. (i.e., a 3dB reduction in fundamental level should result in 6dB fall in second-order distortion and 9dB in third.) Such behaviour as observed here denies the possibility of a simple, low-order-dominant model for the nonlinear transfer function.

The amplifier associated with figures 7.5 and 7.6 is a transformer-coupled, all-solid-state, commercial design manufactured in the mid-1960s, a Monarch SA-400. This is another solid-state circuit, to set the stage, but of the sort common when transistors first became commercially used. Here several interesting characteristics appear.

Firstly, loading with a frequency-dependent, reactive impedance has minimal effect upon gain. However, it affects the nonlinearity performance a great deal. Most strikingly, the GDS for individual harmonics are quite convoluted, suggesting that the distortion is not at all "simple". There is a small dip in the fundamental surface for low frequencies and high levels. This may be introduced by the transformers. Finally, the overall level of

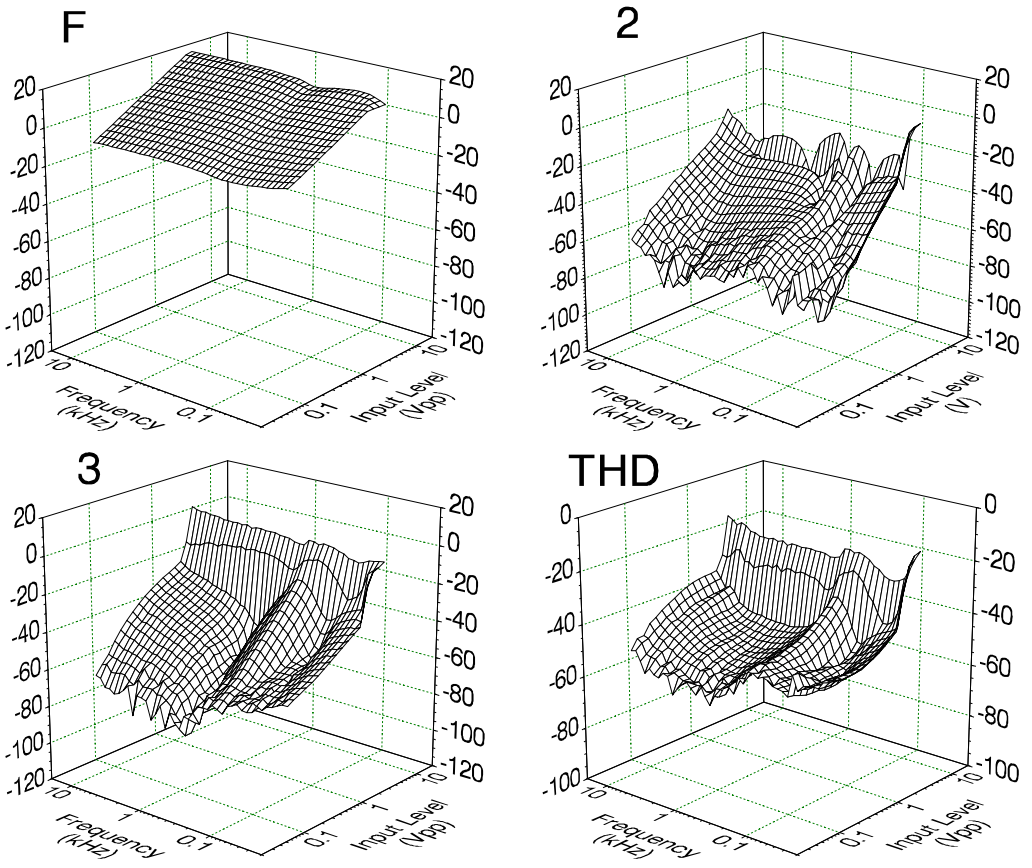

Figure 7.5: Fundamental, second-, third-order and THD surfaces for the Monarch amplifier with a loudspeaker load.

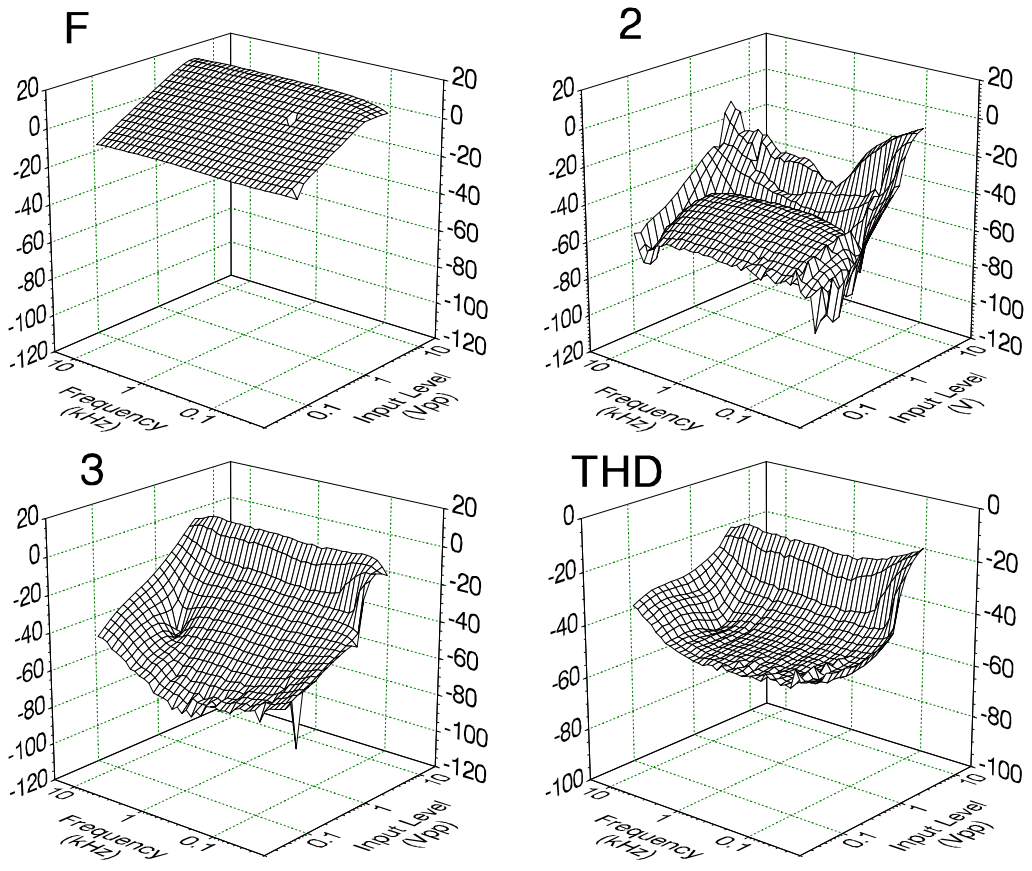

Figure 7.6: Fundamental, second-, third-order and THD surfaces for the Monarch amplifier with an 8Ω resistive load.

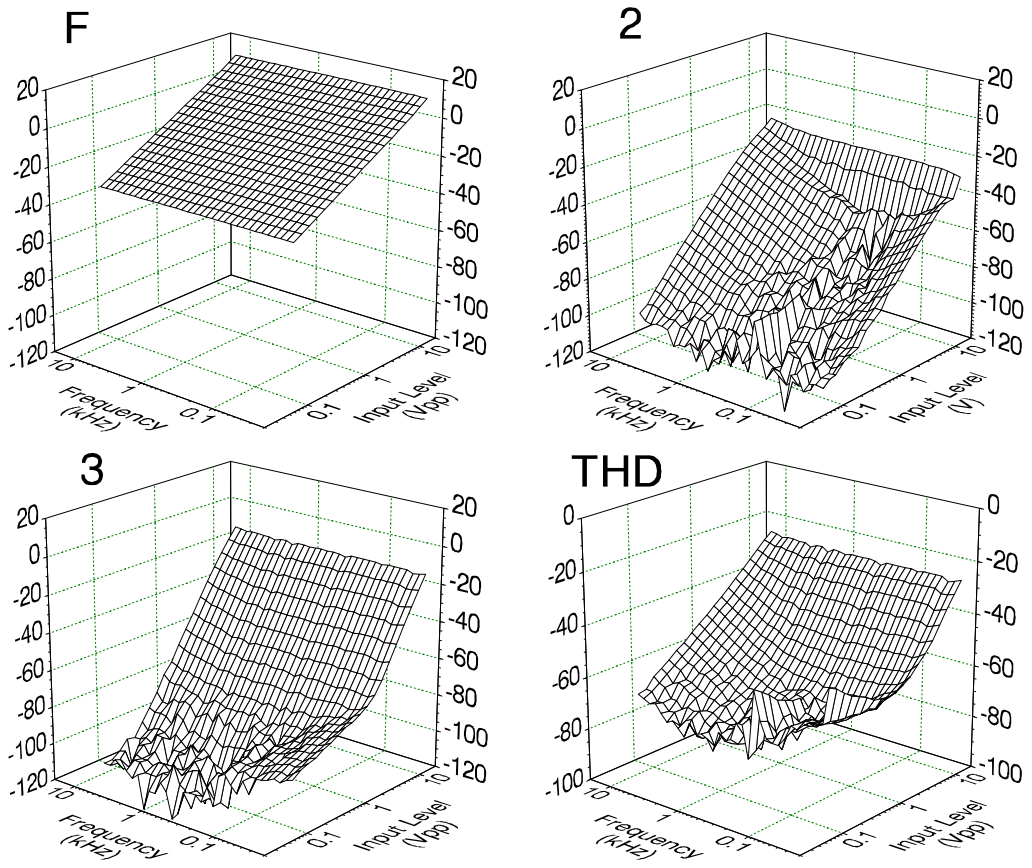

Figure 7.7: Fundamental, second-, third-order and THD surfaces for the Lenard prototype amplifier with a resistive load.

distortion is relatively high, often near 1%. This is characteristic only of early solid-state designs, such as this one.

The Lenard prototype associated with figures 7.7, 7.8, and 7.9 is a high-fidelity, domestic-market, valve-FET hybrid amplifier of 1995-vintage design. With resistive load, figure 7.7, the amplifier is relatively linear (0.01% to 0.2% in the working region), despite the design having no *overall* feedback,[1] and the second- and third-order surfaces are smooth and possess the slopes expected for systems with dominant low-order terms in the nonlinear transfer function. It is interesting to note that the second-order surface has a distinct minimum for mid-band frequencies, rising at 20Hz and 10kHz, but that this is totally absent from the odd-order surface.

[1]Feedback is confined to individual stages.

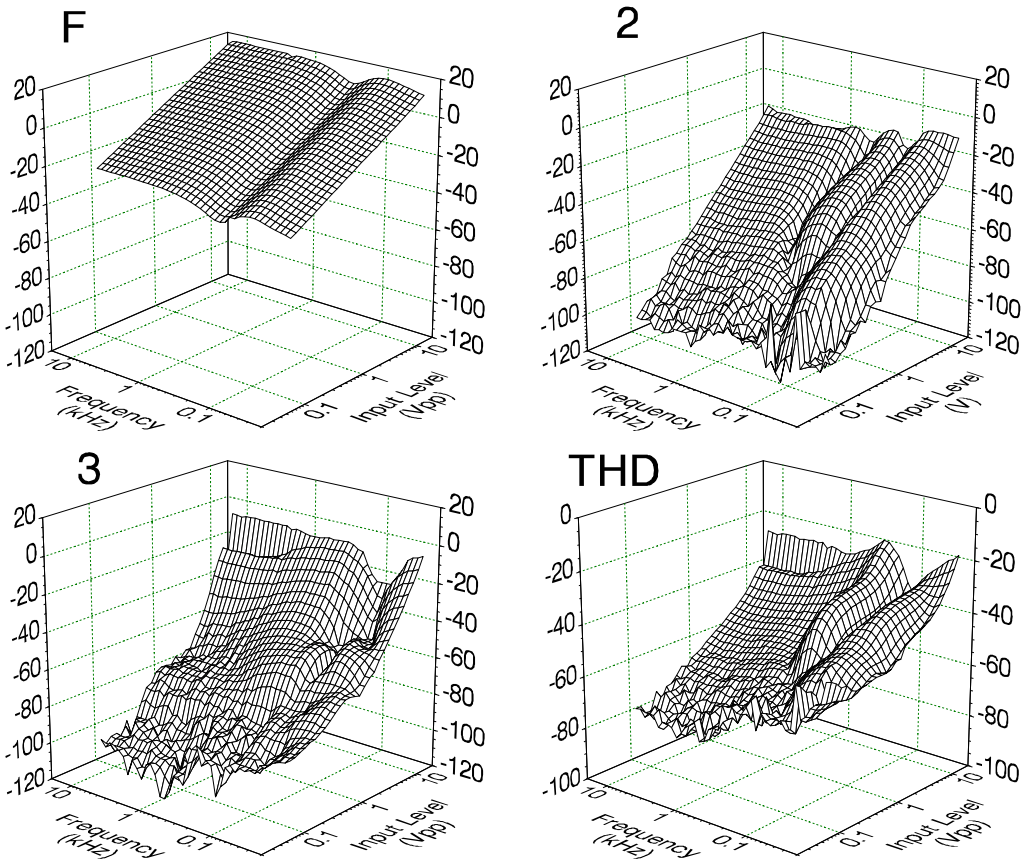

Figure 7.8: Fundamental, second-, third-order and THD surfaces for the Lenard prototype amplifier with a loudspeaker load.

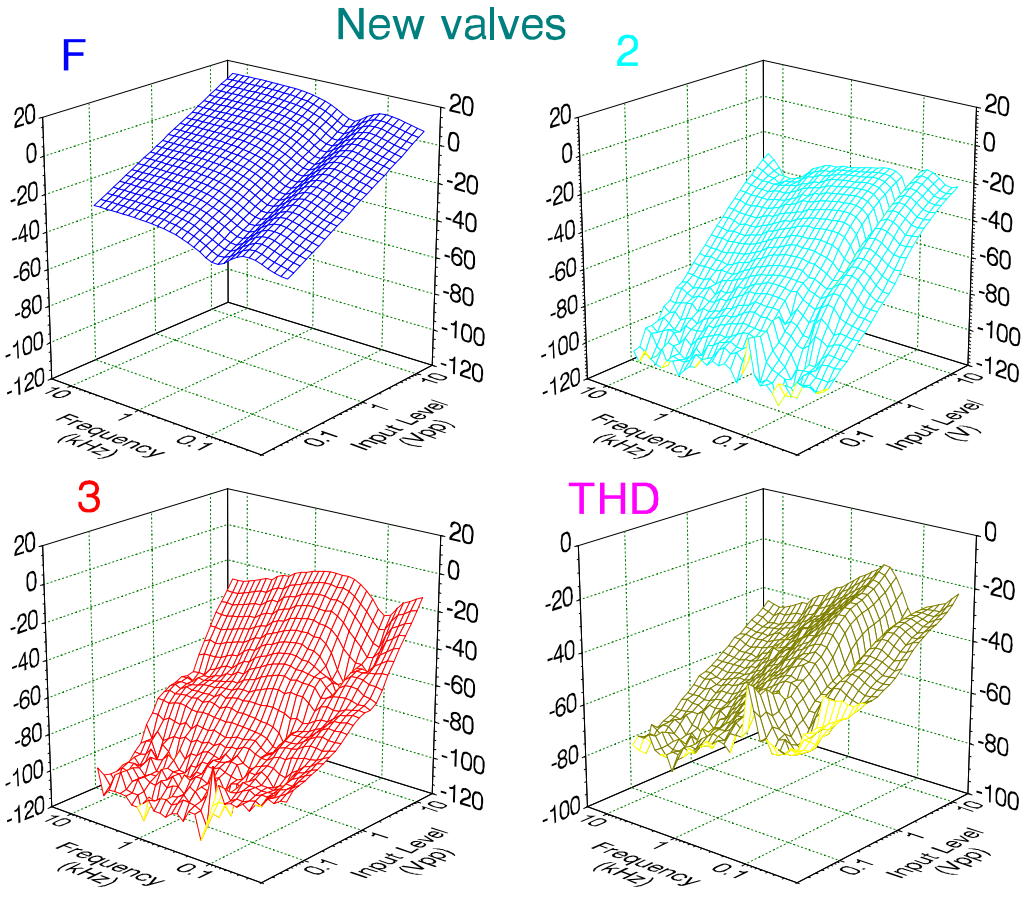

Figure 7.9: Fundamental, second-, third-order and THD surfaces for the Lenard prototype amplifier with a loudspeaker load, after replacement of the valves.

With the reactive test load, the effect of output impedance appears. The second-order surface acquires a pair of "rivers", but the third-order surface acquires a convoluted appearance. Comparing figure 7.8 with the later figure 7.21, note the difference in effect that the *same* reactive load has, especially in the case of the second-order surface. It is evident that there is a complex interaction occurring with the load, and that it differs with different circuits and for different orders of distortion.

Figure 7.9 measures the amplifier identically in all respects to the case for figure 7.8, except that the valves were replaced with new ones of the same model number but from a different manufacturer. Strikingly, the second-order surface appears with only one river, where before it had two. It also drops a few dB in level. The third-order surface becomes marginally smoother, and loses the sharp onset of clipping, suggesting an increase in output power capability.

Figures 7.10–7.13 are associated with two VTA all-valve "monoblocks" (single channel power amplifiers), rated at 30 and 80 Watts respectively. These are recent, Australian designed and manufactured commercial products.

There is a difference in the disturbance produced with the loudspeaker-like load, and the third-order surface does not quite possess twice the slope of the second-order one, and one amplifier obviously has greater output capability. However, both have "smooth" or "well-behaved" surfaces, steadily declining with level, and having the same slope at all frequencies. Note the overall similarities in this respect between these two units and the Lenard prototype amplifier.

An all-valve, integrated (pre- and power), stereo amplifier of recent English design and manufacture, an Audio Note model OTO SE, has GDS shown in figures 7.14 and 7.15. This is a more established, more widely-marketed, valve, high fidelity amplifier. The GDS are *remarkably* similar to those of the VTA30 amplifier. Clipping onsets a little more gently in the case of a resistive load, and the lowest level of THD is slightly higher, but these are the only noticeable differences.

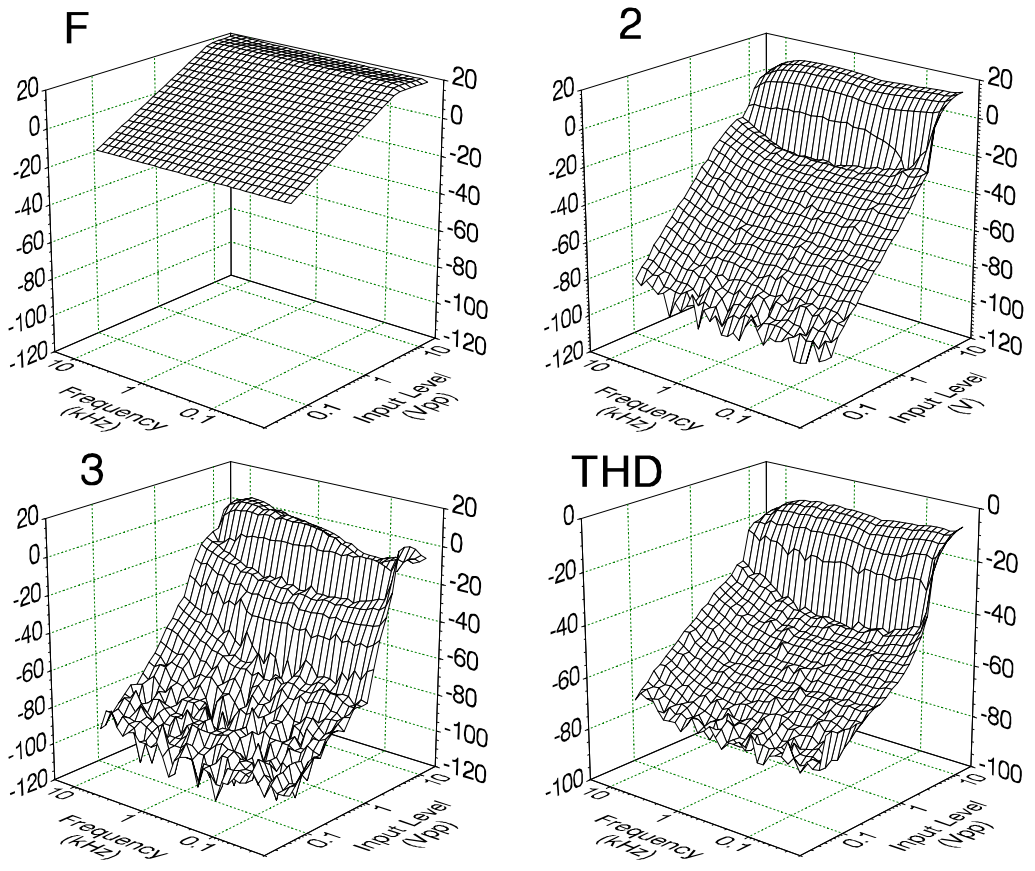

Figure 7.10: Fundamental, second-, third-order and THD surfaces for the VTA 30 amplifier with a resistive load.

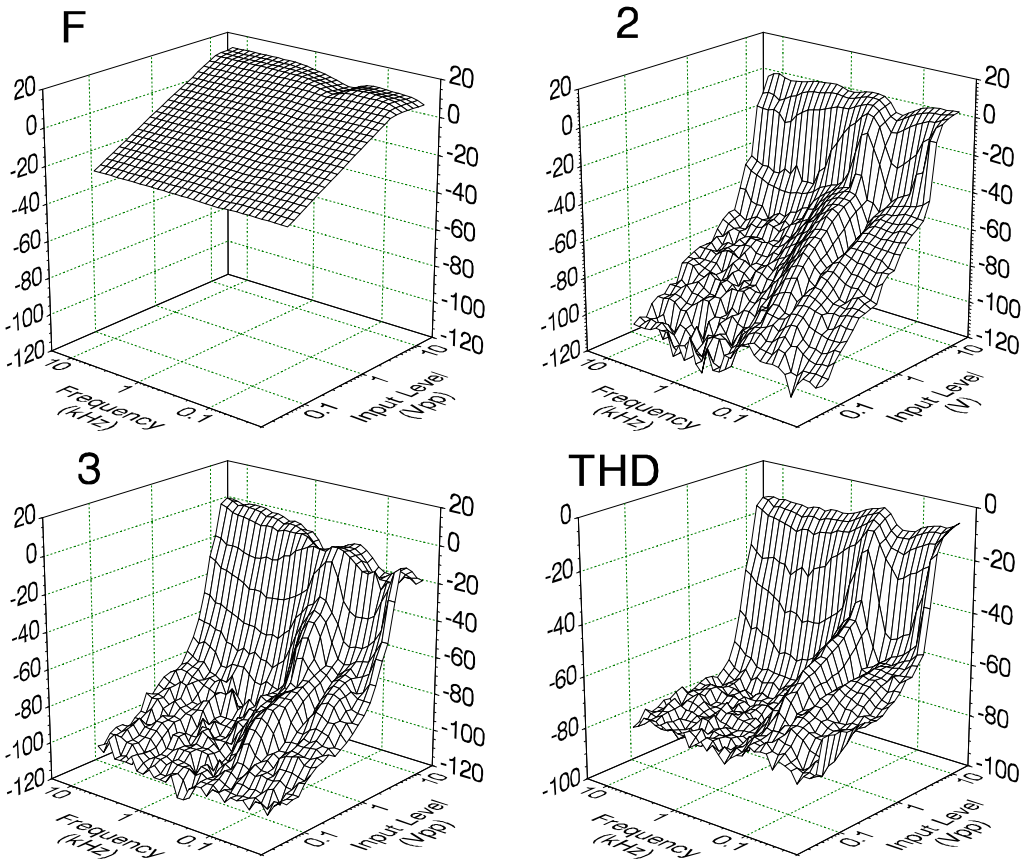

Figure 7.11: Fundamental, second-, third-order and THD surfaces for the VTA 30 amplifier with a loudspeaker load.

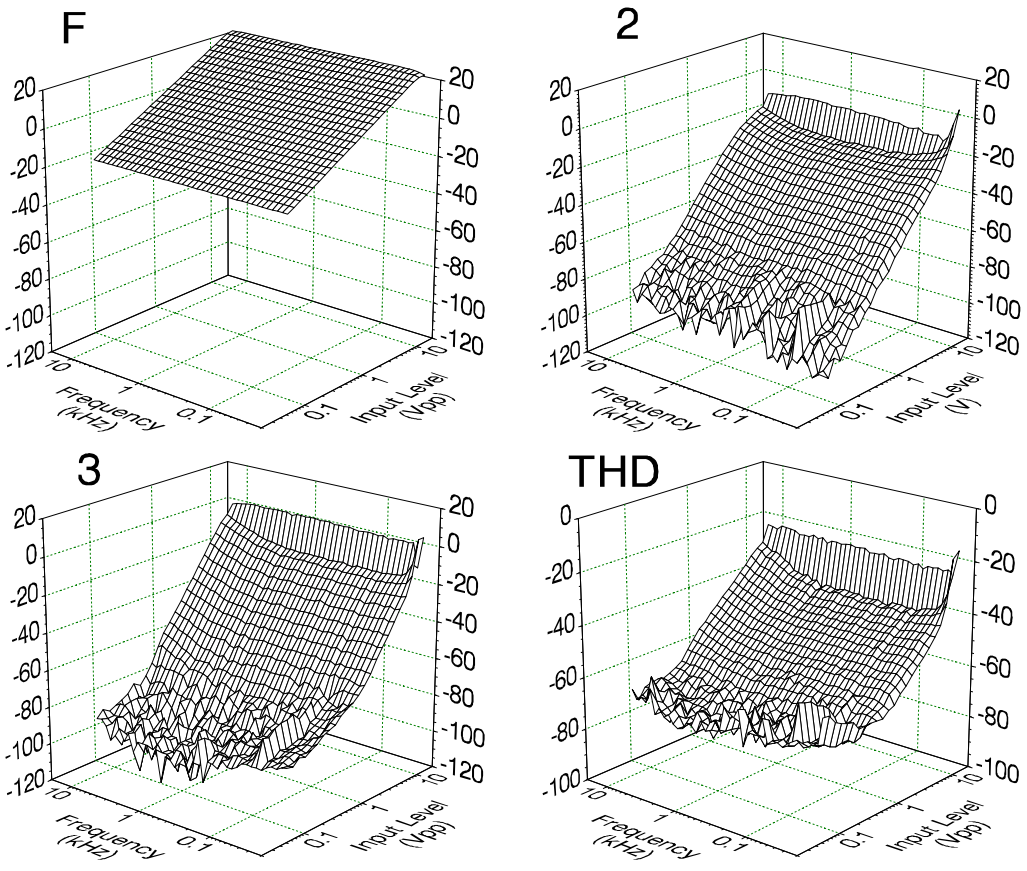

Figure 7.12: Fundamental, second-, third-order and THD surfaces for the VTA 80 amplifier with a resistive load.

