CONTENTS

1. Introduction 1-27

1.1 Introduction 1
1.2. The Communication Process: Elements of a Communication System .. 2
1.3. Sources of Information 3
1.4. Communication Channels 6
1.5. Baseband and Passband Signals 9
1.6. Representations of Signals and Systems 9
1.7. Classification of Communication 11
1.7.1. Line Communication 11
1.7.2. Wireless or Radio Communication 11
1.8. The Modulation Process 11
1.8.1. Types of modulation 12
1.8.2. Need for modulation or Benefits of Modulation 12
1.9. Primary Communication Resources 13
1.10. Analog and Digital Signals 14
1.11. Conversion of Analog Signals to Digital Signals 15
1.12. Analog and Digital Comunication 16
1.12.1. Analog Communication 16
1.12.2. Digital Communication 17
1.12.3. Model of a Digital Communication System 17
1.13. Advantages and Disadvantages of Digital Communication 21
1.14. Fundamental Limitations of Communication System 21
1.14.1. Noise Limitations 22
1.14.2. Bandwidth Limitation 22
1.14.3. Equipment Limitation 22
1.15. Historical Perspective : The Evolution of Modern Electronic Communication ... 23
Summary 24
Review Questions 26
Objective Type Questions 27

2. Signal Analysis and Transmission 28-113

2.1. Introduction 28
2.2. Signal and Signal Classification 29
2.2.1. Continuous-Time and Discrete-Time Signals 29
2.2.2. Real and Complex Signals 30
2.2.3. Deterministic and Random Signals 30
2.2.4. Periodic and Aperiodic Signals 30
2.2.5. Even and Odd Signals 31
2.2.6. Energy and Power Signals 32
2.3. Singularity Functions 32
2.3.1. Unit Step Function 32
2.3.2. Unit Impulse Function 33
2.3.3. Unit Ramp Function 34
2.4. Representation of Signals 34
2.5. Fourier series 35
2.6. Trigonometric Fourier Series 35
2.6.1. Fourier Series Representation of a Periodic Function over the Entire Interval $(-\infty < t < \infty)$ 36
2.6.2. Symmetry Conditions (Few Points) 37
2.6.3. Polar Fourier Series Representation 42
2.6.4. Complex Fourier Exponential Series 42
2.6.5. Concept of Negative Frequency 44
2.6.6. Conversion of Trigonometric Fourier Series into Complex Exponential Fourier Series ... 44
2.6.7. Conversion of Complex Exponential Fourier Series into Trigonometric Fourier Series 44
2.7. Properties of Continuous-time Fourier Series (CTFS) 47
2.7.1. Linearity 47
2.7.2. Time Shifting Property 48
2.7.3. Frequency Shifting Property . 48
2.7.4. Time Scaling 48
2.7.5. Time-Differentiation 48
2.7.6. Time-Convolution 48
2.7.7. Multiplication 49
2.7.8. Symmetry Properties 49
2.8. Analysis of a Non-periodic Function over Entire Interval; Continuous-time the Fourier Transform 49
2.8.1. Fourier Transform Pair 50
2.8.2. Existence of Fourier Transform 51
2.8.3. Some useful Functions and Their Fourier Transforms 52
2.9. Properties of Fourier Transform 64
2.9.1. Time-Scaling Property 64
2.9.2. Linearity Property 65

2.9.3. Duality or Symmetry Property 65
2.9.4. Time Shifting Property 66
2.9.5. Frequency Shifting Property . 67
2.9.6. Time Differentiation Property 67
Fourier Transform of Periodic Functions 70
Convolution 70
.11.1. The Convolution Theorems 70
.11.2. Time Convolution Theorem ... 71
11.3. Frequency Convolution Theorem 71
nergy Signal 72
12.1. Parseval's Theorem for Energy Signals 72
12.2. Energy Spectral Density 76
12.3. Properties of Energy Spectral Density Function 78
ower Signals 79
13.1. Parseval's Theorem for Power Signals 79
13.2. The Power Spectral Density .. 80
13.3. Properties of Power Spectral Density Function 82
ross-Correlation 82
14.1. Cross-Correlation of Energy and Power Signals 83
14.2. Cross Correlation of Energy Signal 83
.4.3. Properties of Cross-correlation Function for Energy Signals .. 84
4.4. Cross-Correlation of Power or Periodic Signals 84
5.5. Properties of Cross-Correlation Function for Periodic Signals 84
tocorrelation Function 85
5.1. Autocorrelation Function of Energy Signals. 86
15.2. Properties of Auto-Correlation Function for Energy Signals . 86
15.3. Autocorrelation Function of Power (i.e., Periodic Signals ... 87
.15.4. Properties of Auto-Correlation Function for Periodic Signals 87
pplication of Fourier Transform to the nalysis of LTI Networks 89
2.16.1. Definition of Frequency-Response or Transfer-Functions 90
2.16.2. Concept of Distortionless Transmission through a System 92
2.16.3. Filters 93
Summary 108
Short Questions with Answers 110
Review Questions 112
Numerical Problems 112

