Table of Contents

1. Introduction 1
   1.1 Basic Physical Principles 1
   1.2 Acoustical Waveform Examples 3
   1.3 Speech Analysis and Synthesis Models 5
   1.4 The Linear Prediction Model 10
   1.5 Organization of Book 16

2. Formulations 18
   2.1 Historical Perspective 18
   2.2 Maximum Likelihood 20
   2.3 Minimum Variance 23
   2.4 Prony's Method 25
   2.5 Correlation Matching 31
   2.6 PARCOR (Partial Correlation) 32
      2.6.1 Inner Products and an Orthogonality Principle 35
      2.6.2 The PARCOR Lattice Structure 38

3. Solutions and Properties 42
   3.1 Introduction 42
   3.2 Vector Spaces and Inner Products 44
      3.2.1 Filter or Polynomial Norms 46
      3.2.2 Properties of Inner Products 47
      3.2.3 Orthogonality Relations 48
   3.3 Solution Algorithms 50
      3.3.1 Correlation Matrix 51
      3.3.2 Initialization 53
      3.3.3 Gram-Schmidt Orthogonalization 54
      3.3.4 Levinson Recursion 55
      3.3.5 Updating $A_m(z)$ 56
      3.3.6 A Test Example 57
   3.4 Matrix Forms 58

4. Acoustic Tube Modeling 60
   4.1 Introduction 60
   4.2 Acoustic Tube Derivation 61
      4.2.1 Single Section Derivation 63
      4.2.2 Continuity Conditions 65

4.2.3 Boundary Conditions 68
4.3 Relationship between Acoustic Tube and Linear Prediction 71
4.4 An Algorithm, Examples, and Evaluation 77
4.4.1 An Algorithm 78
4.4.2 Examples 80
4.4.3 Evaluation of the Procedure 82
4.5 Estimation of Lip Impedance 84
4.5.1 Lip Impedance Derivation 84
4.6 Further Topics 88
4.6.1 Losses in the Acoustic Tube Model 88
4.6.2 Acoustic Tube Stability 90

5. Speech Synthesis Structures 92
5.1 Introduction 92
5.2 Stability 93
5.2.1 Step-up Procedure 94
5.2.2 Step-down Procedure 95
5.2.3 Polynomial Properties 98
5.2.4 A Bound on $|F_m(z)|$ 99
5.2.5 Necessary and Sufficient Stability Conditions 101
5.2.6 Application of Results 102
5.3 Recursive Parameter Evaluation 103
5.3.1 Inner Product Properties 103
5.3.2 Equation Summary with Program 110
5.4 A General Synthesis Structure 113
5.5 Specific Speech Synthesis Structures 117
5.5.1 The Direct Form 118
5.5.2 Two-Multiplier Lattice Model 118
5.5.3 Kelly-Lochbaum Model 121
5.5.4 One-Multiplier Models 123
5.5.5 Normalized Filter Model 123
5.5.6 A Test Example 126

6. Spectral Analysis 129
6.1 Introduction 129
6.2 Spectral Properties 130
6.2.1 Zero Mean All-Pole Model 130
6.2.2 Gain Factor for Spectral Matching 130
6.2.3 Limiting Spectral Match 132
6.2.4 Non-uniform Spectral Weighting 134
6.2.5 Minimax Spectral Matching 136
6.3 A Spectral Flatness Model 139
6.3.1 A Spectral Flatness Measure 139
6.3.2 Spectral Flatness Transformations 141
6.3.3 Numerical Evaluation 142
6.3.4 Experimental Results 143
6.3.5 Driving Function Models 144

6.4 Selective Linear Prediction 146
6.4.1 Selective Linear Prediction (SLP) Algorithm 148
6.4.2 A Selective Linear Prediction Program 149
6.4.3 Computational Considerations 150
6.5 Considerations in Choice of Analysis Conditions 151
6.5.1 Choice of Method 151
6.5.2 Sampling Rates 153
6.5.3 Order of Filter 154
6.5.4 Choice of Analysis Interval 156
6.5.5 Windowing 157
6.5.6 Pre-emphasis 158
6.6 Spectral Evaluation Techniques 159
6.7 Pole Enhancement 161

7. Automatic Formant Trajectory Estimation 164
7.1 Introduction 164
7.2 Formant Trajectory Estimation Procedure 165
7.2.1 Introduction 165
7.2.2 Raw Data from $A(z)$ 167
7.2.3 Examples of Raw Data 169
7.3 Comparison of Raw Data from Linear Prediction and Cepstral Smoothing 172
7.4 Algorithm 1 176
7.5 Algorithm 2 180
7.5.1 Definition of Anchor Points 181
7.5.2 Processing of Each Voiced Segment 181
7.5.3 Final Smoothing 183
7.5.4 Results and Discussion 184
7.6 Formant Estimation Accuracy 185
7.6.1 An Example of Synthetic Speech Analysis 185
7.6.2 An Example of Real Speech Analysis 187
7.6.3 Influence of Voice Periodicity 188

8. Fundamental Frequency Estimation 190
8.1 Introduction 190
8.2 Preprocessing by Spectral Flattening 191
8.2.1 Analysis of Voiced Speech with Spectral Regularity 191
8.2.2 Analysis of Voiced Speech with Spectral Irregularities 193
8.2.3 The STREAK Algorithm 197
8.3 Correlation Techniques 199
8.3.1 Autocorrelation Analysis 200
8.3.2 Modified Autocorrelation Analysis 201
8.3.3 Filtered Error Signal Autocorrelation Analysis 203
8.3.4 Practical Considerations 206
8.3.5 The SIFT Algorithm 206

9. Computational Considerations in Analysis 212
9.1 Introduction 212
9.2 Ill-Conditioning 213
9.2.1 A Measure of Ill-Conditioning 214
9.2.2 Pre-emphasis of Speech Data 216
9.2.3 Prefiltering before Sampling 216
9.3 Implementing Linear Prediction Analysis 217
9.3.1 Autocorrelation Method 217
9.3.2 Covariance Method 219
9.3.3 Computational Comparison 220
9.4 Finite Word Length Considerations 222
9.4.1 Finite Word Length Coefficient Computation 223
9.4.2 Finite Word Length Solution of Equations 224
9.4.3 Overall Finite Word Length Implementation 225

10. Vocoders 227
10.1 Introduction 227
10.2 Techniques 229
10.2.1 Coefficient Transformations 229
10.2.2 Encoding and Decoding 233
10.2.3 Variable Frame Rate Transmission 235
10.2.4 Excitation and Synthesis Gain Matching 239
10.2.5 A Linear Prediction Synthesizer Program 242
10.3 Low Bit Rate Pitch Excited Vocoders 245
10.3.1 Maximum Likelihood and PARCOR Vocoders 246
10.3.2 Autocorrelation Method Vocoders 249
10.3.3 Covariance Method Vocoders 255
10.4 Base-Band Excited Vocoders 260

11. Further Topics 263
11.1 Speaker Identification and Verification 263
11.2 Isolated Word Recognition 265
11.3 Acoustical Detection of Laryngeal Pathology 267
11.4 Pole-Zero Estimation 271
11.5 Summary and Future Directions 275

References 278

Subject Index 285