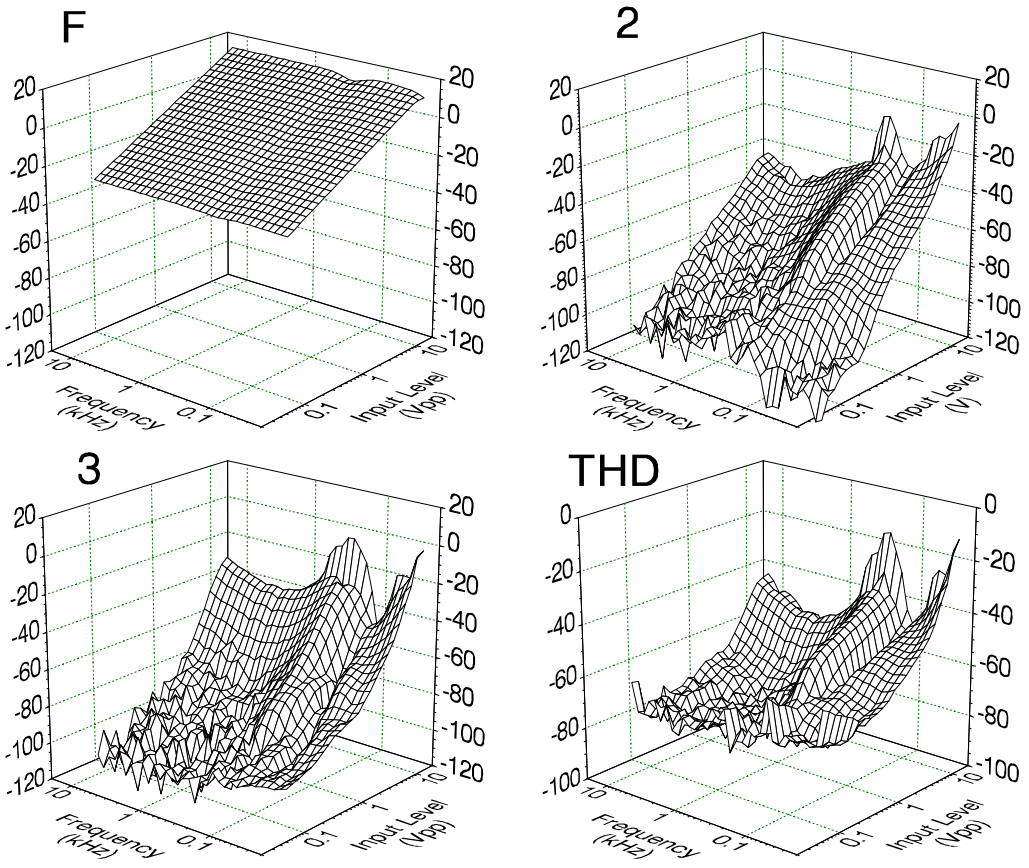

Figure 7.13: Fundamental, second-, third-order and THD surfaces for the VTA 80 amplifier with a loudspeaker load.

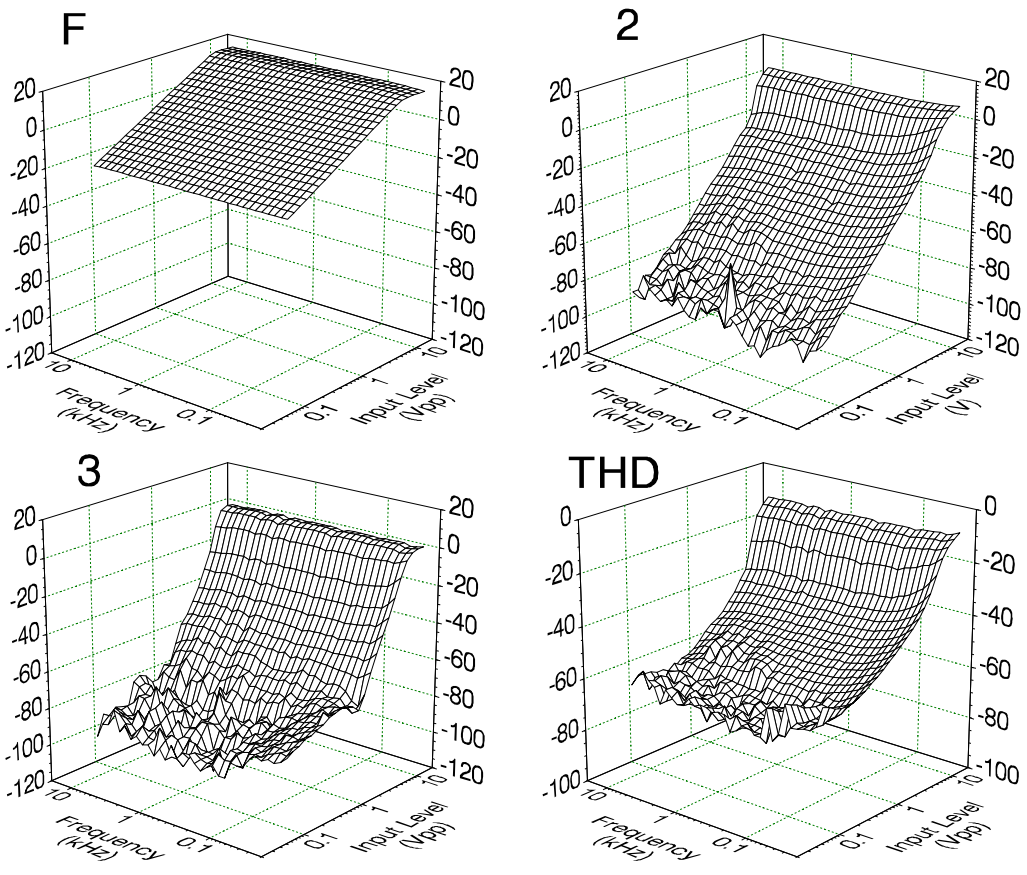

Figure 7.14: Fundamental, second-, third-order and THD surfaces for the Audio Note amplifier with a resistive load.

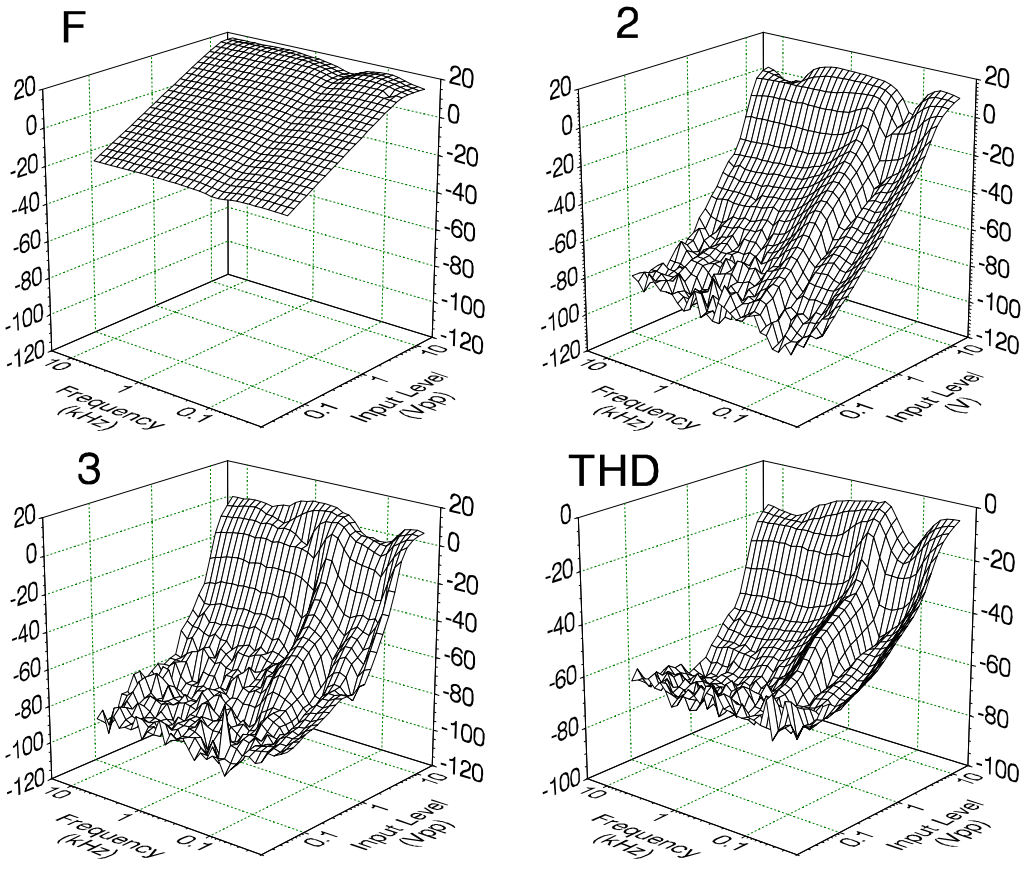

Figure 7.15: Fundamental, second-, third-order and THD surfaces for the Audio Note amplifier with a loudspeaker load.

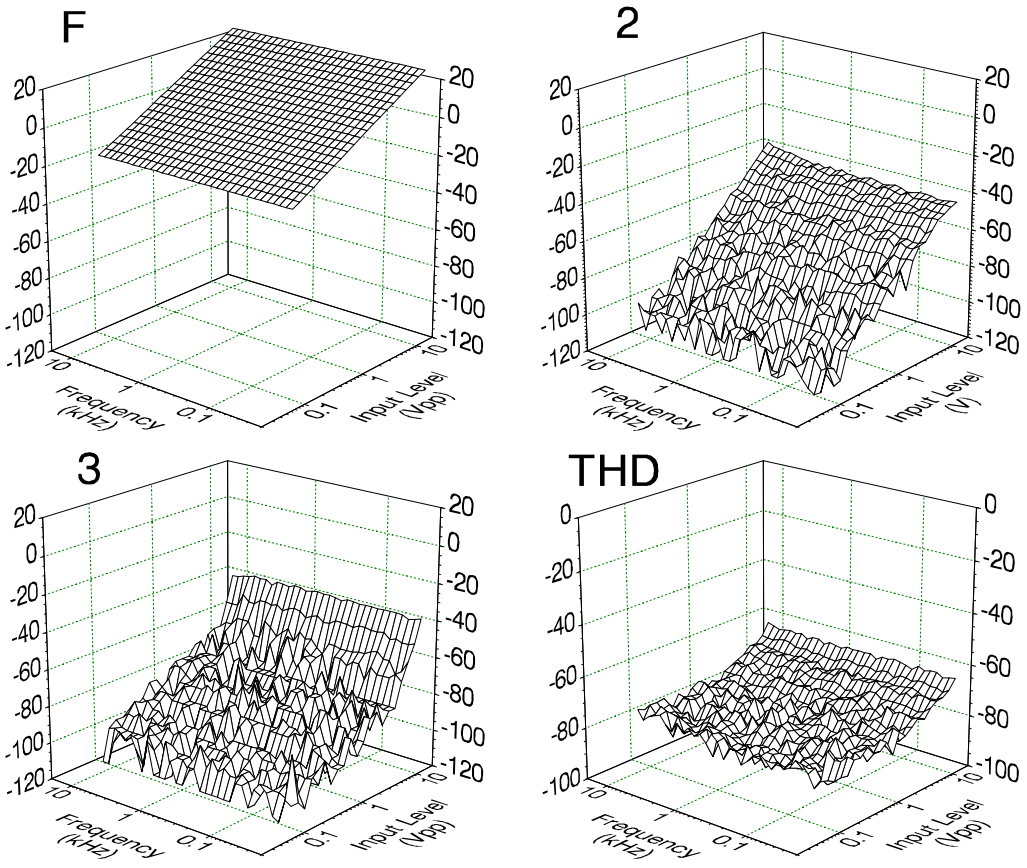

Figure 7.16: Fundamental, second-, third-order and THD surfaces for the Vacuum State preamplifier with a 10kΩ resistive load.

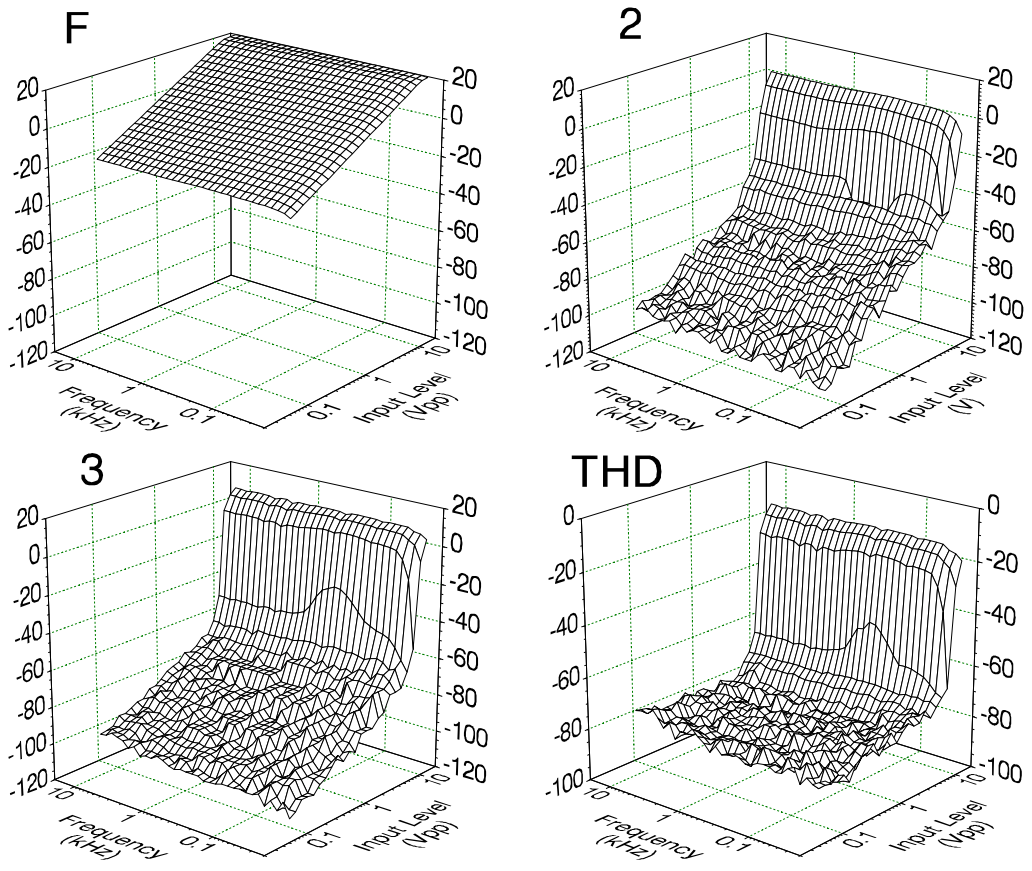

Figure 7.17: Fundamental, second-, third-order and THD surfaces for the NAD preamplifier with a 10kΩ resistive load.

Next, two preamplifiers will be presented. A Vacuum State valve preamplifier of late 1980s manufacture, using solid-state circuits solely limited to ancilliary functions,[2] will be compared to a NAD 1020 solid-state preamplifier of 1980s manufacture. These amplifiers have only one set of surfaces, shown in figures 7.16 and 7.17, since they are preamplifiers and cannot work into a loudspeaker.

With the valve unit, one is struck by the extraordinarily high input and output levels needed to reach clipping, and the very gradual onset of overload. (Indeed it is necessary to look at the spectra carefully to convince oneself that "clipping" has onset at all, and there is 10dB of unreported output attenuation in the figure.) Note next that the distortion that is visible falls away with decreasing level, at the rates expected from simple theory, assuming dominant low-order terms in the model, and in contrast with the GDS of the solid-state amplifier shown in figure 7.4. In short, the nonlinearity is "simple".

The NAD is a comparable-vintage preamplifier. The two sets of GDS are reasonably similar. Apart from the sharp and relatively early onset of overload, which should not come into play in normal listening, there is a slightly flatter THD surface in the solid-state circuit. (Much of it is below the residual floor, but it is possible to comment on the visible part.) One might speculate that this level-independence is caused by a different characteristic in the circuits handling positive and negative excursions, more associated with class B over class A operation than with the state of matter in the amplifying devices.

7.2.3 Conclusion

It would be foolish to rely upon conclusions drawn from so brief a study as reported here. Nevertheless, the techniques presented here serve to reveal much about an amplifier in a quick glance, and some observations can be made.

It has been observed that solid-state amplifiers exhibit features absent from valve designs, and vice versa. In particular,

[2] i.e., there are no transistors occur in the signal path.

- vacuum-tube amplifiers are characterised by soft clipping;
- vacuum-tube amplifiers show nonlinearity smoothly varying with level, apparently limited to low-order effects, giving "well-behaved" surfaces;
- vacuum-tube amplifiers have higher output impedance, and circuit characteristics that make the gain and sometimes the linearity transfer function sensitive to load impedance.

7.2.4 A Pseudovalve Circuit

Figure 7.18 shows one set of GDS for a commercial, all-valve, guitar amplifier manufactured in the late 1960s. This amplifier is distinguished because it is sufficiently popular that it is still known and supported today. It is a 1968, Fender "Bandmaster". The circuit of the power stage that was measured is given in figure 7.19.

Distortion levels are quite high in this amplifier, which has minimal feedback. However, the distortion surfaces do not have sharp convoluted portions. It is easy to guess from the graph that it is a valve design. This amplifier has been used as a circuit model for the next example.

An all-solid-state circuit has been designed by the author, using JFET and MOSFET transistors in the signal path, but with the circuit topology of the Fender amplifier, adapted suitably in order to retain the transformer-coupling and overall level of feedback desensitivity. The circuit appears in figure 7.20.

The pseudo-valve amplifier has its characteristics shown in figures 7.21 and 7.22. Clipping onsets somewhat earlier at low frequencies (below 50Hz), as a consequence of the output transformer. This is visible in the resistive load case. The complex reactive load has its effects on gain, but does not seem to have any effect upon the overall linearity, which adheres to theory where low-order terms dominate. Most significantly, the pseudovalve circuit resembles its predecessor, the Fender circuit.

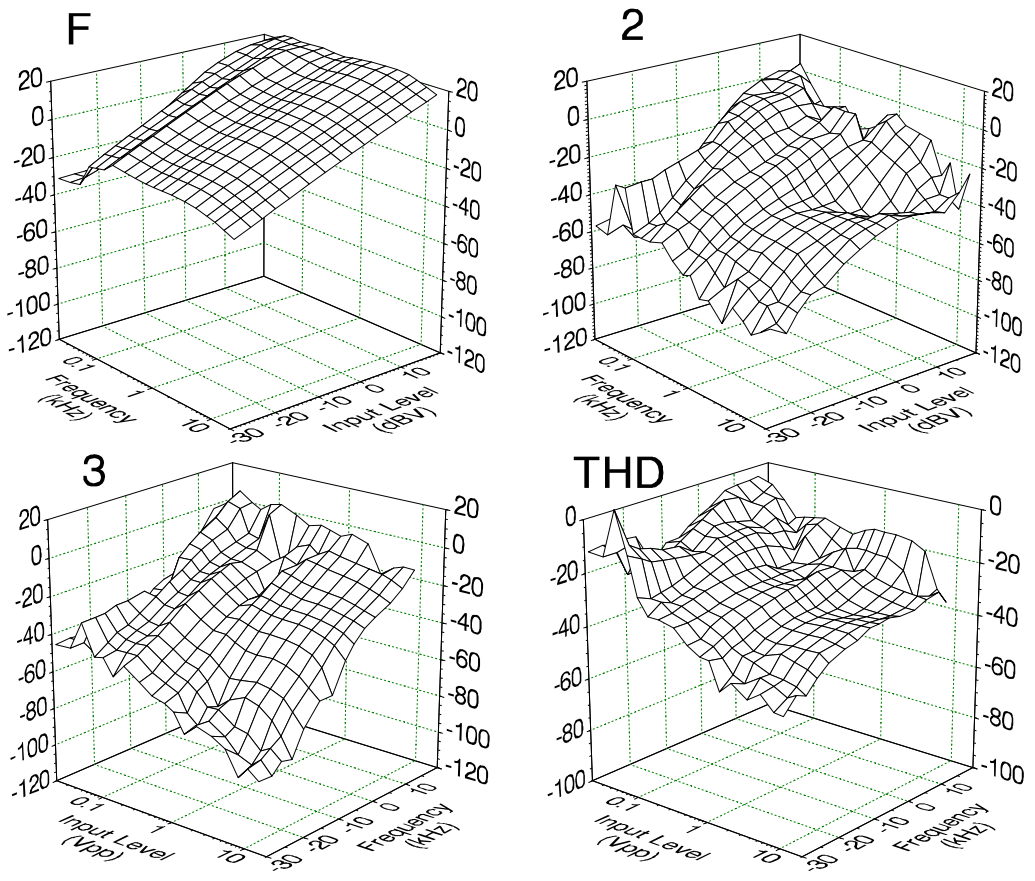

Figure 7.18: Fundamental, second-, third-order and THD surfaces for the Fender guitar amplifier with a loudspeaker load.

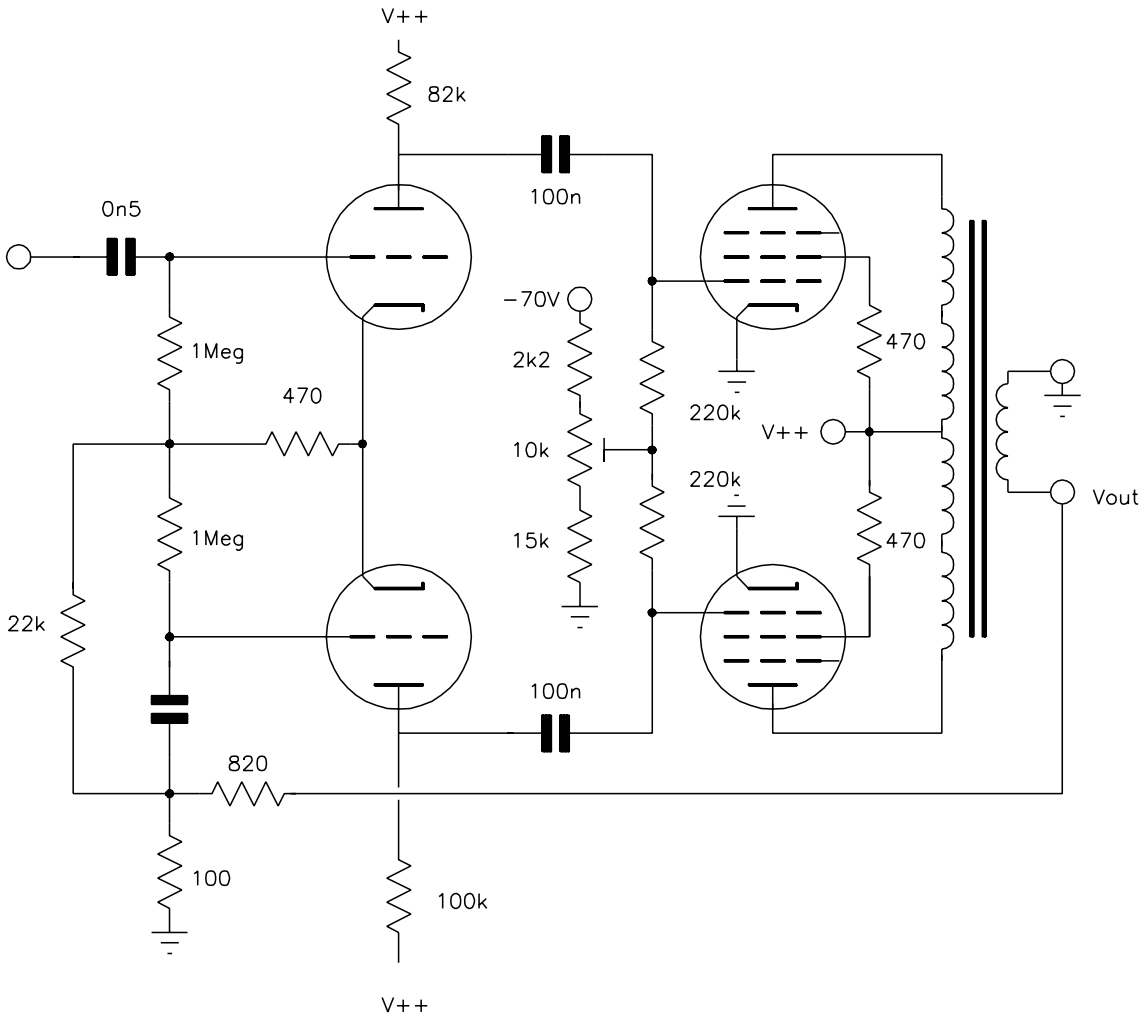

Figure 7.19: Circuit of the final stages of the Fender guitar amplifier.

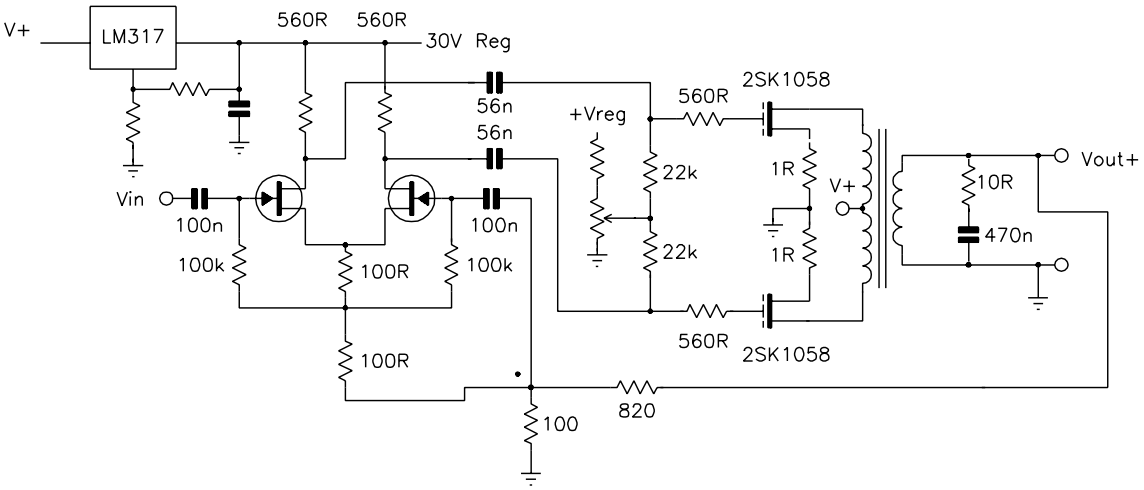

Figure 7.20: Circuit of the author's "pseudo-valve" amplifier, based on the Fender amplifier circuit.

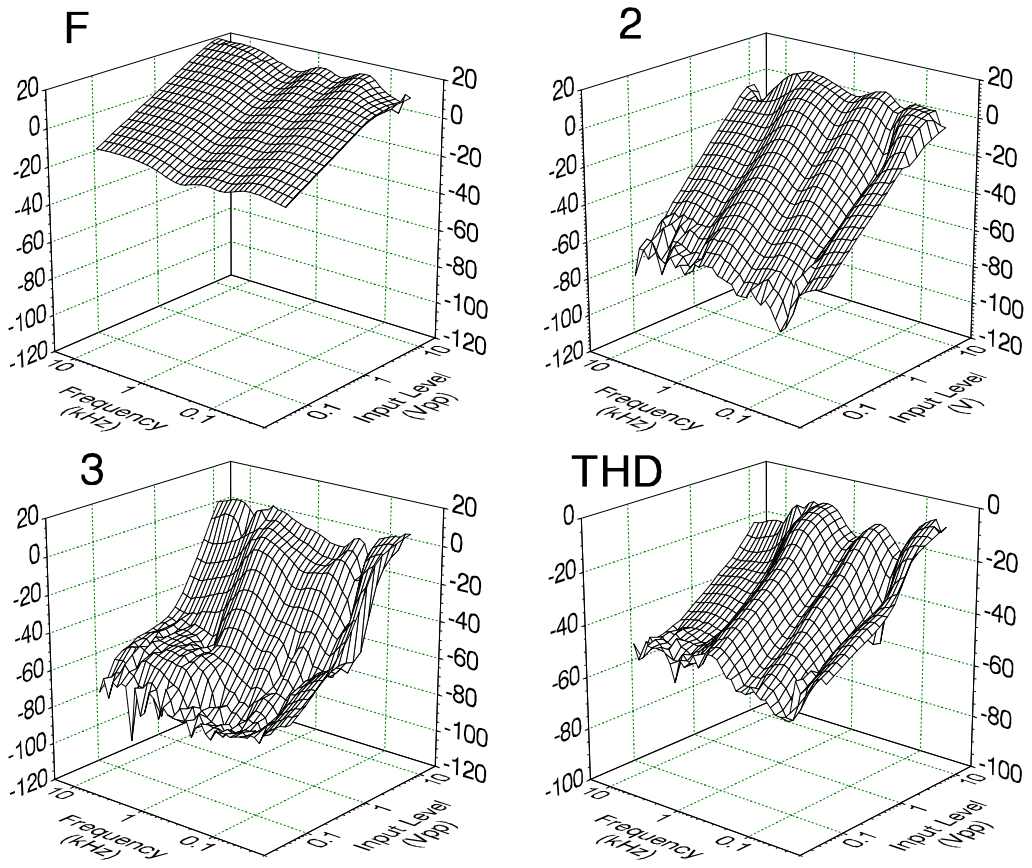

Figure 7.21: Fundamental, second-, third-order and THD surfaces for the author's pseudovalve amplifier with a loudspeaker load.