3. Amplitude-Linear Modulation 114-228
3.1. Introduction 115
3.2. Multiplexing 115
3.3. Amplitude Modulation 116
3.3.1. Spectrum of AM Wave or Frequency-Domain Representation 117
3.4. Modulation Index 120
3.4.1. Overmodulation 120
3.5. Single Tone Amplitude Modulation (AM) 121
3.6. Power Content in AM Wave 123
3.7. Transmission Efficiency of Amplitude Modulated Signal 124
3.8. Power of a Single Tone Amplitude Modulated (AM) Signal 125
3.9. Current Calculation for Single Tone AM .....
3.10. Power Content in Multiple-tone Amplitude Modulation (AM) 12
3.10.1. Total or Net Modulation Index for Multiple-tone Modulation 129
3.11. Generation of Amplitude Modulation (AM) 130
3.11.1. Low Level Amplitude Modulation 130
3.11.2. High Level Amplitude Modulation 131
3.11.3. Non-linear Resistance or Non-Linear Circuits 131
3.11.4. Square Law Diode Modulation 133
3.12. Collector Modulation Method 135
3.13. Demodulation of AM Wave 136
3.13.1. Square-Law Detector 137
3.13.2. Linear Diode or Envelope Detector 138
3.14. Double Side Band Suppressed Carrier (DSB-SC) System 139
3.15. Generation of DSB-SC Signal 141
3.15.1. The Balanced Modulator 141
3.15.2. Ring Modulator 143
3.16. Demodulation of DSB-SC Signals 145
3.16.1. Synchronous Detection Method 145
3.16.2. Envelope Detection after Suitable Carrier Re-insertion 146
3.17. Quadrature-Amplitude Modulation 148
3.18. Effect of Phase and Frequency Errors in Synchronous Detection 149
3.19. Carrier Acquisition in DSB-SC System or Synchronization Techniques in DSB-SC system 150
3.20. Single sideband suppressed-Carrier (SSB-SC) Modulation 152
3.20.1. Time-Domain Description of the SSB-SC Wave 153
3.21. Hilbert Transform 155
3.22. SSB-SC for a General Modulating Signal 157
3.23. Generation of SSB-SC Signal 158
3.23.1. Frequency Discrimination Method 158
3.23.2. Phase-Shift Method 159
3.24. Demodulation of SSB-SC Signals 160
3.24.1. Phasor Diagram and Waveform of SSB-SC Signals 161
3.24.2. Waveform of SSB-SC with Large Carrier Signal 163
3.24.3. Detection of the SSB-SC Signal having Large Carrier 163
3.24.4. Compatible Single Sideband Signal 164
3.25. Vestigial Sideband (VSB) Modulation Systems 164
3.26. Generation of VSB Signals 165
3.26.1. The Time-Domain Representation of the VSB Signal 167
3.27. Demodulation of VSB Signals 169
3.28. Comparison of Various AM Systems .. 170
3.29. Frequency Translation and Mixing 170
Summary 220
Review Questons 224
Short Questions with Answers 224
Numerical Problems 227