This example shows that it should be possible to design circuits free of the drawbacks of valves—short lifespan, inefficiency, fragility, etc.—but with characteristics that draw positive comment in subjective assessment.

7.3 GaAs FET Amplifier Design

In GaAs FETs, both the transconductance and the output conductance contribute nonlinear components to the drain current. The author has previously been involved in work where distortions arising from these two mechanisms were arranged to cancel, given appropriate circumstances.[68] The cumulative effect of various mechanisms in a GaAs FET can give rise to a complex variation in nonlinearity with gate and drain bias conditions.

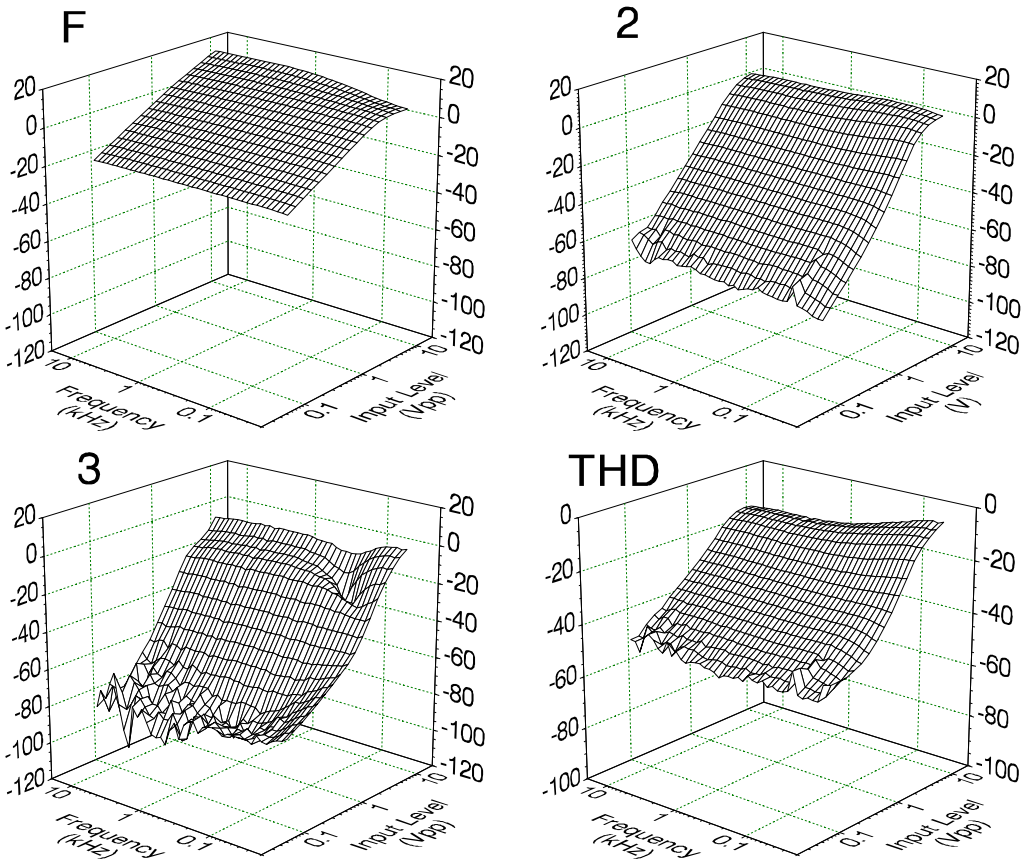

Figure 7.22: Fundamental, second-, third-order and THD surfaces for the author's pseudovalve amplifier with a resistive load.

In this section GDS are used to visualise the nonlinear behaviour, and a novel variation of a solid-state travelling-wave amplifier, offering performance optimised for low distortion, is reported.[106]

The linearity of a MESFET configured as a common-source amplifier may be visually summarised in a set of 3D surface plots with gate and drain bias points as the independent variables, and gain and its odd- and even-order RF-TDFD components as the dependent (z-axis) variables.

7.3.1 RF-TDFD Measurements

It is common for distortion in RF amplifiers to be measured with a two-tone test and expressed as an intercept point, as noted in Chapter 2. The two-tone test lends itself to narrowband systems, since stimulus frequencies and wanted response tones can be confined in a narrow bandwidth. However, it effectively requires a spectrum analysing capability to resolve the result, even if this is simply provided by the narrow-band receiver itself. Although the two-tone test senses only odd-order nonlinearity, this is not a problem as even-order nonlinearity will not produce tones of concern in many such applications, for the same reasons that the test fails to see them. (They fall out of band.)

In the last decade or two, wideband RF systems have proliferated. This opens the way for different tests to be useful. In particular, the advantages of the TDFD test in audio generalise well to the RF arena. A variant of the system described in Chapter 5 has been prototyped and used.[102]

This instrument uses tones at 40.0000MHz and 59.9750MHz, yielding tones of interest at 20.0250MHz and 19.9750MHz. The sources are achieved with relative simplicity by virtue of a filter, based around the bi-notch filter element shown in figure 7.23. The full filter is shown in figure 7.24. The same passive front-end filter scheme, as used in the instrument reported in Chapter 5, is used in the RF-TDFD meter to extend the linear dynamic range of a Spectrum Analyser (SA) to beyond 100dB. In the prototype, the sources were realised

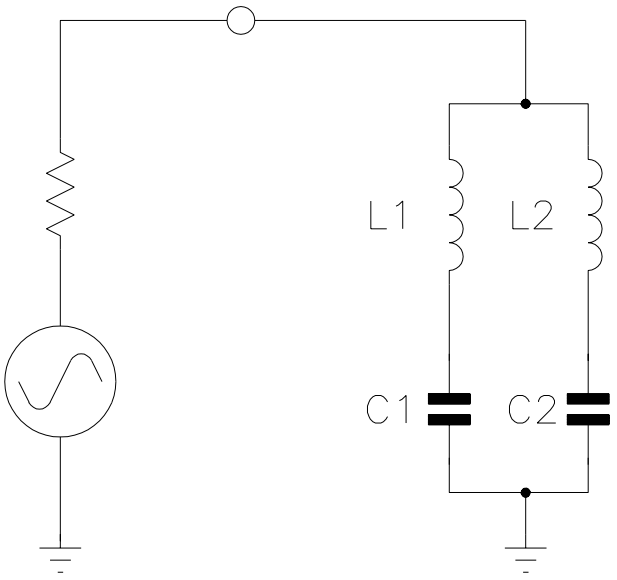

Figure 7.23: Circuit of the bi-notch filter structure used in the combiner filters in the RF-TDFD system.

with PLL ICs locked to crystal oscillators, followed by Mini-Circuits wideband amplifiers, and combiner filters. The sources are provided with buttons to permit momentary pulling by 5kHz, to allow easy identification of the order of tones when viewed singly on a SA display.

The binotch filter circuit of figure 7.23 has two frequencies at which it presents zero impedance, on either side of a frequency where it presents infinite impedance. The two notch frequencies lie at

$$f_1 = \frac{1}{\sqrt{C_1 L_1}} \tag{7.1}$$

and

$$f_2 = \frac{1}{\sqrt{C_2 L_2}} \tag{7.2}$$

while the pass frequency is

$$f_p = \frac{1}{2\pi}\sqrt{\frac{\frac{1}{C_1}+\frac{1}{C_2}}{L_1+L_2}} \tag{7.3}$$

Appropriate selection of component values will place one notch at an unwanted test frequency and the pass-frequency, f_p, at the other, wanted test frequency. (Alternatively, the circuit can be arranged to block both at the price of some attenuation at the wanted, TDFD result-tone frequencies.)

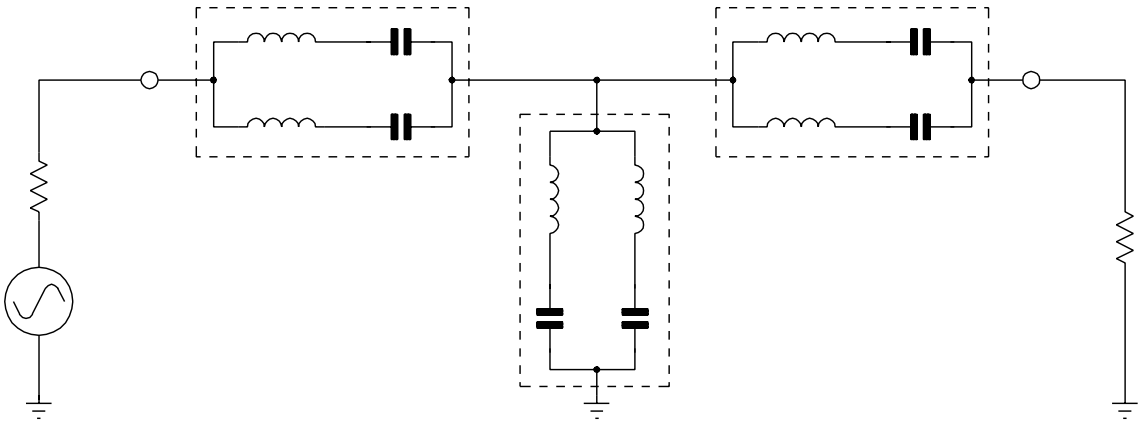

Figure 7.24: Circuit of the combiner filters used to isolate sources in the RF-TDFD system.

The circuit of figure 7.24 utilises a T-network of bi-notch circuits to produce a filter that passes one signal unattenuated while eliminating the other. The parallel bi-notch network is chosen to present low impedance at the unwanted test-tone frequency and high impedance at the wanted test-tone frequency; the series arms are set for the reverse.

RF-TDFD measurements can also be made with lesser dynamic range, but under automated control, using standard instrumentation. The plots appearing in the next section were made under computer control using the setup shown in figure 7.25. HP8657 sources combine with an HP4195 to make an RF-TDFD instrument that is controlled over an HPIB bus. The approach is particularly efficient for measuring GDS because the HP4195 has the ability to measure the magnitude of a single specified spectral component as a function of a dc voltage produced by an integral source. (It effectively uses its own dc source as the display x-axis.) The other independent variable is produced by an external supply of some convenient form. When the dynamic range need only approach 60dB, one may dispense with the isolating and notch filters.

7.3.2 Device GDS

The gain-surfaces of an NE33284 HEMT ($|S_{21}|$ plotted against V_{gs} and V_{ds}), and its even- and odd-order derivative surfaces), obtained from RF-TDFD measurements, appear in figures 7.26, 7.27, and 7.28. This complex kind of behaviour has been observed elsewhere for MESFETs and HEMTs at 2.5GHz.[46]

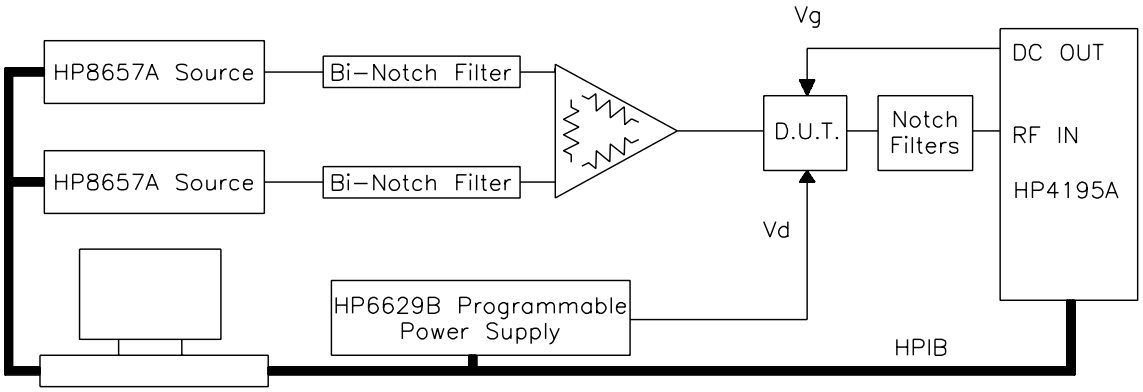

Figure 7.25: Equipment block diagram for automated RF-TDFD measurement.

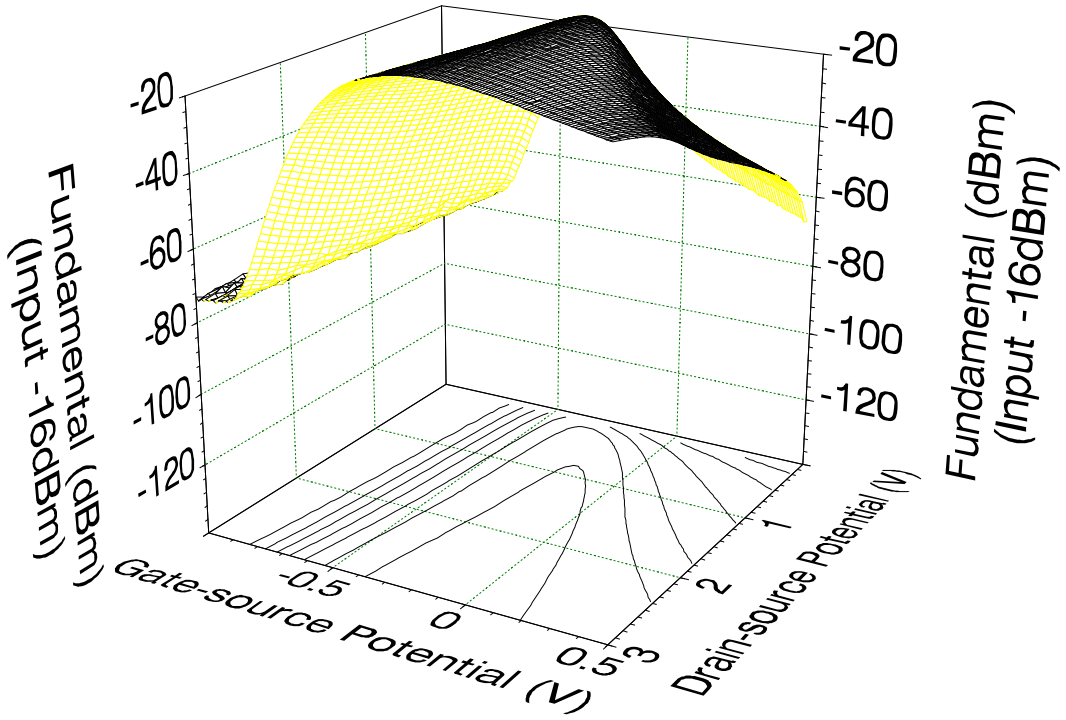

Figure 7.26: Gain surface of an NE33284 HEMT with 50Ω load. A contour plot of the surface is projected onto the XY-plane for clarity.

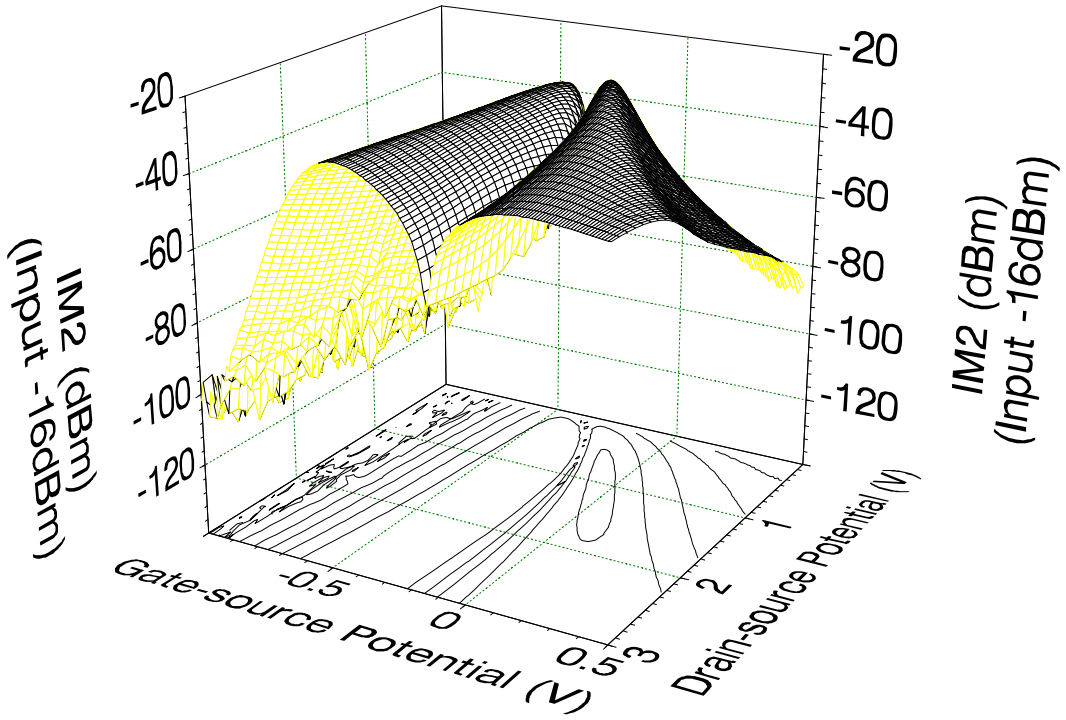

Figure 7.27: Second-order intermodulation (IM2) surface for the same transistor and conditions as figure 7.26.

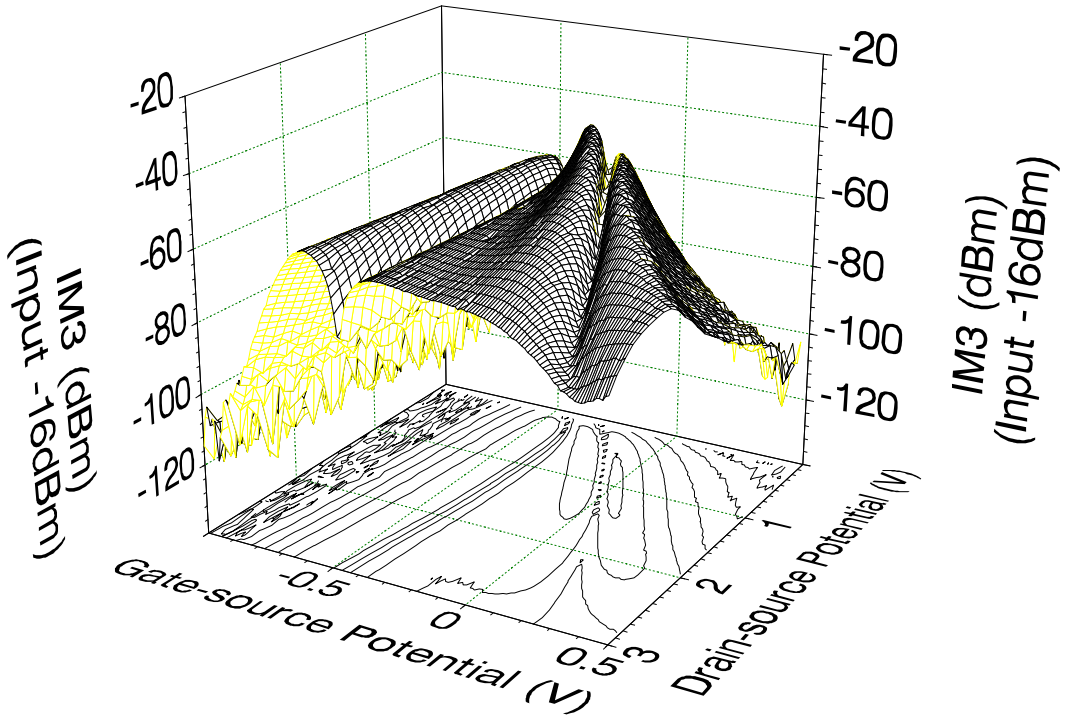

Figure 7.28: Third-order intermodulation (IM3) surface for the same transistor and conditions as figure 7.26.

Note that the surface depicted in figure 7.28 shows two loci of notches, or "rivers", and that of figure 7.27, one. There is a phase reversal in the intermodulation component as such a river is crossed. This is predicted by more advanced nonlinear FET models.[46, 97, 14]

It is now obvious that it should be possible to scale and bias two devices appropriately so as to have either the even or odd intermodulation component cancel in their summed output when they are operated in parallel. In fact, more subtle and useful possibilities can be realised with more than two devices.[41] For instance, simultaneous even and odd cancellation is possible, or maximisation of even with minimisation of odd components for a mixer, etc. The author and colleagues have biassed and combined devices to produce a gain element with greatly improved third-order intermodulation performance.[106]

7.3.3 The Webster Design

A circuit of the form shown in figure 7.29 may be used to produce summed output of several devices of different gate widths and with differing gate bias. The problem of design is then to select the bias voltages and device widths to produce the desired characteristic. The first three curves in figure 7.30, identified with open symbols, show a slice through the three surfaces for a single HEMT, with $V_{ds} = 2.6$V.

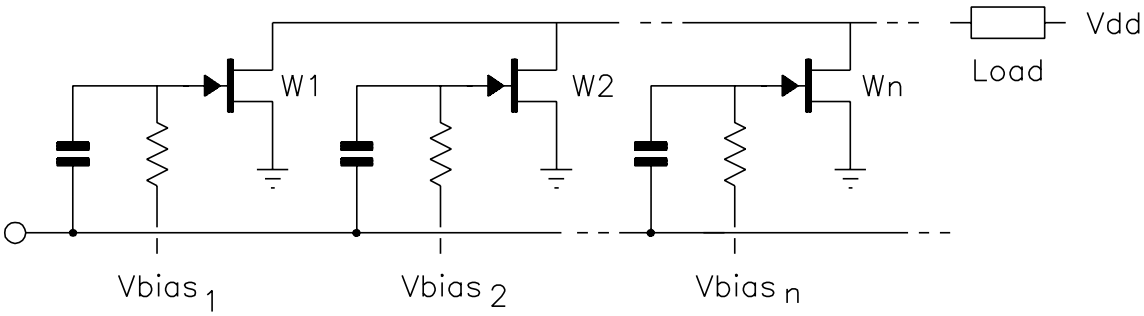

Figure 7.29: A circuit suitable for realising a gain stage with desired distortion characteristics fixed by the derivative superposition method.

Having selected a suitable drain-bias region from consideration of the GDS, the measured variation of the derivatives of a device with gate-source voltage, for a given drain-source voltage, may be expanded into signed linear form in a suitable mathematical package. Scaled and bias-shifted copies of the characteristics are summed together to predict the

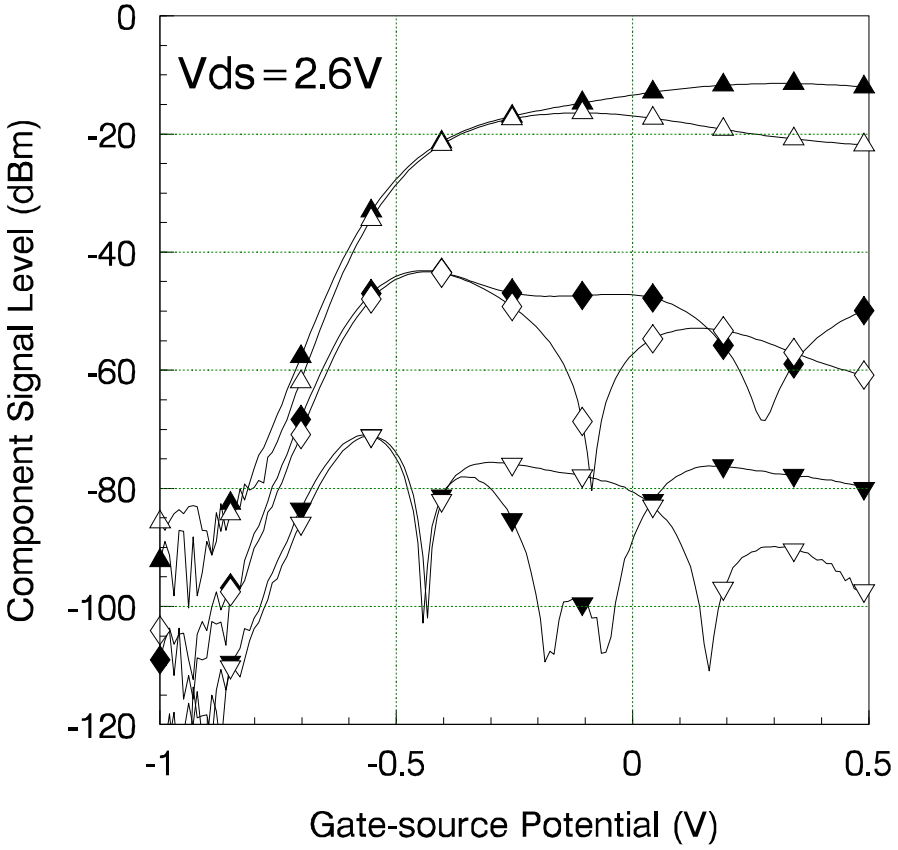

Figure 7.30: Fundamental ($\triangle$), even-order ($\diamond$) and odd-order ($\triangledown$) intermodulation components plotted against gate-source potential for a single device (open symbols) and the designed amplifier (solid symbols).

output of a multi-device circuit. The magnitudes and shifts are varied so as to minimise the undesirable part of this predicted output. The relative magnitudes then define device widths, and the relative shifts are the gate-bias offsets.

Such a combination of device widths and offsets may be determined manually, as opposed to numerically. This approach can yield a solution that is close enough to optimal, and carried out in a number of stages, serves to make the method clear. Initially, one secondary device is added to the first, or main device. Its relative bias is chosen such that its positive (first) IM3 peak is added to the high negative IM3 peak of the main device. The position of this peak determines the left hand margin of the resultant IM3 null. The width of this secondary device is then chosen such that the magnitude of its positive IM3 peak gives a deep null in the resultant IM3, at the position of the secondary device's positive IM3 peak. Next, a third device is added, such that its positive IM3 peak occurs at a higher bias than that of that of the first secondary device, and it produces a second deep null adjacent to the first. If the offset voltage between the two secondary devices is too high, a double null will occur. This relative offset is reduced until the two nulls form a canyon, the peak between nulls at the bottom of the canyon becoming sufficiently small. Further devices may then be added, extending the resultant canyon (possibly requiring minor adjustments to all the secondary device widths and offset voltages). Experimentation shows that trying to place the resultant IM3 canyon too close to the original IM3 minimum of the main device can lead to larger secondary devices and can lead to using more devices for a given result.

7.3.4 Results

A four-device, broadband amplifier circuit verifies the technique. The amplifier has the topology of a solid-state, travelling-wave amplifier, and is thus a good model of a MMIC. It is optimised to produce low third-order intermodulation across a relatively wide input voltage variation. The aim of this is to extend the dynamic range of the improvement. (In the absence of frequency dispersion, input signal can be visualised as movement around a fixed operating point on the X-axis of figure 7.30. Thus one might expect a wide region of low IM3 to preserve the low level of IM3 for higher input signal levels.)