4. Angle Modulation 229-299
4.1. Introduction 229
4.2. Concept of Angle Modulation: Basic Definition 230
4.2.1. Phase Modulation (PM) 231
4.2.2. Frequency Modulation 231
4.3. Frequency Deviation 232
4.4. Relationship Between Phase Modulation (PM) and Frequency Modulation (FM) 233
4.5. Single Tone Frequency Modulation 234
4.5.1. Mathematical Expression for Single Tone Frequency Modulation 236
4.6. Phasor Representation of Angle Modulated (i.e., FM and PM Waves) ... 239
4.7. Types of Frequency Modulation (FM) . 239
4.7.1. Narrowband FM 240
47.2. Generation of Narrowband FM 240
4.7.3. Expression for Single-Tone Narrowband FM 241
4.8. Wideband FM 242
4.9. Effect of Variation of Modulation Index m_f on the Spectrum of FM Signal 246
4.10. Transmission Bandwidth of FM Signal 247
4.10.1. Carson's Rule 248
4.11. FM bandwidth for an Arbitrary Modulating Signal $x(t)$ 249
4.12. Narrowband FM Versus Wideband FM 249
4.13. Multiple Freqeuncy Modulation 251
4.14. Square Wave Modulation 252
4.15. Linear and Nonliner Modulation 254
4.15.1. Linearity of FM Signal for Small Modulation Index 254
4.16. Phase modulation : An Analytic View 25
4.16.1. Transmission Bandwidth of PM 25
4.17. Comparison of Angle Modulated Waves and Amplitude Modulated Wave 25
4.18. Comparison of Frequency Modulation and Amplitude Modulation 25
4.19. FM generation 25
4.19.1. The Direct Method or Parameter Variation Method 25
4.19.2. Varactor Diode Method for FM Generation 25

20. Drawbacks of Direct Method for FM Generation 260
21. The Indirect or Armstrong Method of FM Generation 260
22. Practical Armstrong Method for FM Generation 260
23. FM Demodulators or Detectors 262
24. Types of FM Demodulators 262
4.24.1. Sloped Detector 262
4.24.2. Balanced Slope Detector 263
4.24.3. Disadvantages of Balanced Slope Detector 264
25. Phase Difference Detectors 264
4.25.1. Foster-Seeley Detector 264
4.25.2. Ratio Detectors 266
4.25.3. PLL-FM Demodulator 267
Summary 293
Short Questions with Answers 296
Review questions 297
Numerical Problems 298

Probability, Random Signals & Random Process 300-360
1. Introduction 300
2. Basic Definitions Related to Probability 301
5.2.1. Experiment 301
5.2.2. Sample Space 301
5.2.3. Event 302
3. Probability 302
4. Properties of Probability 302
5. Conditional Probability 303
. Probability of Statistically Independent Events 303
Random Variables 307
5.7.1. Discrete Random Variables 307
5.7.2. Continuous Random Variables 307
Probability Function or Probability Distribution of a Discrete Random Variable 308
Cumulative Distribution Function (CDF) 308
5.9.1. Properties of Cumulative Distribution Function (CDF) 309
5.9.2. Cumulative Distribution function (CDF) for discrete random variables 309
0. Probability Density Function (PDF) 310
5.10.1. Properties of Probability Density Function (PDF) 310
1. Joint Cumulative Distribution Function 317
5.11.1. Properties of Joint Cumulative Distribution Function 318
5.12. The Joint Probability Density Function 318
5.12.1. Properties of Joint PDF 318
5.12.2. The Relationship between Joint PDF and Probability 319
5.13. Marginal Densities 321
5.14. Conditional Probability Density Function 322
5.14.1. Properties of Conditional PDF 322
5.15. Statistical Averages of Random Variables 323
5.15.1. Mean or Average 323
5.15.2. Mean Value of Discrete Random Variable 323
5.15.3. Mean Values of Continuous Random Variables 323
5.15.4. Moments and Variance 324
5.16. Uniform Distribution 328
6.17. Gaussian or Normal Distribution 330
5.17.1. Properties of Gaussian PDF 331
5.18. Rayleigh Distribution 332
5.19. Random Process 334
5.19.1. Ensemble and Time Averages 335
5.19.2. Ensemble Averages 335
5.19.3. Time Averages 336
5.19.4. Stationary and Non-Stationary Random Process 336
5.19.5. Wide Sense Stationary (*i.e.*, weakly stationary) Process 336
5.19.6. Ergodic Process 337
5.19.7. Gaussian Process 337
5.20. Sum of Random Processes 338
5.21. Correlation Function 339
5.21.1. Autocorrelation Function 339
5.21.2. Properties of Autocorrelation Function of Random Process 339
5.21.3. Cross-Correlation Functions 340
5.22. Spectral Densities 341
5.22.1. Power Spectral Density (pds) 341
5.22.2. Energy Spectral Density (ESD) 343
5.23. Response of Linear Systems to Random Inputs 343
Summary *350*
Review Questions *355*
Short Question with Answers *355*
Numerical Problems *359*