For the devices used and the chosen drain bias voltage, the design requires width ratios of $\frac{W2}{W1} = 0.4$, $\frac{W3}{W1} = 0.4$, and $\frac{W4}{W1} = 0.7$; The main bias is intended to be $V_{gs_1} \approx -0.15$, with gate bias offsets of $V_{gs_1} - V_{gs_2} = 0.37\text{V}$, $V_{gs_1} - V_{gs_3} = 0.45\text{V}$, and $V_{gs_1} - V_{gs_4} = 0.58\text{V}$. Since the devices used were discrete, the width scaling factors were realised by means of π-section attenuators at the drains. Because secondary devices are more pinched off and narrower, power consumption is virtually unaltered from the case of the main device at the same operating point. The total gate width is increased by a factor of 2.5, but power consumption increases by less than 4 percent.

The second group of three curves in figure 7.30, identified with solid symbols, show a slice taken through the three surfaces, but for the parallel-connection of four devices in the new amplifier. The wide region of low IM3 is clearly visible about $V_{gs} = -0.15$. The shape and position of this characteristic is relatively invariant with V_{ds}.

Figure 7.31 is a plot of output power against input power for the combined 4-device amplifier and for a single-device amplifier scaled to deliver comparable power. For this comparison, the single device has been biassed at the IM3 minimum at $V_{gs} = -0.43\text{V}$, the most advantageous point for a single-device amplifier. In the region below the onset of gain compression, the third-order intermodulation performance is visibly improved over a wide spread of power levels. The measurements show 20–30dB improvement over a substantial range.

The IM2 is lower than the case of a single device biassed for minimum IM3, and delivering the same power. However, it is higher than would be the case with a single device biassed at the same point as the main device of the new amplifier, and of the same size. Nevertheless, with such bias on a single device, the IM3 is much worse, and the gain still not as high.

7.4 Conclusion

The usefulness of GDS for visualising and applying nonlinear characteristics has been verified in a second situation. An amplifier has been produced using four HEMTs to achieve

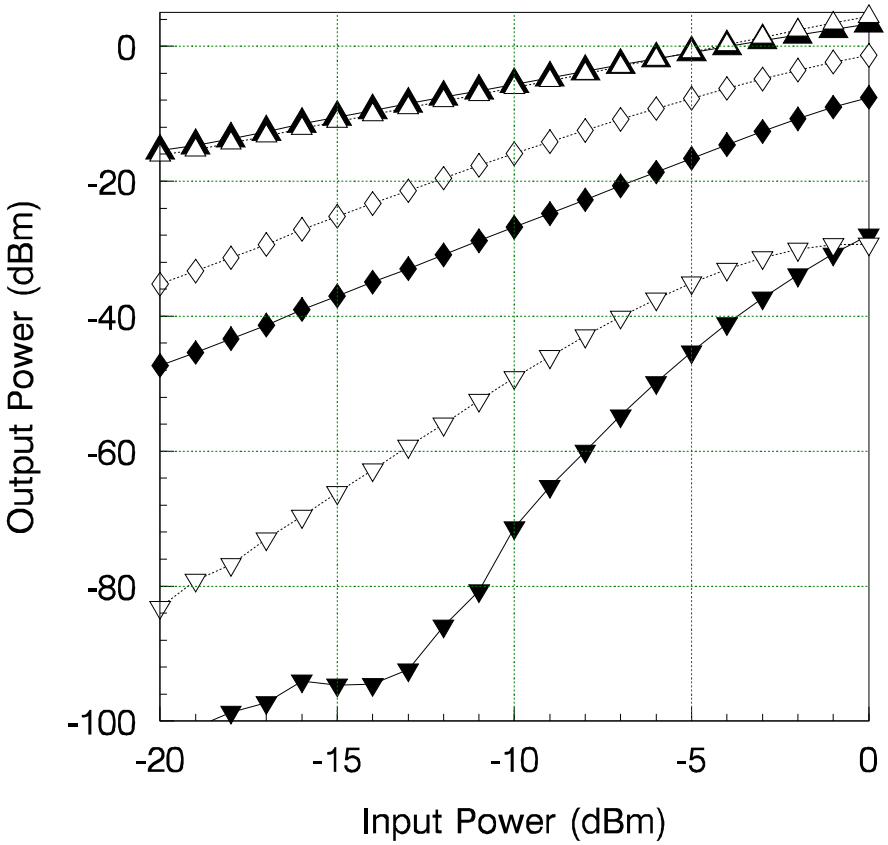

Figure 7.31: Fundamental (△), even-order (◇) and odd-order (▽) intermodulation component power plotted against input power for a single device (open symbols) and the 4-device amplifier (solid symbols).

low IM3 with maximum gain over a wide range of input powers. The technique is especially applicable to broadband MMIC design. The amplifier design method is anticipated to have application in multichannel communication systems where intercarrier interference is of concern.

Chapter 8

Concluding Remarks

Research can continue indefinitely, each step opening more doors. I have chosen to draw the line here. This thesis has presented

- a software tool enabling powerful, real-world prediction of mild distortion with modern simulators;
- a new and eminently practical approach to modelling complex nonlinear systems, applied successfully to loudspeakers;
- a distortion measurment test set with unequalled dynamic range and which is simple, fast and convenient to use;
- a device characterisation system that enables engineers to obtain data permitting GaAs device linearity study, and which is being adopted as a product by Hewlett-Packard;
- a new approach to visualisation of nonlinear behaviour, successfully used to identify properties of valve audio amplifiers and to facilitate linear MMIC design.

These achievements in the field of low-signal engineering have established several lines of promising research. The most appealing future avenues are already being pursued, and can be expected to bear fruit in the future. These are:

- developing software to extract time constants from pulse-profile (signal-time) data,
- verifying the precision of the PIV data by comparing derivative information obtained from APSPA PIV data with that from RF-TDFD measurements,
- identifying circuit topologies and device responses that are responsible for characteristic GDS attributes (such as valve amplifier GDS shapes),
- adding a frequency dimension to RF GDS obtained from RF-TDFD measurements and exploring amplifier characteristics with this,
- improving or adding simulator models based on linearity/GDS information (such as an improved magnetic core model and a thermionic valve model); and
- exploring improved ways to fit models to devices, assisted by linearity measures (especially for GaAs devices, loudspeakers, black-box amplifiers and multiterminal devices such as pentodes).

The author hopes this work or parts of it will answer people's questions as well as add to the sum of the world's knowledge.

Appendix A

Specifications of VXI-based APSPA

Specification	VXI	VXI & Pulser/PSU	Units
Quiescent Stimulus			
Voltage range	±20	0–50	Volts
Resolution	16	12	bits
Current range	> ±200	>1000	mA
Pulse amplitude specifications	(per E1446A)	HP85120-K43/HP6629A	
Voltage range	±20	0–50	Volts (O/C)
Output resistance	≈ 0.5	≈3	Ω
Pulse amplitude resolution	12	12	bits
Maximum peak current (<3ms)	>1	>10	Amps
Pulse timing specifications			
Minimum pulse width	100	400	ns
Maximum pulse width	25	0.4	s
Maximum pulse period	25	25	s
Event timing resolution	6.25	6.25	ns
Digitiser Specifications			
Voltage range	±0.1 to ±100	±0.1 to ±100	Volts (1–2–5 seq.)
Current range	±0.005 to ±25	±1 to ±25	Amps FSD
Resolution	12	12	bits
Accuracy	3	3	%
Measurement time	$S + D + M$	$S + D + M$	
S, software & operating system overhead	≈ 2	$\approx 2 + 2 \times n_d$	s
D, download & storage time	$\approx 0.0003 n_p$	$\approx 0.0003 n_p$	s
M, measurement time = $n_p \times n_a \times T$			
n_d = number of V_{drain} values			
n_p = number of points			
n_a = number of averages			
T = pulse period			
Time for 2000-point grid with 256 averaging and 0.1% duty cycle	6	25	s (typical)
Software Features			
Automatic voltage grids or user list of points	Yes	Yes	
Gridsize	8000	8000	points
Cable length compensation, auto calibration	Yes	Yes	
Background soak tests	Yes	Yes	
Pseudo-random ordering	Yes	gate only	
Script file command execution	Yes	Yes	
Command line parameter override	Yes	Yes	
Soft compliance limiting	Yes	Yes	
Linear pulse profiling	Yes	Yes	
Log-time pulse profiling	Yes	Yes	
S-parameters via HP85108/HP83650A	Yes	Yes	
			JBS&AEP 17-Jan-1996

Appendix B

APSPA MTT-S Presentation

A number of simple demonstrations of the APSPA system were presented at the exhibition of the 1996 IEEE Microwave Theory and Techniques Symposium (MTT-S) in the Hewlett-Packard booth. Figure B.1 shows the author and supervisor at the booth, and figure B.2 the equipment as demonstrated. The demonstrations were:

1. Measurement of a drain characteristic of a test device (a TC724 2-26.5GHz MMIC, effectively a set of parallel MESFETs). The characteristic consisted of approximately 1000 points, 64 times averaged, with a 1μs pulse period and a duty cycle of 0.05%. This took approximately 30 seconds, orders of magnitude less than the previously-available measurement system offered by HP-EEsof. The measurement time is almost entirely consumed by the waits between pulses. The result is shown in figure B.3.

2. Measurement of S-parameters at six points in the characteristic of the same device, for 51 frequencies, with 8 times averaging, in 1μs pulses, with 0.05% duty cycle. The measurement requires about 1 minute. The points are selected to give a spread of gain curves. The result is shown in figure B.4.

3. Three pairs of measurements, taken and presented sequentially, exemplifying the effect of pulse duration, and the use of shuffled-order measurement ("monte carlo"

or "monty" option) to identify pulse durations that eliminate history-dependent effects from characteristic measurement. The first graph (figure B.5) overlays two characteristics similar to that of DEMO 1, but having one use 20μs pulses: Clearly at least one will be misleading with respect to high-frequency operation. The second and third graphs (figure B.6 and figure B.7) showed the breakup using shuffled-ordering with a pulse width of 20μs, and the agreement in the case of 1μs pulses.

4. Measurement of the device characteristics with various duty cycles, showing the change in results that creep in when the duty cycle is too short (the pulse repetition frequency, PRF, is too high). The display is shown in figure B.8.

5. Measurement of a pulse profile with data gathered on a near-logarithmic scale. The display in figure B.9 has a scale covering five orders of magnitude from 0.1μs to 10ms, yet the data is gathered in only a few seconds with 32 times averaging.

6. Measurement of a characteristic on a fine grid, and subsequent calculation of transconductance, g_m, and output conductance, g_{ds}, along any given loadline. The result of the calculations are presented in figure B.10, along with the loadline superimposed over the drain characteristics.

7. Measurement of a limited drain characteristic to show that a representative measurement of a few hundred points with tolerable averaging and duty cycle can be carried out with one second of measurement and less than five seconds of overhead. Such timing shows that production-line use is feasible with the APSPA system. The result is shown in figure B.11.

8. Measurement of device data along an arbitrary V_g/V_{ds} trajectory, in this case the up-down sweeps typical of curve tracers, producing loops in the constant-V_g lines when timing is not such as to deliver isothermal/isodynamic results. The result is shown in figure B.12.

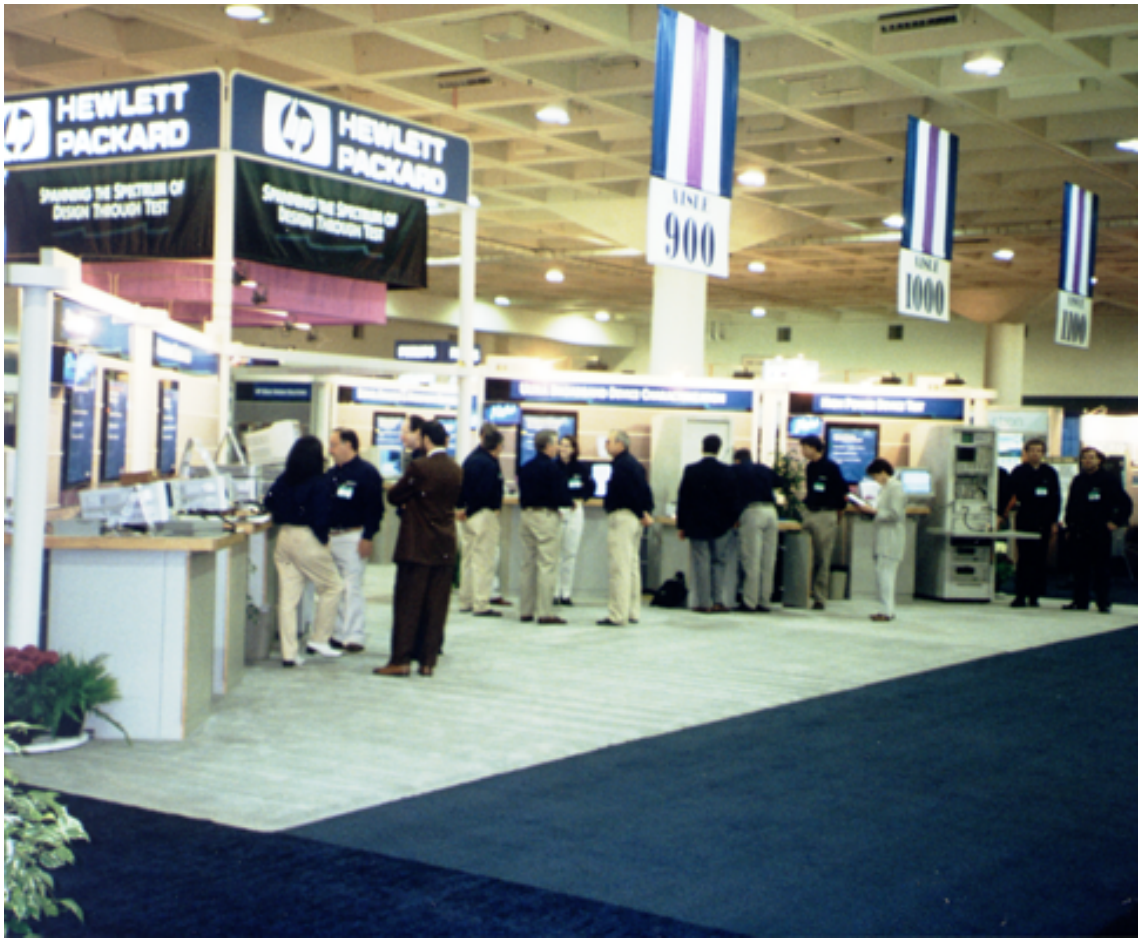

Figure B.1: Photograph of the author and his supervisor at the HP booth at the 1996 IEEE MTT-S exhibition, with the prototype APSPA system.

Figure B.2: Photograph of the prototype APSPA system as demonstrated at the 1996 IEEE MTT-S exhibition.

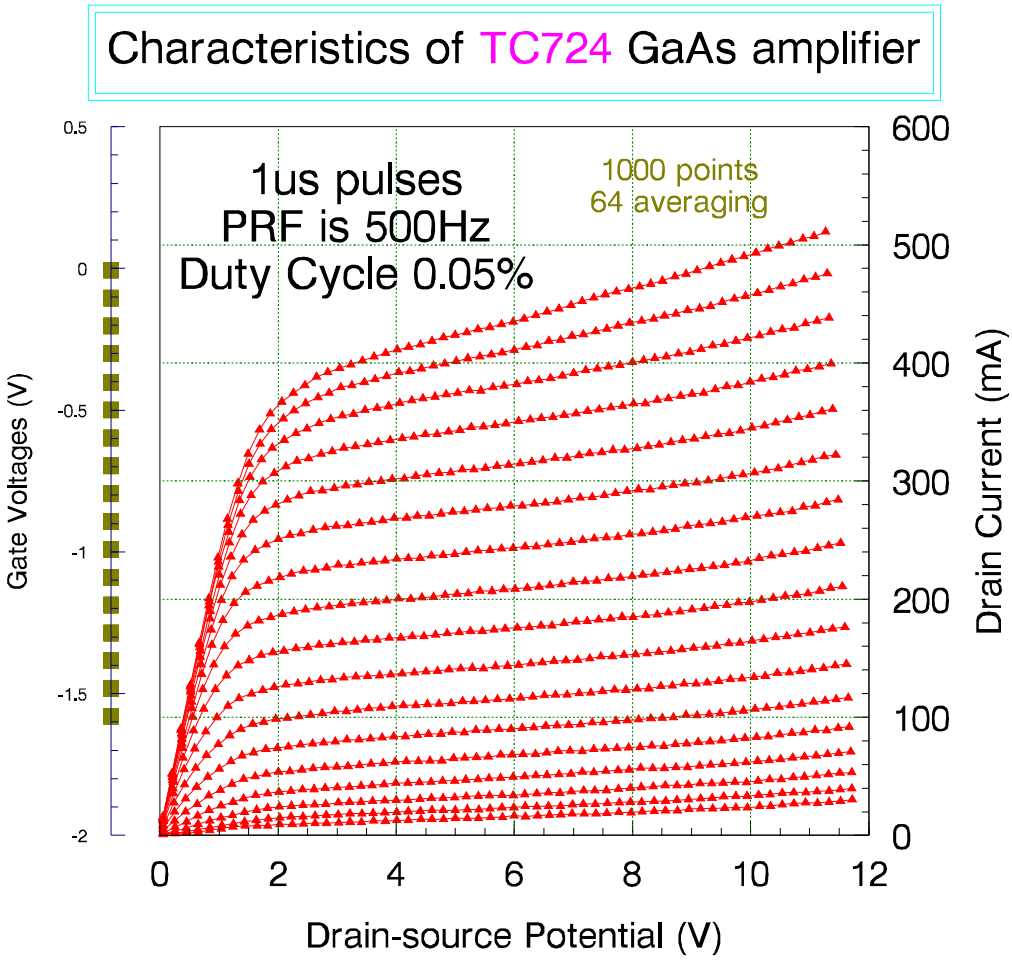

Figure B.3: Display of results from DEMO 1.

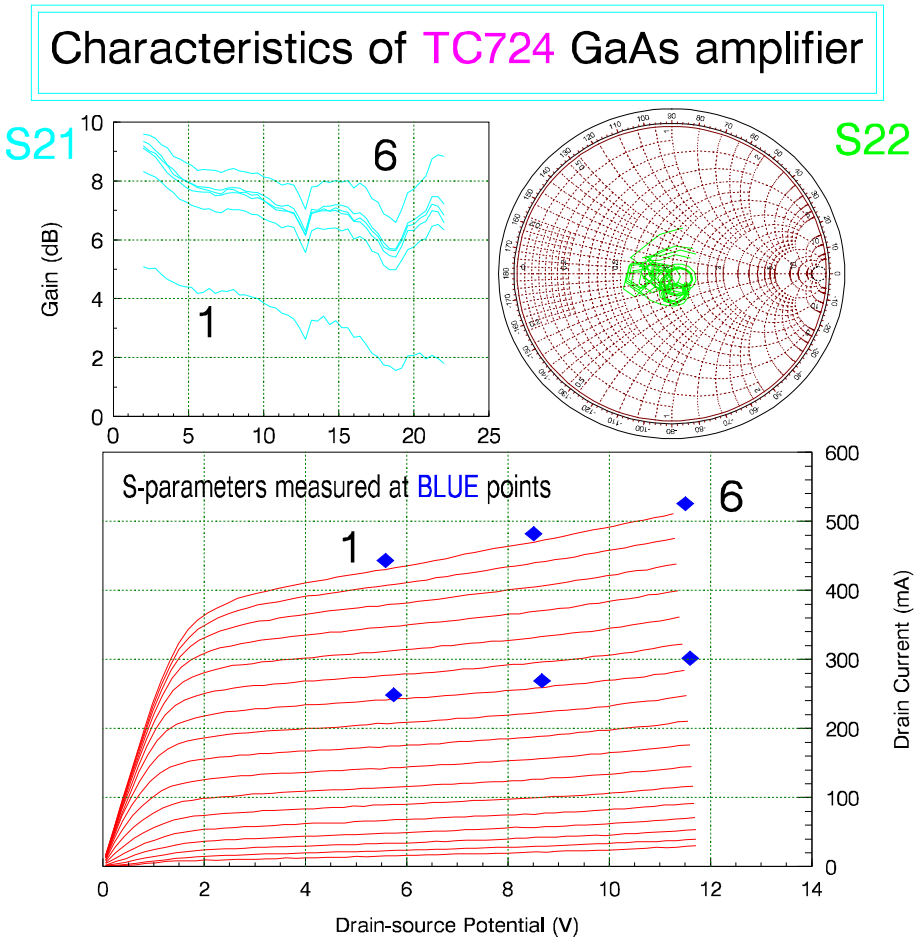

Figure B.4: Display of results from DEMO 2.

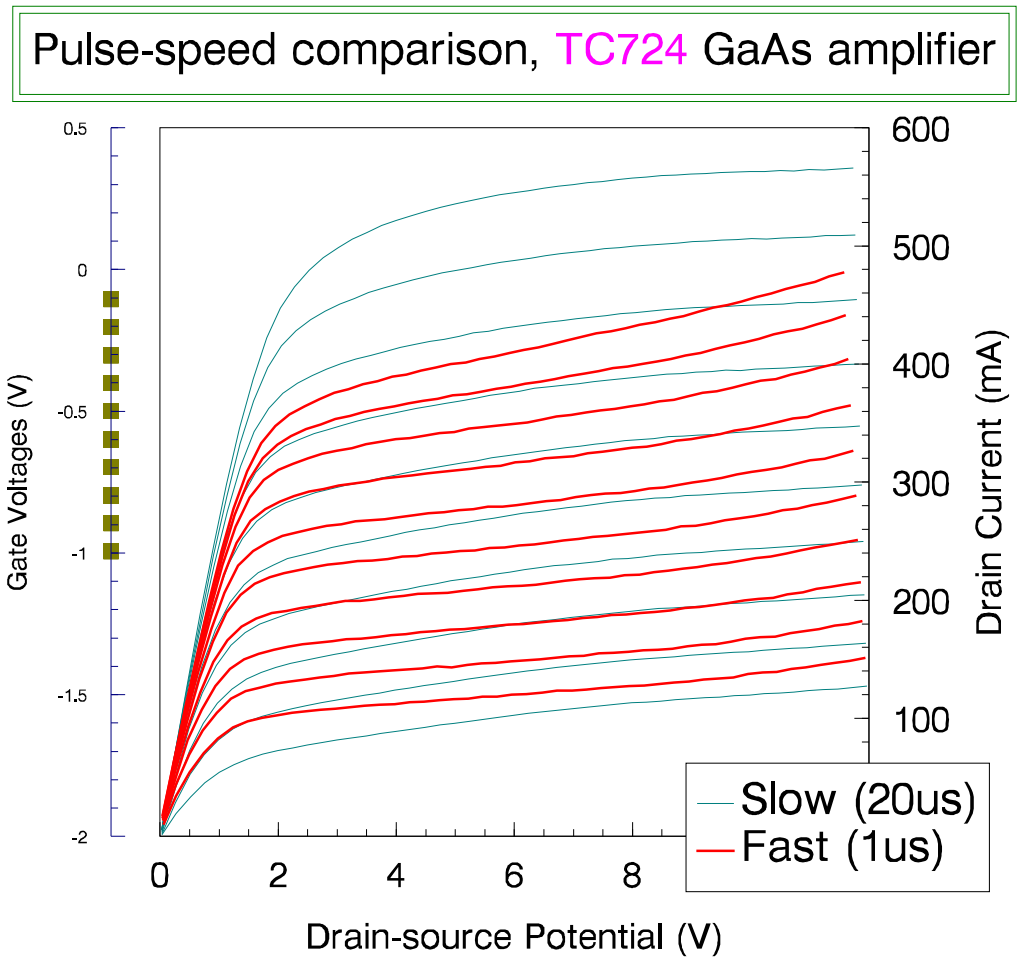

Figure B.5: Display of the first comparison result from DEMO 3, showing the dissimilar characteristics resulting from use of different pulse widths.

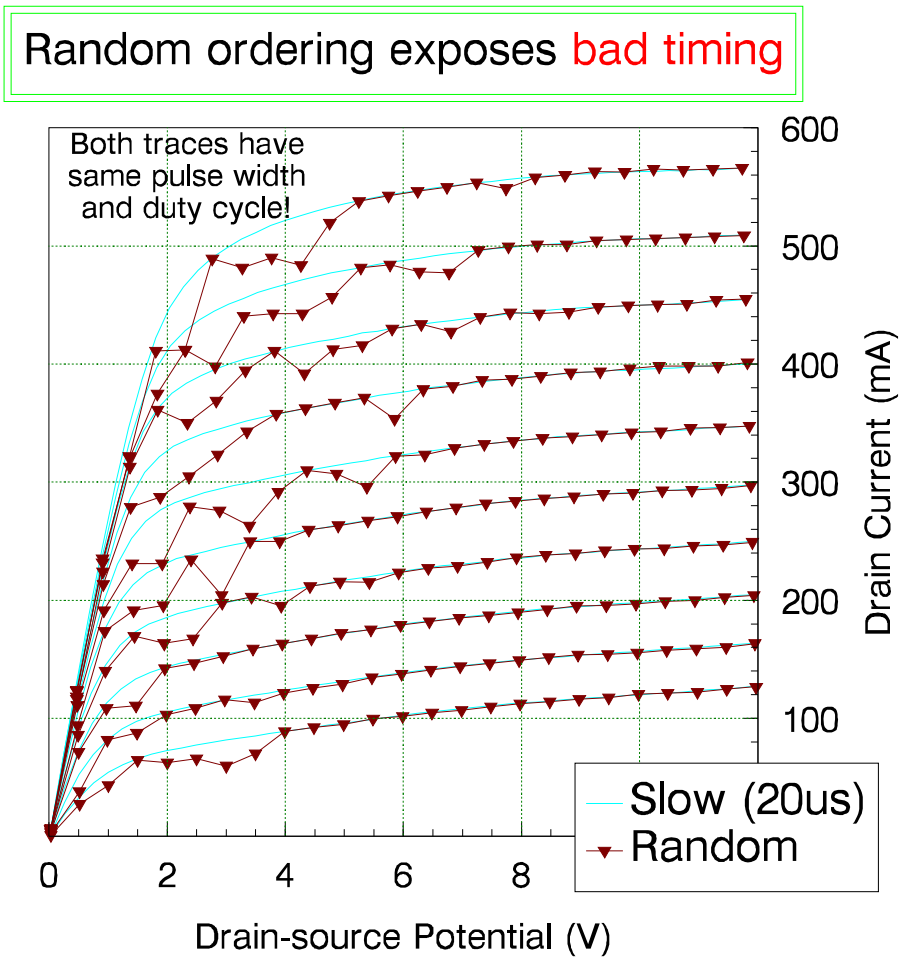

Figure B.6: Display of the second comparison result from DEMO 3, showing the breakup using shuffled-order measurements.

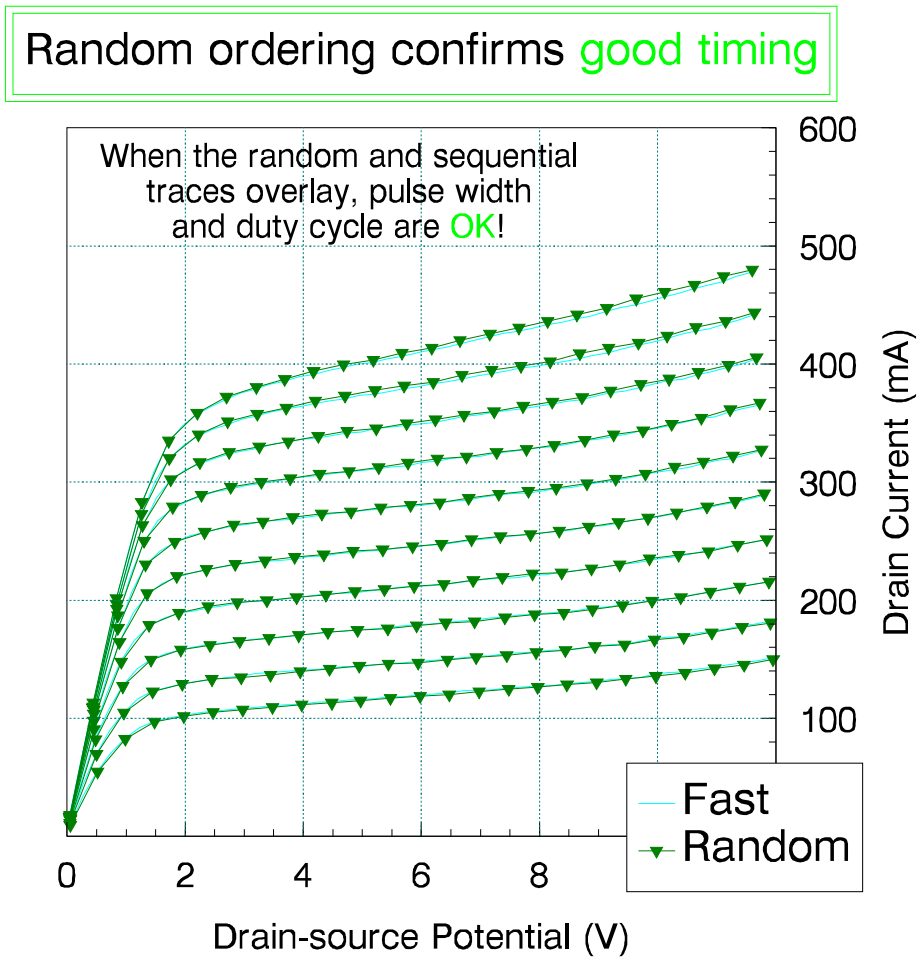

Figure B.7: Display of the third comparison result from DEMO 3, showing a "clean" agreement between sequential and shuffled-order measurements.

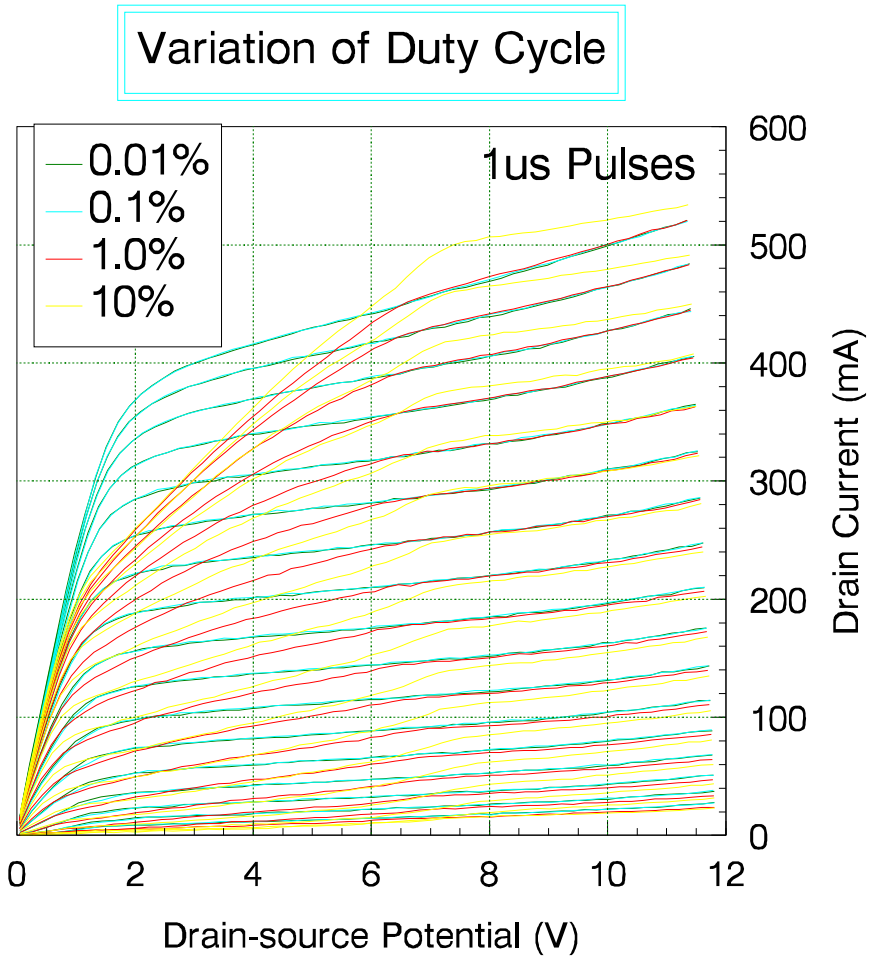

Figure B.8: Display of results from DEMO 4.

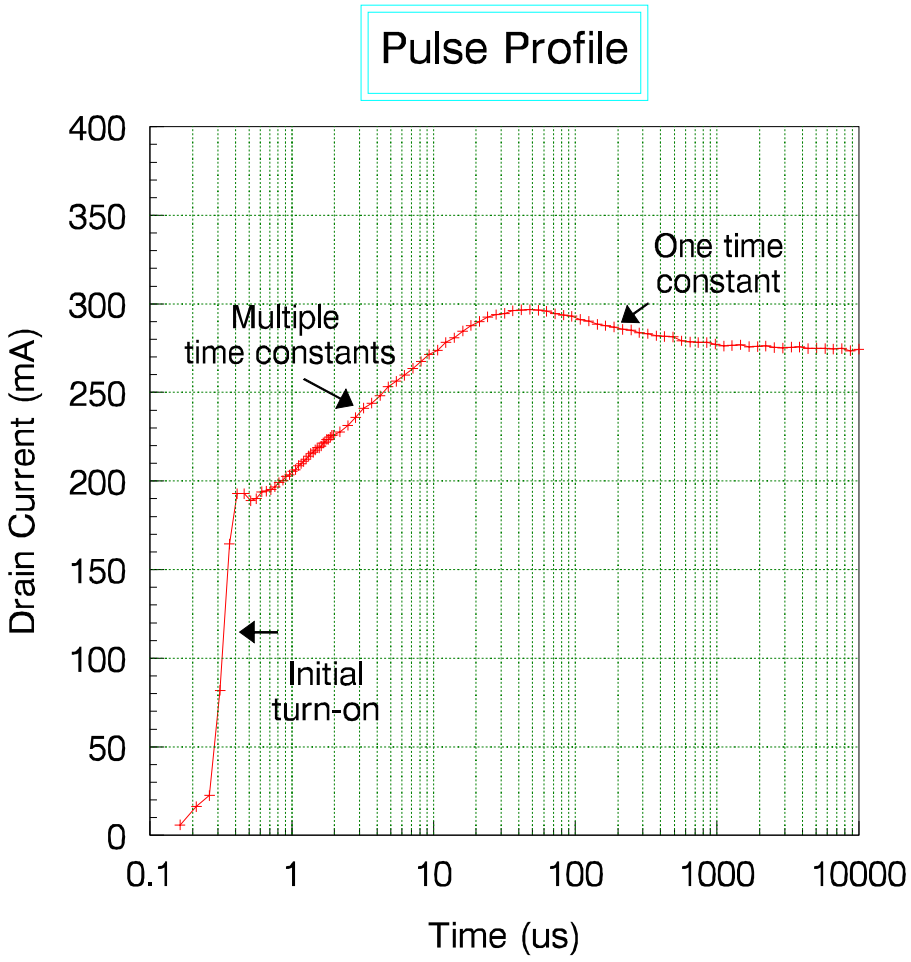

Figure B.9: Display of results from DEMO 5.

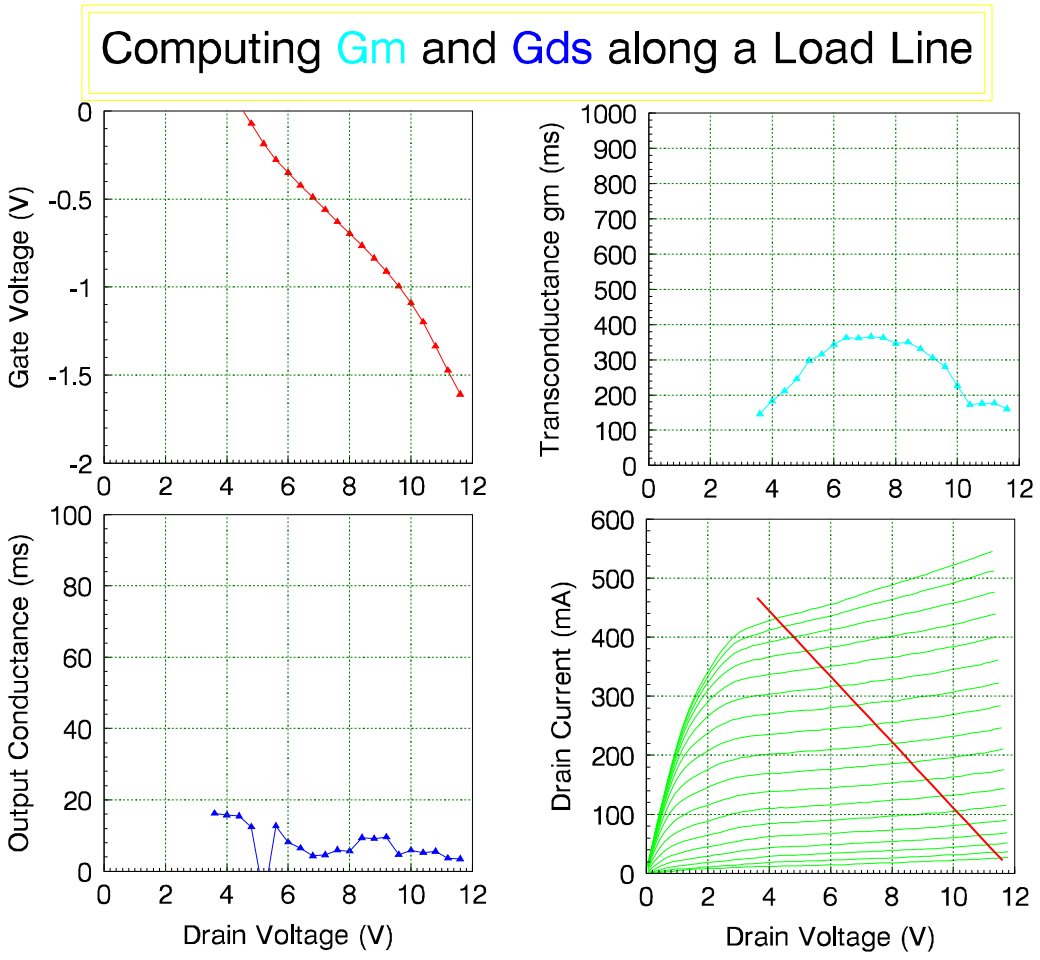

Figure B.10: Display of results from DEMO 6.

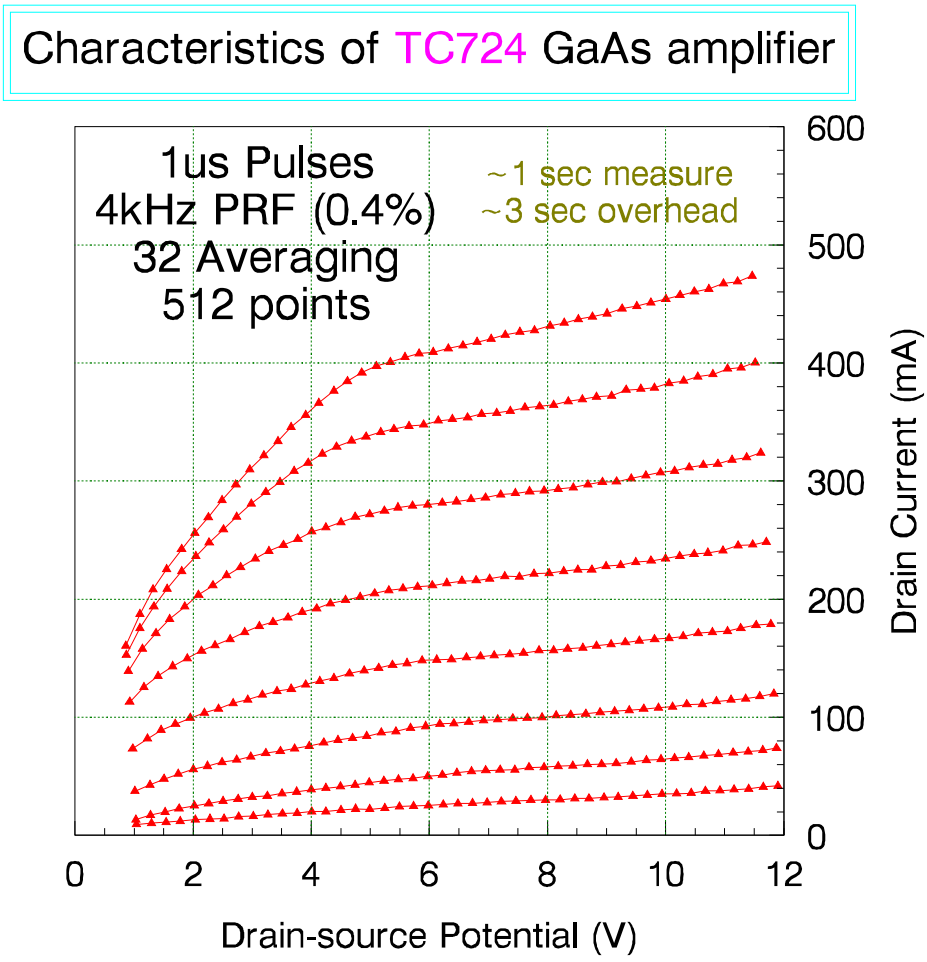

Figure B.11: Display of results from DEMO 7.

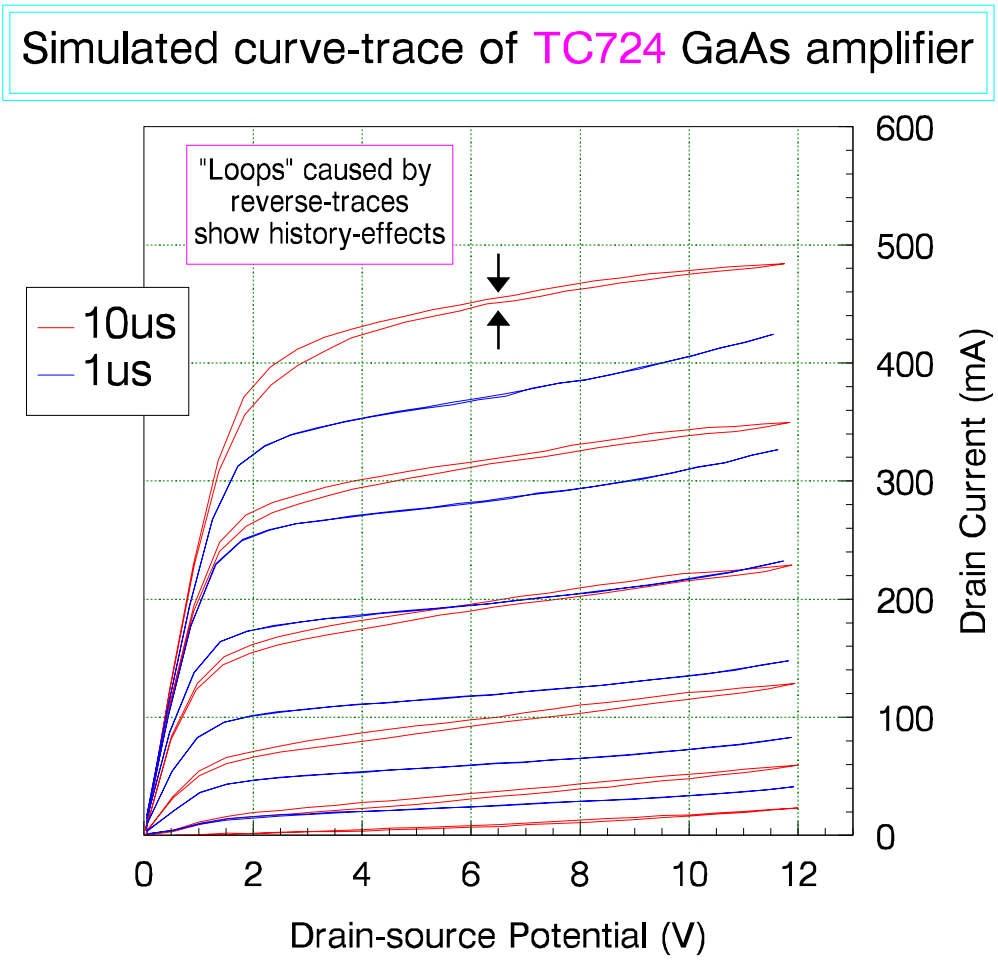

Figure B.12: Display of results from DEMO 8.

Appendix C

APSPA User's Manual

The APSPA system consists of the VXI instruments described in Chapter 6, installed as described in section C.1, with the software described in section C.2.

C.1 Hardware Installation

The Radi-Sys EPC7 embedded VXI 80486 PC must occupy the slot-0 position as controller of the VXI rack. The mechanical position of other modules is not critical, but the layout of figure 6.3 is recommended, and the device ULAs are fixed at compilation time. The patch panel is used purely to bring the timing module's 8 control lines out to appropriate connectors, and contains no active components.

The instrument VXI addresses and external HPIB instrument addresses set to the values set out in table C.1 operate with the default compilation values.

The interconnections between the modules are set out in table C.2.

Instrument	VXI ULA
E1450A Timing module	136
E1340A Arbitrary Function Generator 1	80
E1446A DAC/Amplifier 1	96
E1429B Digitiser 1	48
E1340A Arbitrary Function Generator 2	72
E1446A DAC/Amplifier 2	88
E1429B Digitiser 2	40
Instrument	**HPIB Address**
HP6629B Power Supply	5
HP85108C Pulsed Vector Network Analyser	16

Table C.1: Table of VXI module and HPIB instrument hardware addresses.

Connection from	To (Basic VXI configuration)	To (Using HP85120A/HP6629B)
E1450 control 7	E1450 Delayed Trigger Input	E1450 Delayed Trigger Input
E1450 control 6	NC	NC
E1450 control 5	RF Source Gate (pulse) Input	RF Source Gate (pulse) Input
E1450 control 4	HP85108C Trigger Input	HP85108C Trigger Input
E1450 control 3	NC	NC
E1450 control 2	E1429B Ext 1 inputs	E1429B Ext 1 inputs
E1450 control 1	E1340A AFG (2) Ext 1 input	E1340A AFG Ext 1 input
E1450 control 0	E1340A AFG (1) Ext 1 input	E1340A AFG Ext 1 input
E1446A (1) Main Output	Drain Drive	NC
E1446A (2) Main Output	Gate Drive	Gate Drive
E1340A (1) AFG Aux 1 Output	E1429B Ext 2 (Arm) Input	E1429B Ext 2 (Arm) Inputs
E1429B (1) Ch1 Lo-Z Input	AM503A (drain)	NC
E1429B (1) Ch1 Hi-Z + Input	NC	NC
E1429B (1) Ch1 Hi-Z - Input	NC	HP85120A I-sense
E1429B (1) Ch2 Lo-Z Input	NC	NC
E1429B (1) Ch2 Hi-Z + Input	Drain sense connection	HP85120A V-sense
E1429B (1) Ch2 Hi-Z - Input	NC	NC
E1429B (2) Ch1 Lo-Z Input	AM503A (gate)	AM503A (gate)
E1429B (2) Ch1 Hi-Z + Input	NC	NC
E1429B (2) Ch1 Hi-Z - Input	NC	NC
E1429B (2) Ch2 Lo-Z Input	NC	NC
E1429B (2) Ch2 Hi-Z + Input	Gate sense connection	Gate sense connection
E1429B (2) Ch2 Hi-Z - Input	NC	NC
EPC7 HPIB Interface	HP85108C HPIB Interface	HP85108C & HP6629B HPIB

Table C.2: Table of VXI module interconnections for basic and high-power options. NC means not connected.

C.2 Software Overview

The program "apspa" runs on a Radi-Sys EPC7 embedded VXI 80486 PC under the DOS operating system. It is written in C++ and compiled with The Borland Turbo C++ compiler. Its purpose is to carry out the pulsed measurements described in Chapter 6.

Apspa reads settings in from a file and writes data out to files. Progress messages are directed to stderr. All files in one run have the same name and different suffixes. Possible files are:

file.**SET** This file contains the setup instructions, see below for a full description.

file.**LOG** Interim progress and failure diagnostic reports are placed in this file.

file.**PIV** This file contains the response data, in ASCII form, in four columns giving gate and drain voltages and currents, one point per line. It contains a header, where each line is started with a backslash. The header records data such as the time of measurement, the voltage ranges, etc. When the data quadruplets result from stipulation in the SET file of starting, stopping and step gate and drain voltages, the output file is divided into blocks, corresponding to separate nominal gate voltage values, each block being started with a line containing two integer fields, the number of rows and columns (always 4) in the block. This permits plotting in the usual style, current on the y-axis, drain voltage on the x-axis, with gate voltage as a parameter.

file.**SPM** S-parameters are stored in this file. A header of similar format to the PIV file is followed by lines with nine fields, the first being frequency, and the next eight being the real and imaginary parts of the four S-parameters. The file is divided into blocks corresponding to each bias point. Note that a great deal of data, and a very long measurement sequence, can result if S-parameters are requested across a whole characetristic grid, and this option is typically called along with the list option.

file.**SCR** This is a script file, if the script function is used. Each line may contain any SET parameters that are to be changed during each rerun of apspa. For example, to perform multiple measurements with different pulse widths, each line of the script

file needs to contain only a "meastime=value" statement. The second line might typically contain a "nocal=1" command to cancel recalibration of offsets in the interest of speed. The original setting are not run until the first script line is processed, i.e., the first line takes effect on the first run.

file.**PUL** This file contains pulse-profile data, in time-voltage pairs, in ASCII, with the usual header, if a pulse profile is requested. Similar blocking occurs to divide the data from each pulse profiled. Note that requesting pulse profile on a normal characterisation grid can produce large amounts of data.

file.**LST** The LST file is read when it is desired to take data at user-specified points instead of fixed to a grid. The format is ASCII V_g/V_{ds} pairs, one per line, to a maximum of 2000 points.

C.3 Command-line Options

Running apspa with the -h option will direct the template setup file, similar to that shown in figure C.1, to stdout. Thus the command
apspa -h >fred.set
will produce a setup file for the "fred" set of measurements, which can subsequently be edited with any ASCII text editor, such as vi, EDIT, EDT, etc. The setup variables will be described in groups in the next section.

When wired to use an HP85120A pulser, apspa should be run with the -c option. This enables the power supply programming via HPIB, and switches the digitiser to read drain current from a high impedance input instead of the 50Ω input, and rescales the drain voltage reading appropriately. It also makes the drain AFG output boolean so that can be used as the pulser logic input, and leaves the drain amplifier output disabled. (In coding terms a different routine is called to handle drain configuration, hence -c.)

The software will also accept the -s option, which is equivalent to forcing scriptfile operation with stdin being read for the script commands. This is useful for testing small changes without modifying files.

C.4 The SET File

The # symbol is the comment symbol in set files.

The quiescent operating point is set with the variables **vdsq**, the drain-source quiescent potential, and **vgsq** the gate-source quiescent potential. The default grid of points to which pulsed measurements are to be made are set by **vdstart**, **vdstop**, and **vdstep** for the drain, and **vgstart**, **vgstop**, and **vgstep** for the gate. Irrespective of the sign of the step parameters, both variables will be stepped from the more negative to the more positive limit in steps of magnitude equal to the magnitude of the step parameter. Where there is an exact multiple of the step magnitude between the limits, both limits are included. Thus stepping from 0 to 1 by 0.1 gives eleven values.

The parameter **meastime** sets the pulse width, and **quiestime** sets the interval between pulses. The boolean **swtq** enables software extension of the quiescent time. This needs to be used only if quiescent intervals in excess of about 10 seconds are required, but may compromise the timing precision.

The parameters **ppoints**, setting the largest number of samples that may be taken in each pulse, **pulsetail**, setting the interval at the end of each pulse during which data may be collected, and **tsi**, setting the interval before the end of the gate pulse before which all sampling must be completed, control digitisation within the measurement pulse. The **average** parameter sets the minimum amount of averaging to be carried out on each gridpoint. The timing diagram shown in figure C.2 helps to explain the interrelationships of the timing parameters.

Two parameters control digitisation of the quiescent stimulus and response. The value of

```
#--- A P S P A: Arbitrary Pulsed Semiconductor Parameter Analyser ----
vdsq = 0.0 # Drain source quiescent potential
vgsq = 0.0 # Gate-source quiescent potential
vdstart = 0 # Drain start potential
vdstop = 1 # Drain stop potential
vdstep = 0.1 # Drain step potential
vgstart = -1 # Gate start potential
vgstop = 0.0 # Gate stop potential
vgstep = 0.1 # Gate step potential
measTime = 1.0e-6 # Measurement time
ppoints = 20 # Samples per measurement time
pulsetail = 0.2e-6
# Fraction of pulse tail from which data may be collected
tsi = 0.15e-6 # Sample clock inset from end of t_p
quiesTime = 2000e-6 # Quiescent time
swtq = 0 # Enable use of software tq, if possible
qpoints = 1 # Samples per q time
tdrel = -0.1e-6 # Drain pulse delay relative to gate pulse
tpre = 0.20e-6 # Quiescent sample time before t_p
tri = 0.10e-06 # RF pulse inset from end of t_p
trm = 1e-06 # Sample (NWA) inset from end of RF pulse
leaveon = 1 # Leave outputs running on exit
loop = 1 # Timer continuously loops
spms = 0 # Whether S-parameters are to be gathered
average = 64 # Data bursts to be averaged
vdmax = 20 # Absolute maximum permitted drain voltage
vgmax = 0 # Absolute maximum permitted gate voltage
vdmin = 0.0 # Absolute minimum permitted drain voltage
vgmin = -20 # Absolute minimum permitted gate voltage
idmax = 0.1 # Maximum drain current
igmax = 0.1 # Maximum gate current
shuffle = 0
# Shuffle measurement order? (if not zero this is seed to random)
listfile = 0 # Read file.LST for datapoints?
script = 0 # Use a script file? (0=N, 1=stdin, 2=ask)
digits = 5 # Significant figures for output
wait = 1 # Seconds idling with signal before measurement
nocal = 0 # skip calibration of digitisers
zoutg = 0 # Output impedance of GATE source
zoutd = 0 # Output impedance of DRAIN source
profile = 0 # Perform a linear pulse profile measurement
logprof = 0 # Perform a log pulse profile measurement
calibrate = 1 # Apply calibration to timer outputs
cal0 = -150e-9 # Calibration time displacement factor for output CGT0
cal1 = -150e-9 # Calibration time displacement factor for output CGT1
cal2 = 0 # Calibration time displacement factor for output CGT2
cal3 = 0 # Calibration time displacement factor for output CGT3
cal4 = 0 # Calibration time displacement factor for output CGT4
cal5 = 0 # Calibration time displacement factor for output CGT5
cal6 = 0 # Calibration time displacement factor for output CGT6
cal7 = 0 # Calibration time displacement factor for output CGT7
#--- ------------------------------------------------------------ ----
```

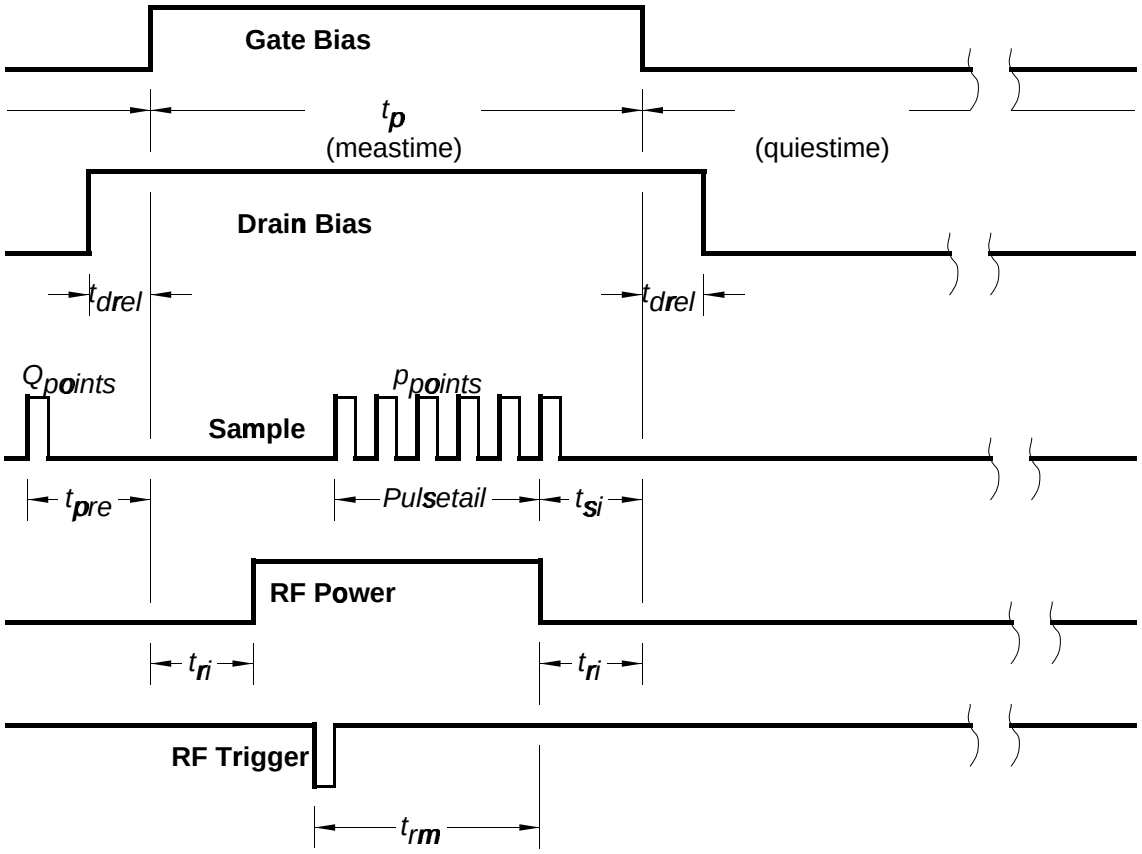

Figure C.2: APSPA timing diagram showing the relationship between timing parameters.

qpoints sets the number of samples of the quiescent levels to be taken before the onset of the gate pulse. (The gate pulse is taken as the master interval to which all others relate.) The value of **tpre** sets the interval before the gate pulse starts, by whih the quiescent sampling must be completed.

Importantly, **tdrel** sets the duration between the onset and termination of the drain pulse relative to gate pulse. When positive, the drain onset is delayed and the termination advanced.

Putting **spms** non-zero instructs S-parameters to be measured at all gridpoints, using an HP85108C NWA. Two parameters control the synchronisation of a network analyser. The relative timing of the RF gate pulse to the RF source is set by **tri** in the same manner as **tdrel**, while **trm** specifies the advance of the sample pulse sent to trigger the NWA with respect to the RF gate pulse. This must be set in conjunction with the HP85108C trigger delay parameter. The RF gate is positive-going as are most of the control logic signals. The NWA trigger pulse, however, is negative-going, as the HP85108 triggers on the negative-going edge of the trigger input.

The boolean parameters **leaveon** and **loop** dictate whether the analog pulse outputs are left active, and whether the timing module is left running. These are chiefly useful in diagnostic and checking roles, but may be used to leave a device "soaking", as the pulse sequence will continue even after apspa has finished running. Note that **leaveon** without **loop** could potentially harm a device, although this is unlikely as the timer is halted at a quiescent point.

A software double-check is included: the parameters **vdmax**, setting the absolute maximum permitted drain voltage, **vgmax**, setting the absolute maximum permitted gate voltage, **vdmin**, setting the absolute minimum permitted drain voltage, and **vgmin**, setting the absolute minimum permitted gate voltage, will abort a run if any gridpoint exceeds any of these values. They are typically set to a device's "absolute maximum ratings".

The current ranges for drain and gate are set by **idmax** and **igmax**. Note that these

may be set high if full 12-bit resolution is not required, or low, if it is desired to expand a low-current part of the characteristic. Autoranging is not supported, as it would require multiple passes.

The integer **shuffle** parameter enables scrambling of the data points when it is not zero. Its value is used to seed the random generator, so any given randomsequence may be reproduced.

When **listfile** is not zero, the contents of the .LST file override the grid as the source of the data point values. This is used to specify arbitrary trajectories.

When **script** is set equal to 1, stdin is used as a source of setup parameters that may override the contents of the mandatory setup file. This allows apspa to be used in a pipe, or in live keyboard mode. When set equal to 2, **script** specifies the use of the .SCR file as the script source.

The parameter **digits** dictates the number of significant figures used in the output. ASCII numbers are generally written in engineering notation (exponents set to multiples of three).

The integer **wait** sets the number of seconds spent idling with signal applied before digitisation is commenced. This is useful for letting a device settle to the average of a measurement sequence when measuring below known time constants in the device response.

Putting **nocal** nonzero skips calibration of the dc offsets in the digitisers (and amplifiers).

The parameters **zoutg** and **zoutd** fix the output impedance of gate and drain amplifiers. Legal values are 0, 50 and 75. Zero actually produces an effective source resistance of about 0.5Ω.

The parameter **profile** will replace measuring pulse value with a linear pulse profile measurement. In this case, the parameters **ppoints**, **pulsetail**, and **tsi** fix the interval over which the pulse is profiled. **Pulsetail** is usually chosen larger than **meastime**. If **logprof** is nonzero, an attempt is made to space the measurements logarithmically across the

interval for the pulse profile measurement. This may not be possible, especially for large numbers of points, because of limitations in the timing module hardware, but it can reduce data file sizes drastically, especially if software methods for time-constant extraction that require transition to a logarithmic domain are to be used.

Finally the boolean **calibrate** causes the algorithm to apply the calibration constants to the timer outputs. These constants, **cal0** through **cal7**, are chiefly used to allow for cable lengths in the instrument, but also allow for characteristic delays in instruments such as the E1446A amplifiers, external pulsers, etc. The values are typically set for a particular set of instruments when first assembled, using a fast oscilloscope at the fixture.

Appendix D

Nonlinear Loudspeaker-like Load

D.1 Introduction

Tests conducted with a plain, nominal-value resistor are convenient, but not characteristic of real operating conditions. Recent work unambiguously shows the need to take into account the reactive nature of real loudspeaker loads when designing and testing amplifiers.[108]

There are three main motives for constructing a load that electrically resembles a loudspeaker, but that does not contain an actual transducer (in preference to simply using a loudspeaker):

- Even brief laboratory tests of an amplifier, using real loudspeakers, creates annoying, loud sounds, be they tones, music or noise;
- Drivers are wont to sustain damage driven continuously by single-frequency or several-tone signals characteristic of standard tests such as THD and TDFD;
- Loudspeakers are bulky and expensive as test equipment goes, especially for powers above a couple of hundred watts.

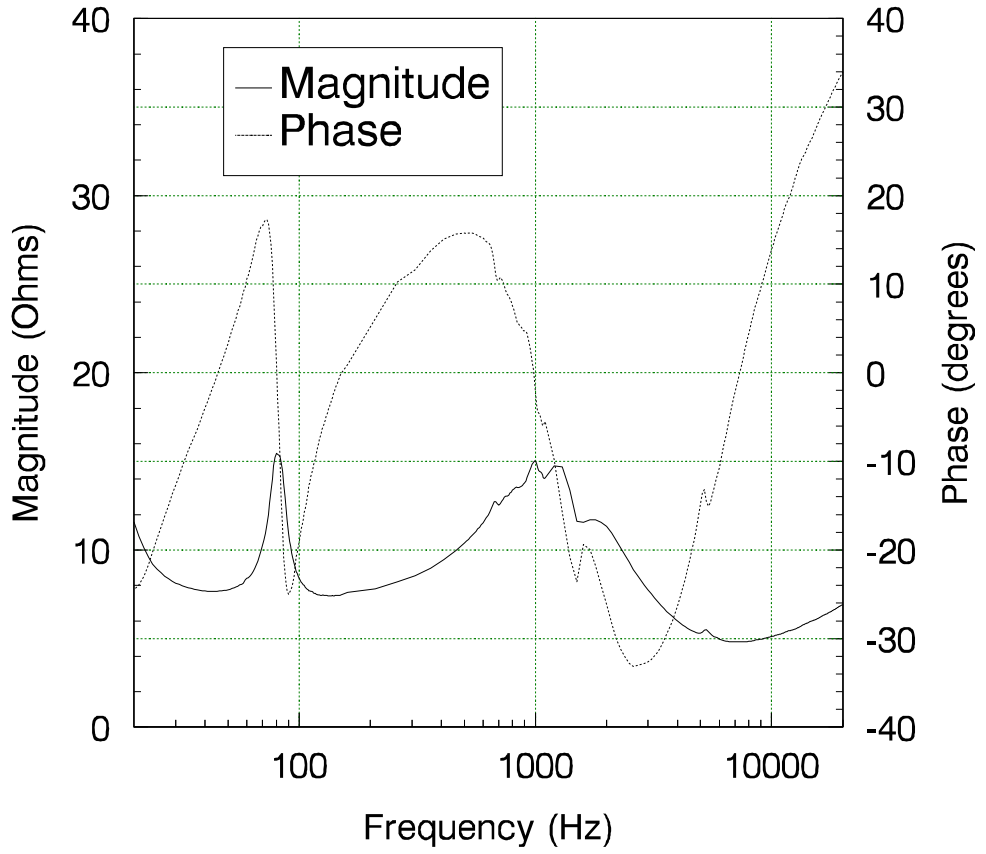

Figure D.1: Impedance of a typical, 3-way, commercial loudspeaker.

This section describes a passive electronic circuit that presents a port impedance representative of a (moving-coil type) loudspeaker. The impedance resembles a real loudspeaker not only in its linear, reactive nature, but also in its non-linear nature. Results of measurements of a number of loudspeakers have been used to identify common characteristics. A particular set of component values is given, but an algorithm by which values may be found is also presented. The equivalent circuit load is silent, portable, relatively inexpensive and robust. As such it represents the best possible laboratory test load for an audio power amplifier yet reported.

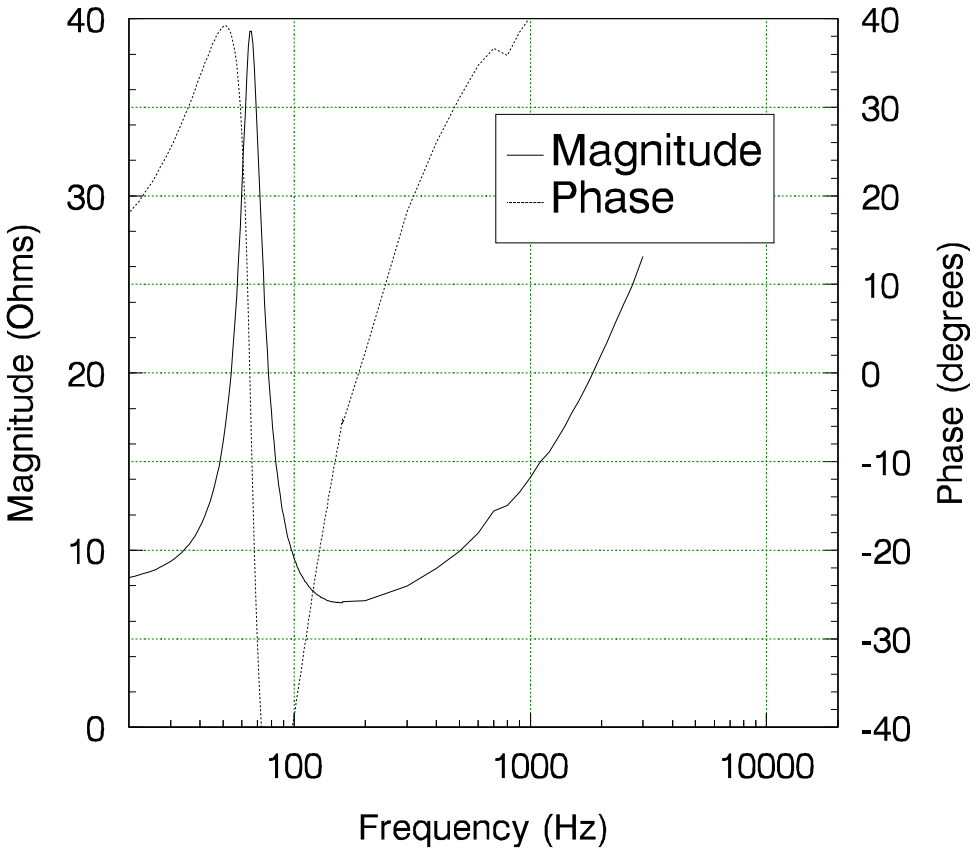

Figure D.2: Impedance of a single-driver system.

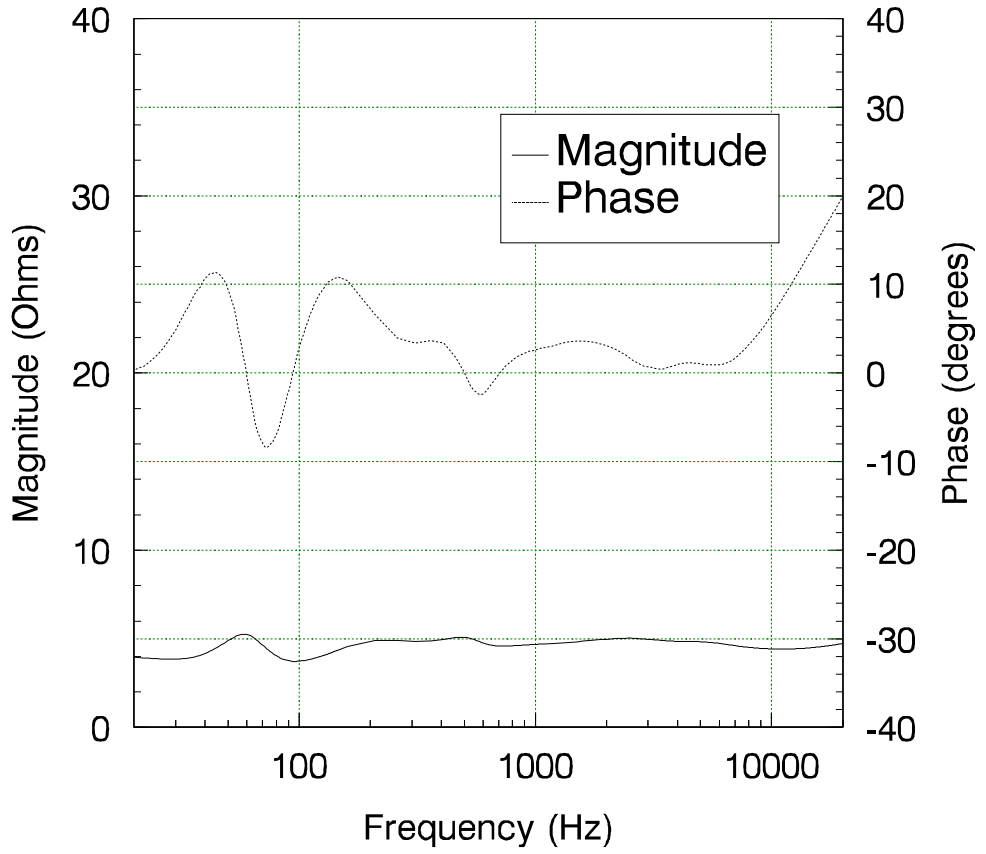

Figure D.3: Impedance of a 2-way loudspeaker fitted with impedance-compensating networks.

D.2 Typical Impedance Characteristics

The (complex) impedance of loudspeakers has been well-reported in the literature over recent years.[34, 108] With the exception of a small number of systems containing compensation networks, such as those manufactured by KEF electronics, the terminal impedance of a loudspeaker box is dominated, especially at low frequencies, by the impedances of the drivers and mechanical properties of the enclosure. The impedance of a nominal 8Ω system will typically start as a resistance of about 6Ω at dc, move inductive and then rise to a resistive resonance of a few tens of Ωs, become capacitive, and fall to a minimum close to the nominal resistance value at a few hundred Hertz. Behaviour above this frequency will depend upon whether the box is a single, two-way or three-way system, how complicated is the crossover network, and whether any pronounced resonances from other drivers (or the enclosure) appear at the terminals. A single driver will typically have overall impedance behaviour consisting of a small gradual rise in magnitude with frequency. Figures D.1, D.2 and D.3 give examples of impedance curves measured on sample systems. (Note that the graphs have been identically scaled to facilitate comparison.)

D.3 Typical Non-linear Characteristics

Any reports in the literature of the linearity of loudspeaker current demands are uncommon. Consequently linearity measurements have been made with numerous systems, including one with impedance-compensation. Figure D.4 shows the block diagram of the measurement setup used.

The linearity of the current drawn by a driver varies with frequency as well as drive level. Also, it may be characterised in various ways, such as the total harmonic content or the content at various individual harmonic frequencies. A reasonable impression of the linearity of a driver is gained from a plot of the fundamental current, the second and third-order distortion products, and the THD calculated from all harmonics below

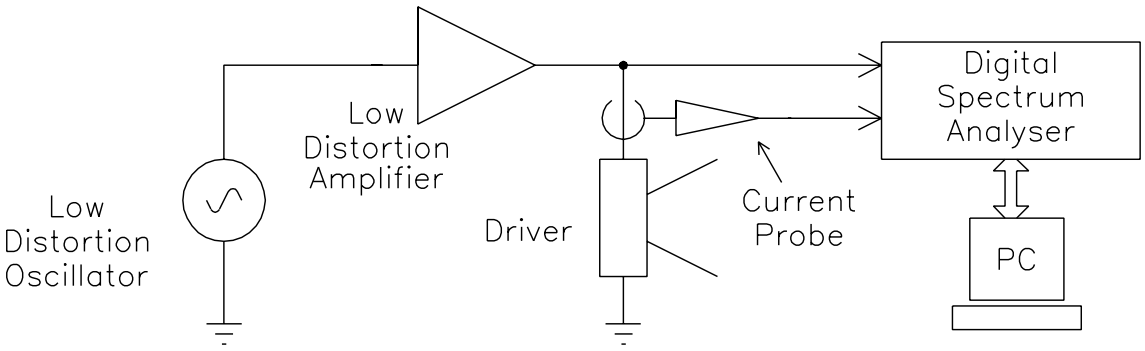

Figure D.4: Measurement setup used to determine current linearity of loudspeakers.

20kHz, against frequency. This produces a crowded graph, but one that provides much information in return for brief study.

Some current distortion measurements are shown in figures D.5–D.7. Each figure shows fundamental, second-, third-order and THD current against frequency, for two drive levels on one example loudspeaker enclosure. (The levels are chosen to suit the particular system.) Although the *fundamental* current drawn by the compensated system is flat in comparison to uncompensated systems, note from figure D.7 that the *harmonic* currents behave similarly to those of the uncompensated systems, as expected.

Many variable factors affect the current distortion characteristics—the number of drivers, various ratings of the drivers, complexity of crossover network, type of enclosure—but several common attributes appear:

1. As frequency rises, second- (even-) order non-linearity falls, and third- (odd-) order non-linearity remains more level, thus eventually becoming the dominant effect.

2. Odd order effects are not as dramatically level-dependent as basic theory might suggest, especially at higher frequencies.[1]

3. In the region of driver resonance, current magnitudes fall, but harmonics tend to rise relative to fundamental, so total distortion rises. Mechanical design of the driver

[1]This seems contrary to expectation, especially as well-known measures such as third-order intercept are based on the expectation from theory that third-order components should rise as the *cube* of fundamental component magnitude. However, the assumptions of time-invariance and dominance of the lower-order terms in the Taylor's series expansion of transfer functions are questionable in this, as other, common instances. Refer to the discussion in section 2.3.3.

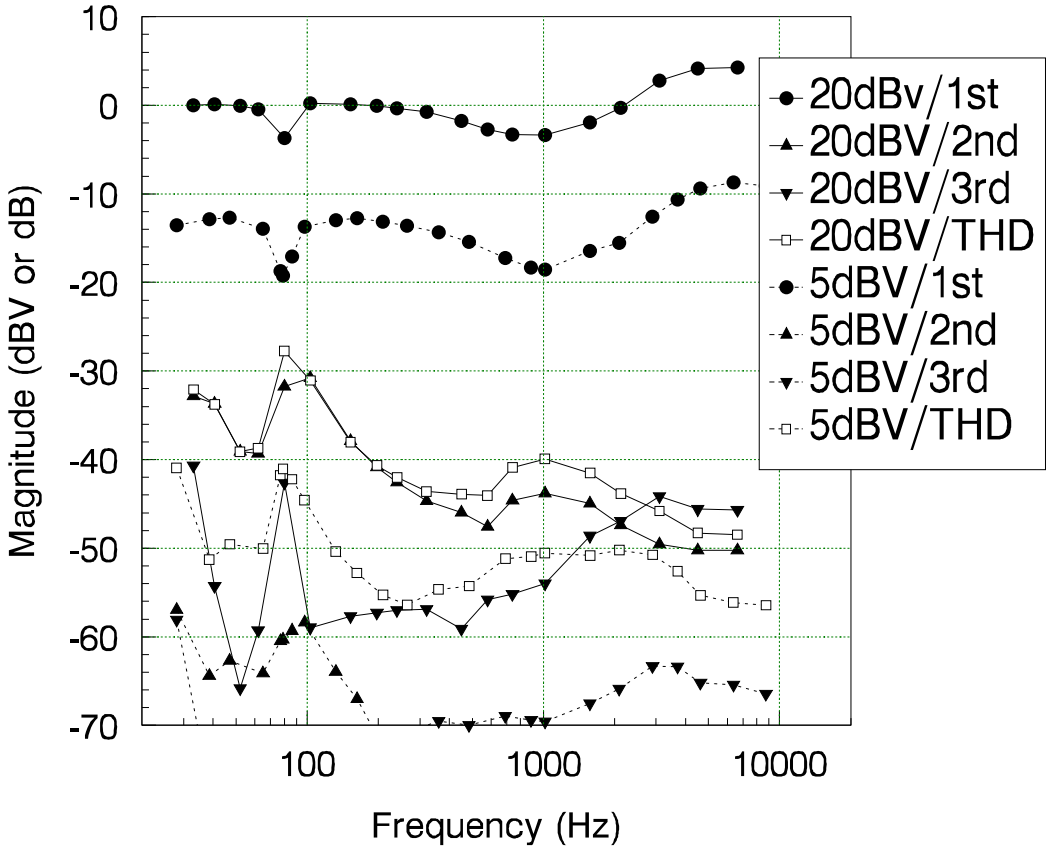

Figure D.5: Harmonic currents and THD in the 3-way loudspeaker.

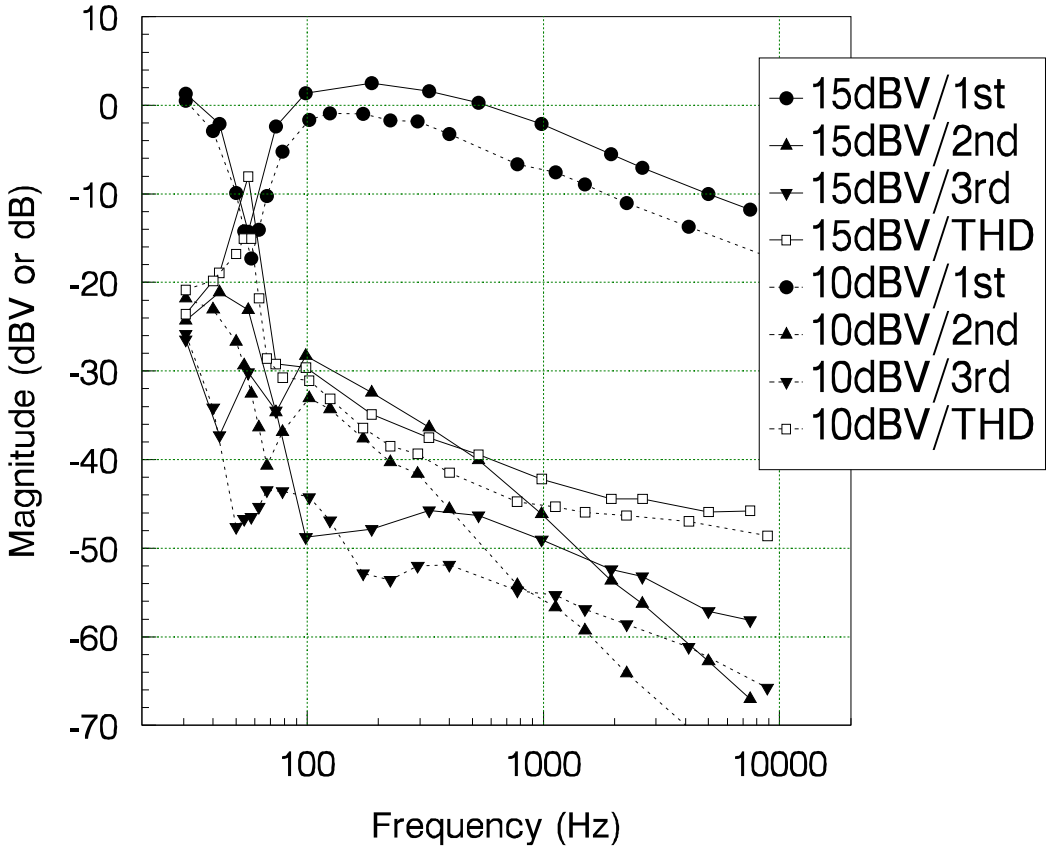

Figure D.6: Harmonic current distortion in the single-driver enclosure.

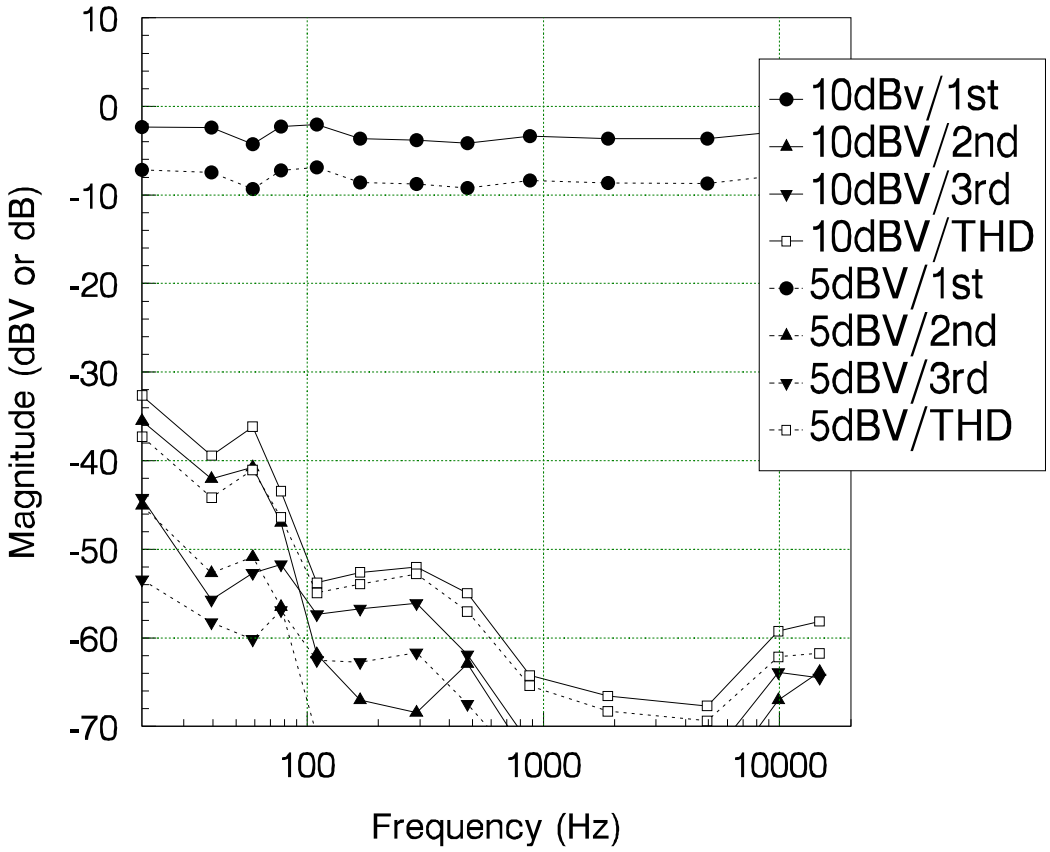

Figure D.7: Harmonic current distortion in the loudspeaker with impedance compensation. Note the relatively flat *fundamental* current.

probably sets whether the second- (even-) or third- (odd-) order effect dominates, but second is commonly dominant.

D.4 Aim of the New Load Circuit

The electronic circuit delivers an impedance that can be set to model a simple driver in both linear and non-linear characteristics. An enclosure may be modelled either by using several driver-model subcircuits and the same crossover network as is used in the enclosure being modelled, or more simply with a single circuit. The latter will necessarily be relatively crude, but has been observed to be quite adequate as most difficulties occur in the frequency range covered by the low-frequency driver.

D.5 Equivalent Circuit

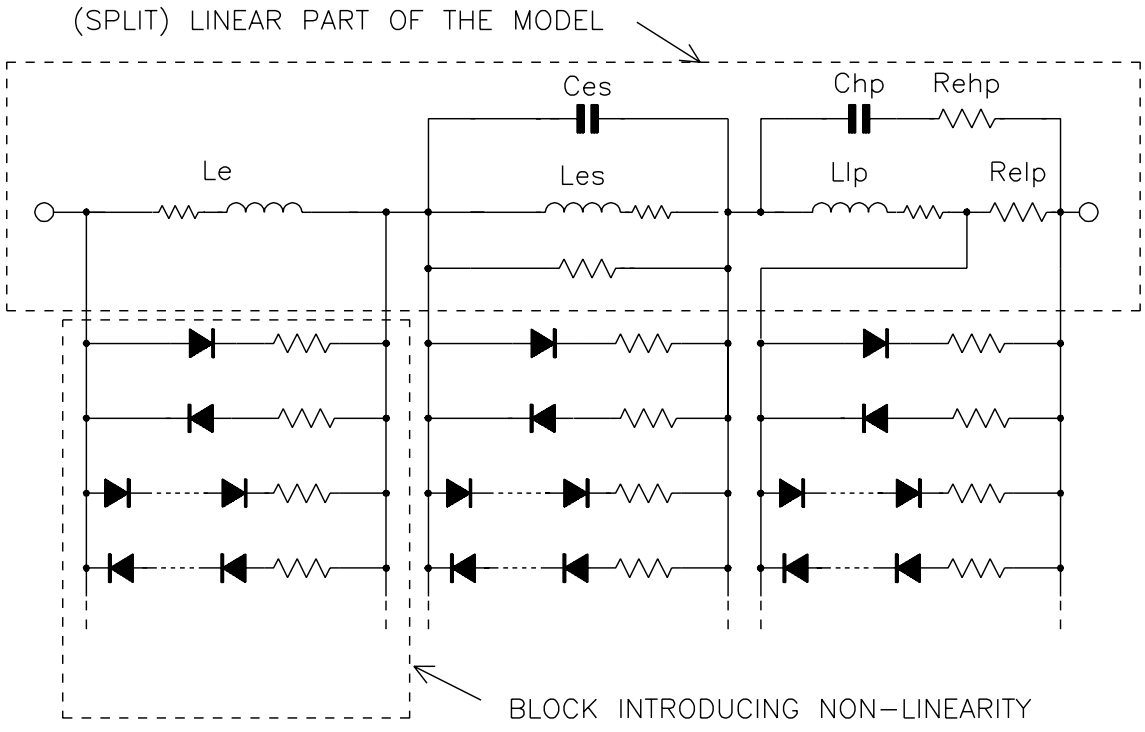

Figure D.8: The circuit of the non-linear loudspeaker-like load.

The general form of the equivalent circuit of the electronic load is shown in figure D.8.

One dotted line in the figure identifies the basic, reactive, small-signal circuit. Another dotted line identifies one of several, potentially identical, non-linearising networks.

D.5.1 Component Selection in the Linear Subcircuit

The reactive circuit roughly resembles the Thiele-Small equivalent circuit of a single driver. The main exception is that R_e, the ohmic coil resistance, has been replaced by a separate network. This network consists of a simple L-C crossover comprised of C_{hp} and L_{lp}, with their associated resistors, $R_{\text{e hp}}$, and $R_{\text{e lp}}$. It can be shown that putting

$$R_{\text{e hp}} = R_{\text{e lp}} = R_e \tag{D.1}$$

and

$$\frac{L_{lp}}{C_{hp}R_e^2} = 1 \tag{D.2}$$

leads to a smooth transition from the current flowing in the low-pass arm to the high-pass arm as frequency passes

$$\omega = \frac{1}{\sqrt{L_{lp}C_{hp}}} \tag{D.3}$$

while the total network impedance remains equal to R_e. This is useful when allowing for the varying non-linear current contribution with frequency.

The reactive elements of the equivalent circuit may be chosen quite straightforwardly to suit any enclosed driver. The parallel tuned circuit is given the same centre frequency and Q as that of the resonance observed at the loudspeaker terminals, with due allowance for the series resistance of the inductor. The value of L_e is selected to approximate the magnitude of impedance at higher frequencies, as in a Thiele-Small fit to a single driver.[2] The crossover frequency of the tail circuit is chosen to be that where the magnitude of the target system's second order distortion falls below that of its third order component. The series resistances of the tails are chosen to provide the correct value of total ohmic impedance, after allowance for losses in the inductors.

[2]The approach could of course be expanded to a more complex driver impedance model such as that in [32], if desired. This degree of refinement is deemed unnecessary for the application here.

A constructional trade-off is likely to arise in practice. This involves the size and cost of the inductor in the main tuned circuit. The coils must of course be *air cored*, so that they do not introduce non-linear effects because of saturation and hysteresis in core material. Achieving the same Q as a driver can demand a huge inductor.

D.5.2 Component Selection in the Non-linear Subcircuits

The reactive linear circuit will draw a good approximation of the required fundamental current component. The task then is to select components in additional circuits which will draw the required currents at higher harmonics. This has been achieved by a two-stage process:

1. To a numerical model of the linear reactive circuit, a current source of even and/or odd conductance is added in parallel with each of the four subsections of the reactive circuit. The coefficients are selected by iterative numerical means to obtain the desired harmonic characteristic.

2. Each ideal source is then approximated by a resistor-diode network.

The current sources are constrained to have zero conductance for zero applied voltage, and to be either perfectly anti-symmetrical ($I_s|_{V<0} = -I_s|_{V>0}$) or to have zero conductance for $V < 0$, making the source one-sided. These constraints make the sources easy to approximate by means of straightforward resistor-diode networks. The equivalent circuit used in the numerical process is given in figure D.9.

About resonance, virtually all of the applied voltage appears across the main parallel tuned circuit of the equivalent reactive model circuit. The non-linear current at this point must be produced by the non-linear elements in parallel with the parallel resonant circuit. The current source coefficient(s) are selected to produce behaviour matching the system to be modelled at the resonant frequency. This can be carried out substantially independently of performance at other frequencies. (This makes the iterative procedure orthogonal enough to be done manually in many cases.)

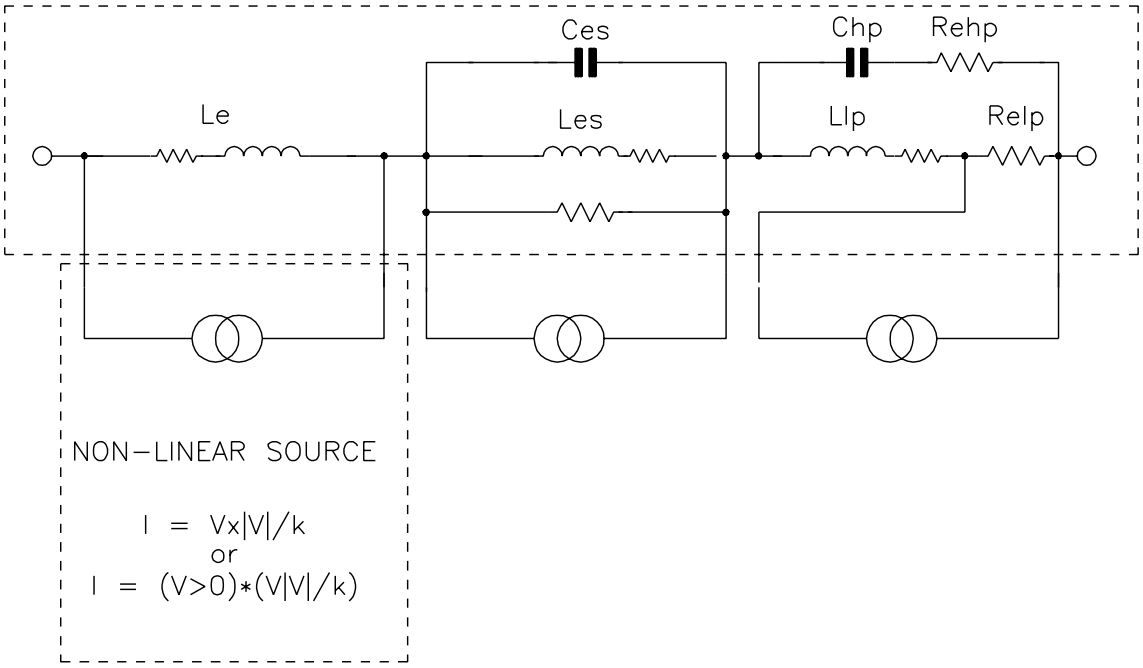

Figure D.9: Circuit used for numerical determination of suitable nonlinear elements.

Above the main driver resonance, but below the point where odd order effects overtake even, there is a peak in even order components; this is generated by the low-pass tail circuit. The peak arises because even order components are rising in comparison to the fundamental components with falling frequency, but the overall magnitudes fall with rising impedance. This is a suitable point from which to use data to fix the extent of the even-dominant distortion generated by the low-pass tail network.

Finally, the high-frequency non-linear behaviour is produced first in the high-pass tail circuit, but then chiefly by the voltage dropped across L_e. It has been found that a non-linearising network in parallel with L_e is generally sufficient. Thus no non-linearising block appears in parallel with $R_{\mathrm{e\ hp}}$ in figure D.8.

R-D network approximation

Each non-linear approximation network consists of parallel arms, each arm consisting of a resistor in series with a string of diodes, representing an ideal diode in series with a voltage source. If the arms come in symmetrical pairs (which may of course be achieved

with a single resistor, power dissipation permitting), the effect will be symmetrical and hence will favour odd harmonic generation. If all the diodes in a network face one way (the other direction arms in figure D.8 being deleted), the distortion will be even order.

The non-linear current sources have the form

$$I = \frac{V \times |V|}{k} \tag{D.4}$$

in the symmetric case. In the asymmetric case the equation is the same for $V > 0$ and 0 otherwise.

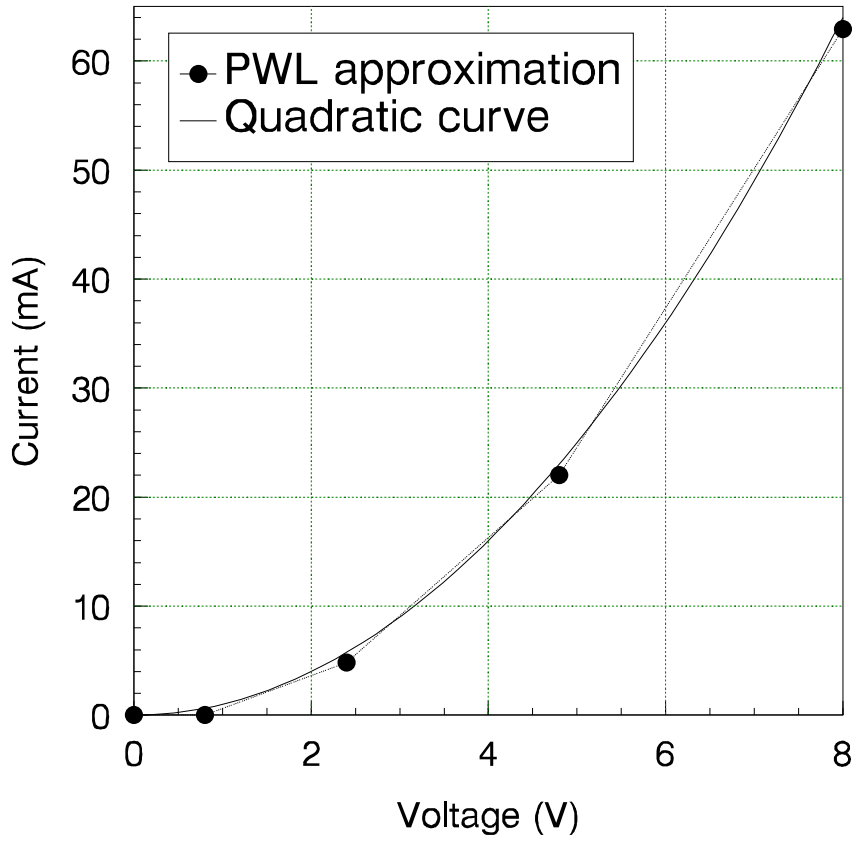

Figure D.10: A piecewise-linear approximation to the quadratic V-I characteristic.

Suppose we wish to approximate this in a piecewise-linear fashion using the circuit described. An example is depicted in figure D.10, where the curve is $\frac{V^2}{1000}$. The first step is to estimate the voltage level applied across the source when the load is being subjected to its nominal maximum working condition. In this example, let us assume that a peak

voltage of about 8 volts is to be expected. Next, suitable breakpoints are selected for the piecewise-linear approximating function. These are set by the available voltage drops. Since we use silicon diodes, these voltages are multiples of 0.8V, and the first is generally at 0.8V, the lowest possible value. Two or three are generally sufficient. Knowing that

$$I = kV^2 \tag{D.5}$$

the slope midway between breakpoints i and j is found by simple calculus to be

$$m_{\mathrm{i,j}} = \left.\frac{\delta I}{\delta V}\right|_{V=\frac{V_i+V_j}{2}} \tag{D.6}$$

The total resistance is then the reciprocal

$$\frac{1}{m_{\mathrm{i,j}}} = R_{\mathrm{total}} = \frac{k}{V_i + V_j} \tag{D.7}$$

and

$$R_{\mathrm{total}} = R_1 \parallel R_2 \parallel \cdots \parallel R_{i-1} \tag{D.8}$$

allowing the resistance of each arm in the nonlinear subcircuit to be found.

Between the first and second breakpoints this is simply the resistor in series with the first diode; subsequent resistors are found by taking the (ever decreasing) total resistance, and calculating what resistor must be paralleled with the previous ones to arrive at the total. This in the example above, between 0.8 and 2.4 volts the slope may be taken as the tangent at 1.6V, giving $R_{\mathrm{total}} \approx 310$, say 330Ω. If the next breakpoint is at 4.8 volts (6 diodes), R_{total} becomes ≈ 140 giving the next resistor a value of about 240Ω. Finally, the slope between 4.8 and 8 volts will demand a total resistance of about 78Ω, giving a final resistor of 180Ω.

D.6 Results

Figure D.12 shows the simulated and measured impedance, while figure D.13 and figure D.14 the simulated and measured current distortion, of the load with the values shown in figure D.11. There is good comparison between expected and measured values. The

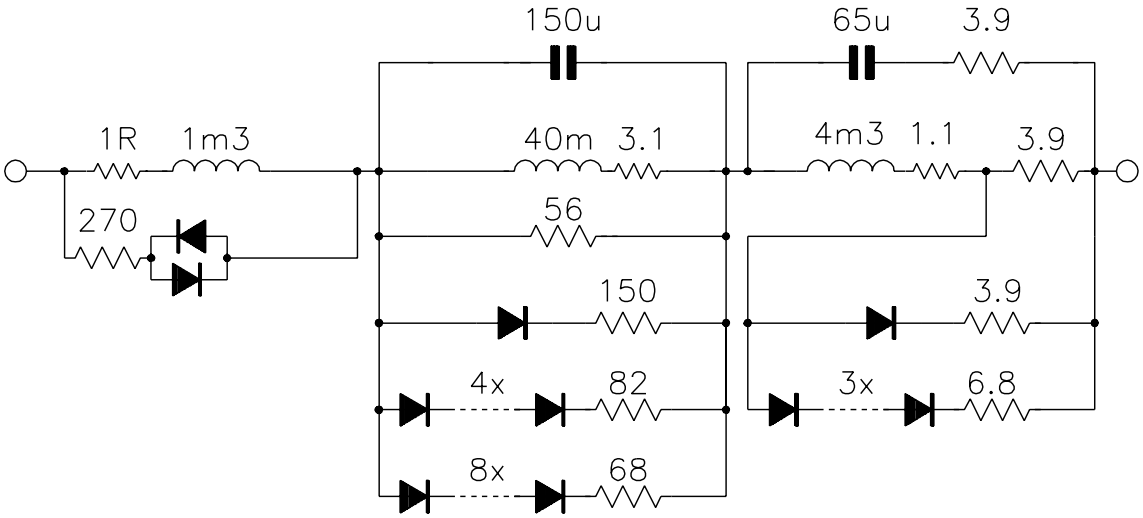

Figure D.11: Non-linear load circuit modelling a single-driver enclosure.

component values have been chosen to model a typical, modern 200mm driver in a simple enclosure; overall characteristics resemble those of the single-driver, laboratory system referred to earlier. Repeated measurements suggest that variation in impedance and linearity varies from driver to driver by the same degree as the differences in measured and emulated distortion currents. Variation from system to system, of course, occurs to much greater degrees.

D.6.1 Power Dissipation

The question arises as to what power level is possible with the equivalent circuit constructed with components of given ratings, or conversely what components to use to guarantee a given power handling capability. It is not easy to specify power dissipation of the whole circuit, since the power is dissipated in different components at different frequencies, and often within the inductors to a significant degree. It is possible to make an estimate of the power capability for single-frequency drive signals numerically. Figure D.15 plots power dissipation against frequency in the three main resistors and the 40mH inductor. Clearly, the high-pass-tail resistor needs to be capable of dissipating about 60 Watts, the low-pass-tail resistor about 15 Watts,[3] while using 10 Watt resistors elsewhere will be

[3]The graph suggests this figure to be 30-odd Watts. This is a pessimistic estimate since it does not account for the assistance provided by the loading of the resistors in the non-linearising network, as the calculation has been carried out with SPICE's .AC command and is therefore a purely linear estimate. A

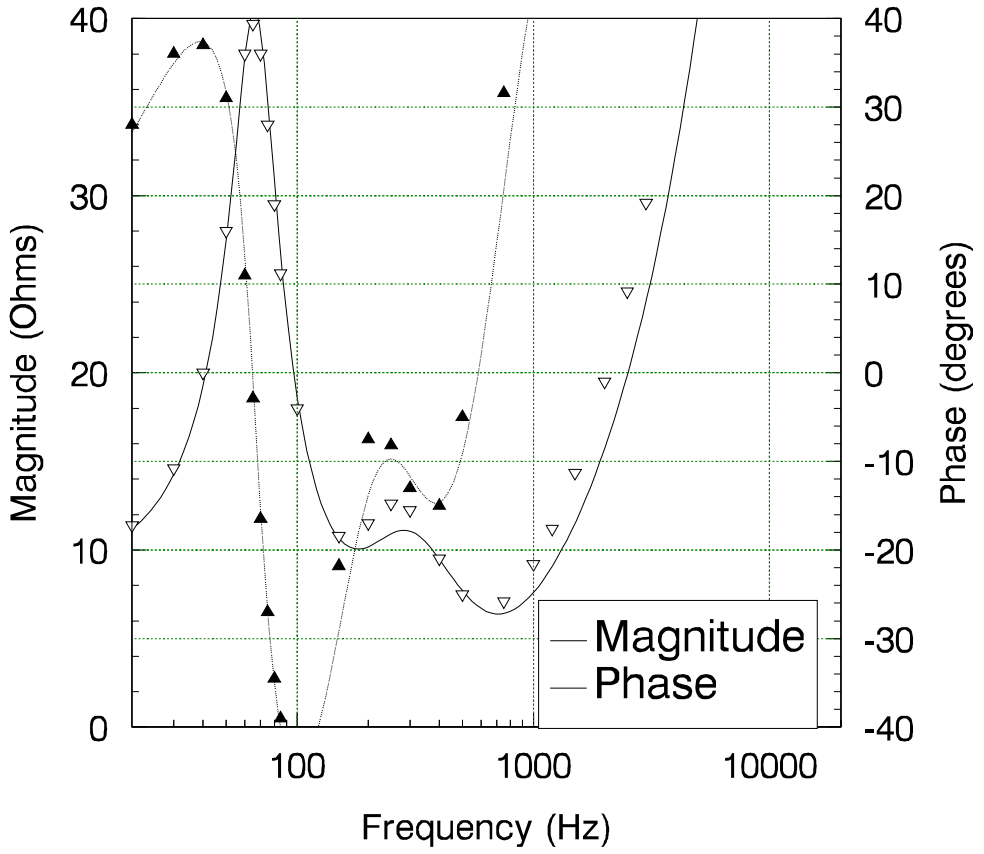

Figure D.12: Simulated (lines) and measured (symbols) impedance of the non-linear load.

quite satisfactory if the model is required to handle up to 25Vrms of drive. The 40mH coil can be expected to dissipate at most about 18 Watts at 20Hz. Recall that the coils must be *air cored.*

The prototype was constructed with resistors continuously rated at 10 Watts and with diodes rated for three Amperes average. The tail resistors are formed by using simple networks of several 10 Watt resistors. Being air-cored, the coils are sufficiently heavy that they do not enter into the power dissipation problem. The capacitors are formed using 10μF or smaller polyester devices paralleled on PC boards. The whole is mechanically supported by a wooden structure that offers no permeability or magnetic linearity influences. All components and materials are thus ones readily available. Although the prototype has

more thorough power dissipation analysis can be carried out at individual frequency points using a .TRAN analysis.

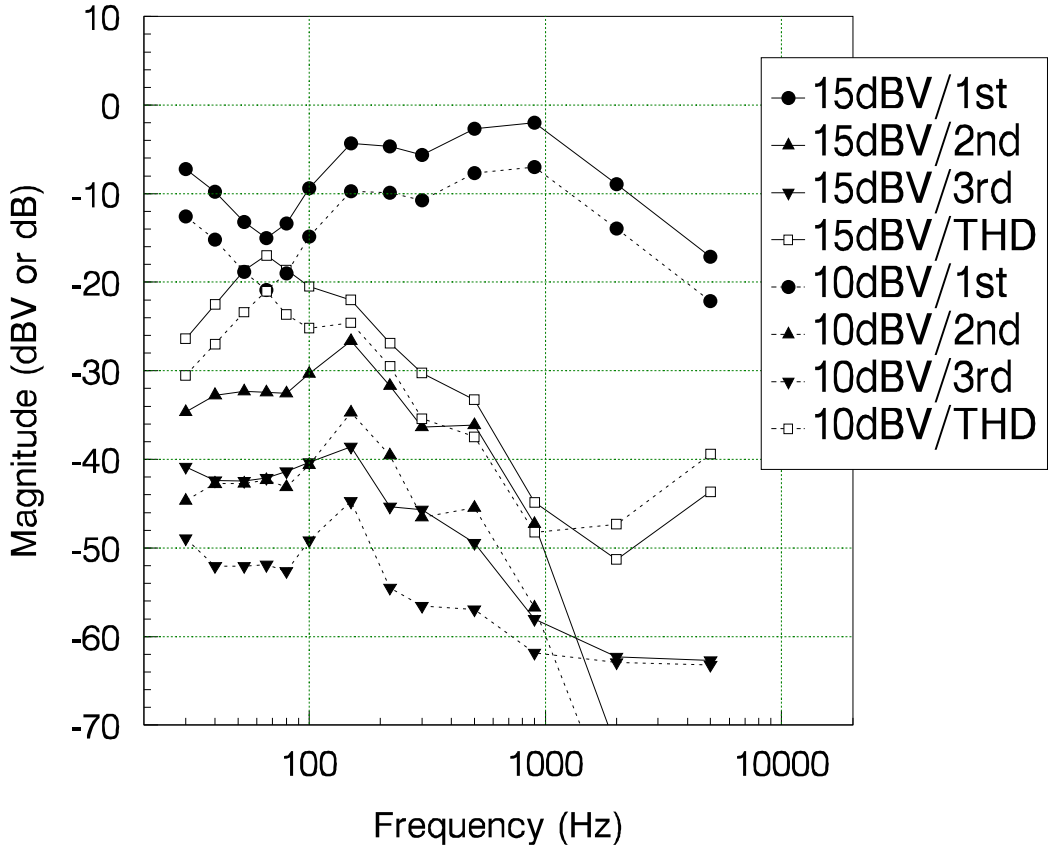

Figure D.13: Simulated fundamental and harmonic currents, and THD, of the non-linear load.

been given particular characteristics, component values required to adjust the model to mimic any specific loudspeaker load may be determined.

D.7 PSPICE Listing

The example listing below shows an implementation of the non-linear sources required for optimisation.

```
Example equivalent ckt      JBS Jan-1995
.tran 1m 20m 0m 0.01m  ;command
.four 200 I(Vsig)         ;command
```

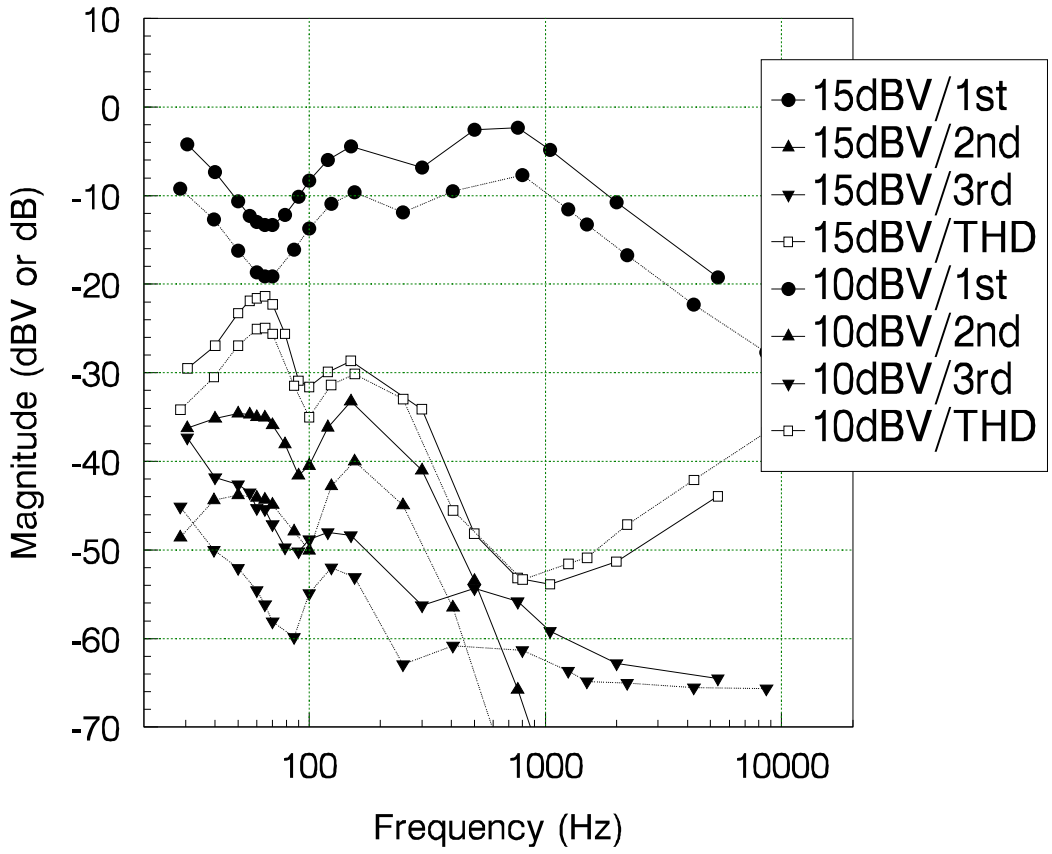

Figure D.14: Measured fundamental and harmonic currents, and THD, of the non-linear load.

```
Vsig in 0 AC 1 SIN(0 14 200Hz)

.FUNC MIN(A,B) (A+B-ABS(A-B))/2

.FUNC MAX(A,B) (A+B+ABS(A-B))/2

*

Le intnl1 vvel 1.3m

RLe in intnl1 1.0

* A non-linearity with Le

GLenlo   in vvel  ;odd

+ VALUE={V(in,vvel)*abs(V(in,vvel))/5500}

GLenle   in vvel  ;even

+ VALUE=

+{MAX(0,V(in,vvel)*abs(V(in,vvel))/4000)}
```

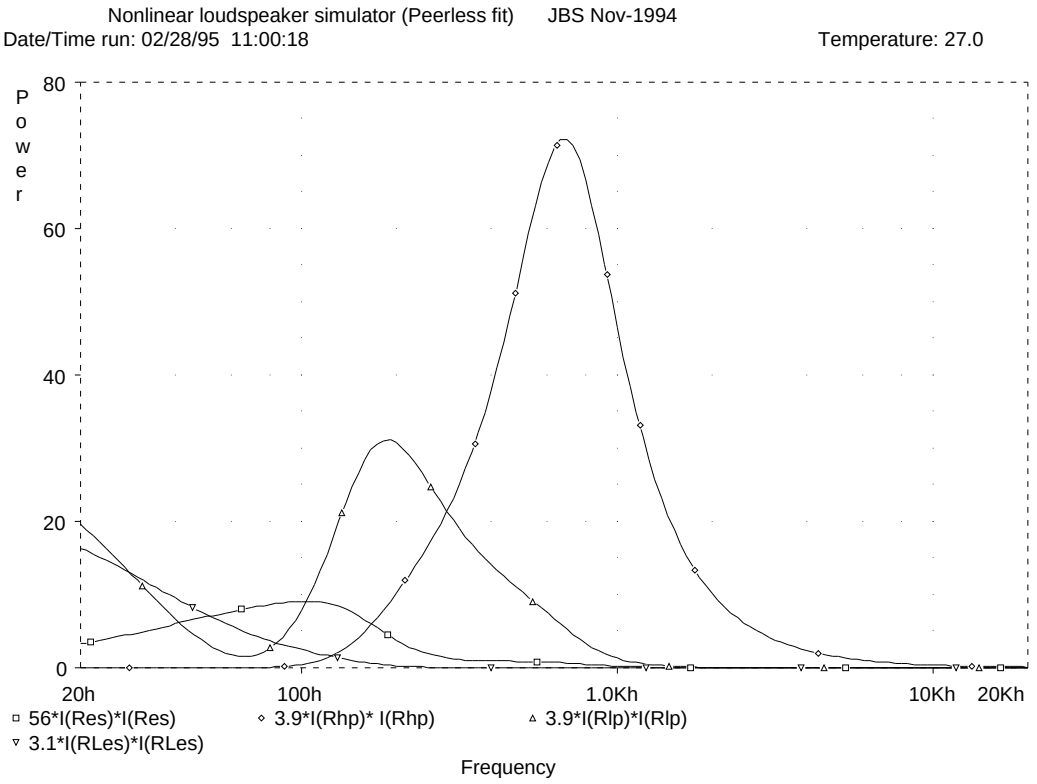

Figure D.15: Power dissipation in several resistances plotted against frequency with 25Vrms drive for the single-driver model circuit of figure D.11.

```
* resonant circuit of LF driver

RLes vvel intnl2 3.1

Les tail intnl2 39.8m

Ces vvel tail 150u

* this value allows for loss in Les

Res vvel tail 56

* A non-linearity at resonance

Gresnle vvel tail

+ VALUE={MAX(0,V(vvel,tail)

+               *abs(V(vvel,tail)))/1000}

* the tail crossover network:

Chi tail hi1 65.3u

Llo tail intnl3 4.3m

Rlo intnl3 lo1 1.1

* high-pass tail circuit
```

```
Rhp hi1 0  3.9
* low-pass tail circuit
Glpnle lo1 0
+ VALUE={MAX(0,V(lo1)*abs(V(lo1))/200)}
Rlp lo1 0 3.9
```

Bibliography

[1] Paul R. Gray and Robert G. Meyer, *Analysis and Design of Analog Integrated Circuits*, 2^{nd} Edition, Wiley, 1993.

[2] P. Horowitz and W. Hill, *The Art of Electronics*, 2^{nd} Edition, CUP, 1989.

[3] F. H. Mitchell, Jr, and F. H. Mitchell, Snr, *Introduction to Electronics Design*, Prentice Hall, 2^{nd} Edition, 1992.

[4] Richard C Dorf, *The Electrical Engineering Handbook*, Chapter 5, CRC Press, 1993.

[5] J B Scott, Analog Electronic Design, Prentice-Hall, 1991.

[6] Richard B. Childs and Vincent A. O'Byrne, "Predistortion linearization of directly modulated DFB lasers and external modulators for AM video transmission", Proc. OFC90, pp79, 1990.

[7] Douglas Self, "Distortion in Power Amplifiers", *Electronics and Wireless World*, August 1993 through March 1994.

[8] D. D. Weiner and J. F. Spina, *Sinusoidal Analysis and Modeling of Weakly Nonlinear Circuits*, Van Nostrand, New York, 1980.

[9] A. N. Thiele, "Measurement of Nonlinear Distortion in a Band-limited System", *Journal of the Audio Engineering Society*, vol. 31, June 1983, pp443–445.

[10] W. Marshall Leach, Jr, "SPICE Models for Vacuum-Tube Amplifiers", *Journal of the Audio Engineering Society*, vol. 43, no. 3, March 1995, pp117–126.

[11] P. W. Tuinenga, *SPICE: A Guide to Circuit Simulation and Analysis using PSPICE*, Prentice–Hall, 1988, or Second Edition, 1992.

[12] Paolo Antognetti and Giuseppe Massobrio, *Semiconductor Device Modelling with SPICE*, McGraw-Hill, 1988.

[13] Muhammad H. Rashid, *SPICE for Circuits and Electronics using PSpice*, Prentice-Hall, 1990.

[14] MicroSim Corporation, *The Design Centre, Analysis—Reference Manual*, Irvine, CA 92718, USA, 1992.

[15] MicroSim Corporation, *The Design Centre, Analysis—User's Guide*, Irvine, CA 92718, USA.

[16] Adly A. Girgis and Frederic M. Ham, "A quantitative study of pitfalls in the FFT", *IEEE Transactions on Aerospace and Electronic Systems*, vol. AES-16, no. 4, July 1980.

[17] A.V. Oppenheim and R.W. Schafer, *Digital Signal Processing*, Prentice Hall, 1975.

[18] G. D. Bergland, "A guided tour of the fast Fourier transform", in *Digital Signal Processing*, ed. L.R. Rabiner and C.M. Rader, IEEE Press, New York, 1972, pp228-239.

[19] Frederick J. Harris, "On the use of windows for harmonic analysis with the discrete Fourier transform", *Proceedings of the IEEE*, January 1978, vol. 66, no. 1, pp51-83.

[20] Roger G. Cox, "Window functions for spectrum analysis", *Hewlett-Packard Journal*, September 1978, pp10-11.

[21] William H. Press, Brian P. Flannery, Saul A. Teukolsky, and William T. Vetterling, *Numerical Recipes in C—The Art of Scientific Computing*, Cambridge University Press, 1988.

[22] Jonathan Scott and Dirk Heuer, "High Performance Total Difference Frequency Distortion Meter", *Journal of the Audio Engineering Society*, vol. 42, no. 6, June 1994, pp483–489.

[23] H. Hilton, "10-MHz ADC With 110dB Linearity," *High Speed ADC Conference*, Las Vegas, April 21–22, 1992.

[24] A. J. M. Kaizer, "Modeling of the Nonlinear Response of an Electrodynamic Loudspeaker by a Volterra Series Expansion", *Journal of the Audio Engineering Society*, vol. 35, no. 6, June 1987, pp421–432.

[25] P. G. L. Mills and M. O. J. Hawkesford, "Distortion Reduction in Moving-Coil Loudspeaker Systems Using Current-Drive Technology", *Journal of the Audio Engineering Society*, vol. 37, no. 3, March 1989, pp129–140.

[26] Wolfgang Klippel, "Dynamic Measurement and Interpretation of Nonlinear Parameters of Electrodynamic Loudspeakers", *Journal of the Audio Engineering Society*, vol. 38, no. 12, December 1990, pp944–954.

[27] Wolfgang Klippel, "The Mirror Filter—A New Basic for Reducing Nonlinear Distortion and Equalizing Response in Woofer Systems", *Journal of the Audio Engineering Society*, vol. 40, no. 9, September 1992, pp675–691.

[28] David R. Birt, "Nonlinearities in Moving-Coil Loudspeakers with Overhung Voice Coils", *Journal of the Audio Engineering Society*, vol. 39, no. 4, April 1991, pp219–231.

[29] Jonathan Scott and Glenn Leembruggen, "Loudspeaker Impedance Non-linearity and Bi-wiring", 95^{th} Convention of the Audio Engineering Society, Melbourne, August 1993.

[30] Erling S. Olsen, "Measurement of Mechanical Parameter Nonlinearities of Electrodynamic Loudspeakers", *98^{th} Convention of The Audio Engineering Society*, Paris, February 25–28, 1995.

[31] Jonathan Scott and Anthony Parker, "Trends in Device Characterisation: A Pulsed Semiconductor Parameter Analyser System for III-V FETs", *Journal of Electrical and Electronics Engineering, Australia*, vol. 14, no. 3, September 1994, pp. 196–205.

[32] J. R. Wright, "An Empirical Model for Loudspeaker Motor Impedance", *Journal of the Audio Engineering Society*, vol. 38, no. 10, October 1990, pp749–755.

[33] Richard H. Small, "Vented-Box Loudspeaker Systems Part 1: Small-Signal Analysis", *Journal of the Audio Engineering Society*, vol. 21, no. 5, June 1973, pp363–372.

[34] Richard H. Small, "Direct Radiator Loudspeaker System Analysis", *Journal of the Audio Engineering Society*, vol. 20, no. 5, June 1972, pp383–395.

[35] Milton Abramowitz and Irene A. Stegun, *Handbook of Mathematical Functions*, Dover Publications, Inc., New York, 1970.

[36] D. B. Keele Jnr., "Low-frequency Loudspeaker Assessment by Near-field Sound Pressure Measurement", *Journal of the Audio Engineering Society*, vol. 22, no. 3, April 1974, pp154–162.

[37] A. N. Thiele, "Measurement of Nonlinear Distortion in a Band-limited System", *IREE Convention Digest*, August 1975, pp480–482; republished in *Journal of the Audio Engineering Society*, vol. 31, June 1983, pp443–445.

[38] R. H. Small, "Total Difference-Frequency Distortion: Practical Measurements", *Journal of the Audio Engineering Society*, vol. 34, June 1986, pp427–436.

[39] J B Scott, "On Low Distortion Oscillator Design", *International Symposium on Circuits and Systems,* Portland, May 1989, pp1419-1422.

[40] M. E. Van Valkenburg, *Analog Filter Design*, Holt Saunders, 1982.

[41] D. R. Webster, D G Haigh and A E Parker, "Novel Circuit Synthesis Technique using Short Channel GaAs FETs giving Reduced Intermodulation Distortion", *IEEE International Symposium on Circuits and Systems*, April 30^{th}–May 3^{rd} 1995, Seattle, pp1348–1351.

[42] Haruhiko Koizumi, Shunsuke Nagata and Kunihiko Kanazawa, "A Technique for Minimising Intermodulation Distortion of GaAs FETs", *IEEE Transactions on Microwave Theory and Techniques*, vol. 43, no. 2, February 1995, pp463–466.

[43] W. Struble, S. L. G. Chu, M. J. Schindler, Y. Tajima and J. Huang, "Modelling Intermodulation Distortion in GaAs MESFETs using Pulsed I-V Characteristics",

Technical Digest of the IEEE GaAs IC Symposium, Monterey, California, October 20–23, 1991, pp179–182.

[44] Malcolm Hawkesford, "Reduction of Transistor Slope Impedance Dependent Distortion in Large-signal Amplifiers", *Journal of the Audio Engineering Society*, vol. 36, no. 4, April 1988, pp213–222.

[45] P.H. Ladbrooke, A.K. Jastrzebski, R.J. Donarski, J.P. Bridge and J.E. Barnaby, "Mechanism of drain current droop in GaAs MESFETs", *Electronics Letters*, vol. 31, no. 21, 12^{th} October, 1995, pp 1875–1876.

[46] Iltcho Angelov, Herbert Zirath and Niklas Rorsmann, "Validation of a Nonlinear Transistor Model by Power Spectrum Characteristics of HEMTs and MESFETs", *IEEE Transactions on Microwave Theory and Techniques*, vol. 43, no. 5, May, 1995, pp1046–1052.

[47] J. Rodriguez Tellez, S. Loredo and R. W. Clarke, "Self-Heating in GaAs FETs: A Problem?", *Microwave Journal*, September 1995, pp76–88.

[48] J. A. C. Stewart and J. G. Leckey, "Verification and Extraction of Device Model Parameters from Harmonic Amplitude and Phase", *High Performance Electron devices for Microwave and Optoelectronic Applications*, IEEE UK/ROI Section Workshop Proceedings, Oct 1993.

[49] J. G. Leckey et al., "Nonlinear MESFET Parameter Estimation Using Harmonic Amplitude and Phase Measurements", *IEEE MTT-S Digest*, May 1994, pp1563–1566.

[50] T. M. Barton, C. M. Snowden, J. R. Richardson and P. H. Ladbrooke, "Narrow pulse measurement of drain characteristics of GaAs MESFETs", *Electronics Letters*, vol. 23, pp686–687, 1987.

[51] M. Paggi, Paul H. Williams and Jose M. Borrego, "Nonlinear GaAs MESFET Modeling Using pulsed Gate Measurements", *IEEE Transactions on Microwave Theory and Techniques*, vol. 36, no. 36, December, 1988, pp. 1593–1597.

[52] A. Platzker, A. Palevski, S. Nash, W. Struble and Y. Tajima, "Characterisation of GaAs Devices by a Versatile Pulsed I-V Measurement System", *IEEE MTT-S Digest*, 1990, pp1137–1140.

[53] Anthony Parker and Jonathan Scott, "New Method for Comprehensive Characterisation of MES/MOD/MOS FETs", *Proceedings of the IEEE International Symposium on Circuits and Systems*, Chicago, 3–6 May, 1993, pp1093–1096.

[54] T. Fernandez, Y. Newport, J. M. Zamanillo, A. Mediavilla, and A. Tazon, "High speed automated pulsed I/V measurement system", *European Microwave Conference*, Madrid, Spain, 1993, pp494–496.

[55] Sam Prichett, Ryan Stewart, James Mason, Gailon Brehm, and David Christie, "Precision Pulsed I-V System for Accurate GaAs Device I-V Plane Characterisation", *IEEE MTT-S International Microwave Symposium Digest*, 23–27 May 1994, pp. 1353–1356.

[56] Jonathan Scott, Mohamed Sayed, Paul Schmitz and Anthony Parker, "Pulsed-bias/Pulsed-RF Device Measurement System Requirements", 24^{th} European Microwave Conference, Cannes, September 1994, pp951–961.

[57] Anthony Parker and Jonathan Scott, "Method to Determine Correct Timing for Pulsed-I/V Measurement of GaAs FETs", *Electronics Letters*, vol. 31, no. 19, 14 September, 1995, pp1697–1698.

[58] D. N. Swingler, "A Differential Technique for the Fourier Transform Processing of Multicomponent Exponential Functions", *IEEE Transactions on Biomedical Engineering*, vol. BME-24, no. 4, July 1977, pp408–410.

[59] James C. M. Hwang, "Gradual Degradation under RF Overdrive of MESFETs and PHEMTs", *GaAs IC Symposium*, San Diego, October 1995, pp81–84.

[60] Gaudenzio Meneghesso, Eros De Bortoli, Paolo Cova and Roberto Menozzi, "On Temperature and Hot Electron Induced Degradation in AlGaAs/InGaAs PM-HEMTs", *EDMO 95*, London, UK, October 1995.

[61] Anthony Parker, Jonathan Scott, James Rathmell and Mohamed Sayed, "New Applications for Pulsed/Isothermal Test System", *ARFTG*, invited paper, San Franscisco, June 21, 1996, pp70–75.

[62] J B Scott "On Low Distortion Oscillator Design", *International Symposium on Circuits and Systems,* Portland, May 1989, pp1419-1422.

[63] José Carlos Pedro and Jorge Perez, "Accurate Simulation of GaAs MESFET's Intermodulation Distortion Using a New Drain-Source Current Model", *IEEE Transactions on Microwave Theory and Techniques*, vol. 42, no. 1, January 1994, pp 25–33.

[64] Andrea M. Crossmun and Stephen A. Maas, "Minimization of Intermodulation Distortion in GaAs MESFET Small-Signal Amplifiers," *IEEE Transactions on MTT*, vol. 37, no. 9, pp. 1411–1417, September 1989.

[65] A. Werthof, F. van Raay, and G. Kompa, "Direct Nonlinear Power MESFET Parameter Extraction and Consistent Modelling", *Digest of the IEEE Microwave Theory and Techniques Symposium*, October 1993, pp645–648.

[66] Dirk Heuer and Jonathan Scott, "A New Total Difference Frequency Distortion (TDFD) Meter offering 135dB Range without Critical Filters", 94^{th} Convention of the Audio Engineering Society, Berlin, March 1993.

[67] Danny R Webster, Anthony E Parker, David G Haigh and Jonathan B Scott, "Effect of Circuit Parameters and Topology on Intermodulation in MESFET Circuits", IEEE GaAs IC Symposium, October 10–13, 1993, San Jose, California, pp255–258.

[68] Anthony Parker and Jonathan Scott, "Intermodulation Nulling in GaAs MESFETs" *Electronics Letters*, vol. 29, no. 22, 28 October, 1993.

[69] J B Scott, D R Webster, A E Parker, D G Haigh, James P M Proctor and Oliver J Ridler, "Improved Characterisation of FET Non-linear Behaviour", *Asia-Pacific Microwave Conference*, December 6–9, Tokyo, Japan, pp635–638.

[70] "A 60+60 Watt Audio Amplifier", *Electronics Australia*, May, 1986.

[71] Jonathan B. Scott and Greg Lemon, "A Nonlinear Load for Power Amplifier Testing", *Convention of the Audio Engineering Society*, Sydney, April, 1995.

[72] Jonathan Scott, Jonathan Kelly, and Glenn Leembruggen, "New Method of Characterising Driver Linearity", *Journal of the Audio Engineering Society*, vol. 44, no. 4, April 1996.

[73] A. S. Sedra and K. C. Smith, *Microelectronic Circuits*, Holt–Saunders, 1982.

[74] J. Millman and C. C. Halkias, *Intergrated Electronics: Analog and Digital Circuits and Systems*, McGraw–Hill, 1972.

[75] V. D. Hwang and T. Itoh, "An Efficient Approach for Large-Signal Modeling and Analysis of the GaAs MESFET", *IEEE Transactions on Microwave Theory and Techniques*, vol. 35, no. 5, pp. 396–402, April 1987.

[76] Paul Birdsall, BT Research Labs, Martlesham Heath, IP5-7RE, UK, private communication.

[77] George D. Vandelin, "Evaluating nonlinear models for microwave GaAsFETs", *IEEE Spectrum*, September 1990, pp 48–50.

[78] P. H. Ladbrooke and S. R. Blight, "Low-Field Low-Frequency Dispersion of Transconductance in GaAs MESFETs with Implications for Other Rate-Dependent Anomalies", *IEEE Transactions on Electron Devices*, vol. 35, no. 3, March 1988, pp. 257–267.

[79] David G. Haigh and Chris Toumazou, "Synthesis of Transconductor/Multiplier Circuits for Gallium Arsenide Technology", *IEEE Transactions on Circuits and Systems*, vol. 39, no. 2, February, 1992, pp 81–92.

[80] H. Statz, P. Newman, I. W. Smith, R. A. Pucel and H. A. Haus, "GaAs FET device and circuit simulation in SPICE", *IEEE Transactions on Electron Devices*, vol. ED-34, pp. 160–169, February 1987.

[81] J. Rodriguez-Tellez and P. J. England, "A five-parameter dc GaAs MESFET model for non-linear circuit design", *IEE Proceedings*, Part G, vol. 139, no. 3, pp. 325–332, June 1992.

[82] A. McCamant and D. Smith, "TOM–TriQuint's Own SPICE MESFET model", Internal Memo, TirQuint Semiconductor Inc., 12 June, 1989.

[83] A. J. McCamant, G. D. McCormack and D. H. Smith, "An Improved GaAs MESFET Model for SPICE", *IEEE Transactions on Microwave Theory and Techniques*, vol. 38, no. 6, pp. 822-824, June 1990.

[84] A. E. Parker and D. J. Skellern, "Improved MESFET Characterisation for Analog Circuit Design and Analysis", *IEEE GaAs IC Symposium Technical Digest*, Miami Beach, October 4–7, 1992, pp. 225–228.

[85] David E. Root, "A Measurement-based FET model improves CAE Accuracy", *Microwave Journal*, September 1991.

[86] D. E. Root, S. Fan, and J. Meyer, "Technology Independent Non Quasi-static FET Models by Direct Construction from Automatically Characterized Device Data", 21^{st} European Microwave Conference Proceedings, Stuttgart, Germany, September 1991, pp. 927–932.

[87] Norman Scheinberg, R. J. Bayruns, P. W. Wallace and R. Goyal, "An Accurate MESFET Model for Linear and Microwave Circuit Design", *IEEE Journal of Solid-State Circuits*, vol. 24, no. 2, pp. 532–539, April 1989.

[88] W. R. Curtice, "A MESFET Model for use in the design of GaAs integrated circuits", *IEEE Transactions on Microwave Theory and Techniques*, vol. 28, no. 5, pp. 448–456, May 1980.

[89] W. R. Curtice and M. Ettenberg, "A nonlinear GaAs FET model for use in the design of output circuits for power amplifiers", *IEEE Transactions on Microwave Theory and Techniques*, vol. 33, no. 12, pp. 1383–1394, December 1985.

[90] W. R. Curtice, "GaAs FET modelling and nonlinear CAD", *IEEE Transactions on Microwave Theory and Techniques* vol. 36, no. 2, pp220–230, February 1988.

[91] T. Kacprzak and A. Materka, "Compact DC model of GaAs FETs for Large-signal Computer Calculations", *IEEE Journal of Solid-state Circuits*, vol. SC-18, April 1983, pp. 211–213.

[92] M. Ali Khatibzadeh and R. J. Trew, "A Large-signal, Analytic Model for the GaAs MESFET", *IEEE Transactions on Microwave Theory and Techniques*, vol. 36, no. 2, February 1988, pp. 231–238.

[93] L. E. Larson, "An Improved GaAs MESFET Equivalent Circuit Model for Analog Integrated Circuit Applications", *IEEE Journal of Solid-State Circuits*, vol. 22, no. 4, pp. 567–574, August 1987.

[94] C. D. Hartgring, "An Accurate JFET/MESFET Model for Circuit Analysis", *Solid-State Electronics*, vol. 25, no. 3, 1982, pp. 233–240.

[95] F. Filicori, G. Vannini, A. Mediavilla, and A. Tazon, "Modelling of deviations between static and dynamic drain characteristics in GaAs FETs", *European Microwave Conference*, Madrid, Spain, 1993, pp. 454–457.

[96] A. E. Parker and D. J. Skellern, "GaAs Device Modelling for Design and Applications", *IEEE International Symposium on Circuits and Systems*, Singapore, June 11-14, 1991, pp. 1837–1840.

[97] A. E. Parker, "Implementing SPICE models with high-order continuity and rate dependence", *IEE Proceedings on Circuits, Devices and Systems*, vol. 141, no. 4, August 1994, pp251–257.

[98] Alan Podell, Pacific Monolithics, "Can we make designs work the first time?", Panel discussion D, *IEEE GaAs IC Symposium*, San Diego, October 24, 1989.

[99] M. Paggi, P. H. Williams, and J. M. Borrego, "Nonlinear GaAs MESFET Modeling Using Pulsed Gate Measurements", *IEEE Transactions on Microwave Theory and Techniques*, vol. 36, no. 12, December 1988, pp. 1593–1603.

[100] A. Platzker, A. Palevsky, S. Nash, W. Struble, and Y. Tajima, "Characterisation of GaAs Devices by a Versatile Pulsed I-V Measurement System", IEEE MTT-S Digest, 1990, pp. 1137–1140.

[101] Sam Prichett, Ryan Stewart, James Mason, Gailon Brehm, and David Christie, "Precision Pulsed I-V System for Accurate GaAs Device I-V Plane Characterisation", *IEEE MTT-S Digest*, San Diego, May 23–27, 1994, pp. 1353–1356.

[102] Jonathan Scott, Mohamed Sayed, Paul Smitz and Anthony Parker, "Pulsed-bias/Pulsed-RF Device Measurement System Requirements", to be published, 24^{th} European Microwave Conference, Cannes, September 1994.

[103] N. Jain, J. Gutmann, and D. M. Johnson, "Transient RF Signals During the Switching of MESFET Control Devices", *IEEE Transactions on Microwave Theory and Techniques*, vol. 39, no. 1, January 1991, pp. 18–24.

[104] Mohamed Sayed, Custom Systems Engineering, Hewlet-Packard, 1400 Fountaingrove Parkway, Santa Rosa, CA 95401, USA, private communication.

[105] Barry Fox, "High fidelity ain't what it used to be", *New Scientist*, Technology, 3 February 1996, p18.

[106] Danny R Webster, Jonathan B Scott, and David G Haigh, "Control of Circuit Distortion by the Derivative Superposition Method", *Microwave and Guided Wave Letters* vol 6, no. 2, 1995.

[107] C. T. Murray, "A Low Distortion Single-ended Push-pull Audio Amplifier", Proceedings of the IRE Australia, March 1960, pp134–137.

[108] Eric Benjamin, "Audio Power Amplifiers for Loudspeaker Loads", *Journal of the Audio Engineering Society*, vol. 42, no. 9, September 1994, pp670–683.