6. Noise 361-402
6.1. Introduction 362
6.2. Classification of Noise 362
6.3. External Noise 362
6.4. Atmospheric Noise 362
6.5. Extraterrestrial Noise 362
6.5.1. Solar Noise 362
6.5.2. Cosmic Noise 363
6.6. Industrial Noise 363
6.7. Internal Noise 363
6.8. Shot Noise 363
6.8.1. Power Density Spectrum of Shot Noise in Diodes 364
6.9. Partition Noise 365
6.10. Flicker Noise, or Low-Frequency Noise 365
6.11. Transit-time Noise or High Frequency Noise 365
6.12. Thermal Noise 366
6.13. Voltage and Current Models of a Noisy Resistor 366
6.14. Addition of Noise due to Several Sources in Series 368
6.15. Addition of Noise due to Several Sources in Parallel 368
6.15.1. Addition of Noise due to Several Amplifiers in Series 369
6.16. Spectral Densities 370
6.17. Derivation of Thermal Noise Expressions Using Power Spectrum Density 372
6.18. Generalised Nyquist Theorem for Thermal Noise 373
6.19. Calculation of Thermal Noise for Single Noise Source 375
6.20. Thermal Noise in RC Circuit 375
6.21. Calculation of Thermal Noise for Multiple Noise sources: Super position of power Spectrum Density 376
6.22 White Noise 377
6.23. Equivalent Noise Bandwidth 378
6.24. Noise Equivalent Bandwidth for Low Pass Filter (LPF) 379
6.25. Nose in Reactive Circuits 381
6.26. Signal to Noise Ratio 383
6.27. Noise Figure 384
6.28. Noise Temperature 384
6.29. Calculation of Noise Figure 386
6.30. Noise Figure from Equivalent Noise Resistance 388
6.31. Experimental Determination of Noise Figure 390
6.32. Noise Figure Determination for Cascaded Stages of Amplifiers 392
Summary 398
Short Questions with Answers 401
Review Questions 402
Numerical Problems 402

7. Performance of Communication Systems 403-448
7.1. Introduction 403
7.2. Bandpass Noise Model 404
7.3. Noise in Communication System 405
7.3.1. Few Assumptions to Calculate the Figure of Merit for various Communication Systems 405
7.4. Noise Calculation in Various Amplitude Modulation (AM) Systems . 406
7.5. Double Sideband Suppressed Carrier (DSB-SC) System 407
7.6. Single Sideband Suppressed Carrier (SSB-SC) System 409
7.7. Amplitude-Modulation System (envelope Detector Method) 411
7.8. Evaluation of Noise in Angle Modulated Systems 413
7.9. Evaluation of Noise in Frequency Modulation (FM) Systems 414
7.10. Noise in Phase Modulation (PM) 421
7.11. Noise in Pulse-Modulated Systems 425
7.12. Noise in Pulse Amplitude Modulation (PAM) 425
7.13. Noise in Pulse Position Modulation (PPM) 426
7.14. Coded Systems 428
7.15. Noise in Binary PCM 428
7.16. Comparison of Coded and Uncoded Systems 431
Summary 445
Short Questions with Answers *446*
Review Questions 447
Numerical Problems 447

8. AM and FM Receivers 449-484
8.1. Introduction 449
8.2. Classification of Radio Receivers 450
8.3. Tuned Radio Frequency Receiver 450
8.3.1. Drawbacks of Tuned Radio Frequency (TRF) Receiver 451
8.4. Superheterodyne Receiver 452
8.5. AM Receivers 453
8.6. Sensitivity 453
8.7. Selectivity 453
8.8. Double Spotting 454
8.9. Tracking or Tuning of a Superheterodyne Receiver 454
8.10. Image Frequency and its Rejection 455
8.11. AM Superheterodyne Receiver (Description of Various Blocks) 457
8.12. RF Amplifier 457
8.13. A Series Resonance Circuit 459

8.14. Characteristics of a Series Resonant Circuit 460
8.15. A Parallel Resonance Circuit or Tuned Circuit 461
8.16. Characteristics of Parallel Resonant or Tuned Circuit 462
8.17. The Single Tuned Voltage Amplifier 464
8.18. The Frequency Response of a Single Tuned Voltage Amplifier 465
8.19. Limitation of a Single Tuned Voltage Amplifier 466
8.20. Double Tuned Voltage Amplifier 466
8.21. The Frequency Response of Double Tuned Voltage Amplifier 467
8.22. Stagger Tuned Voltage Amplifier 467
8.23. Local Oscillator 468
8.23.1. Frequency Mixers 469
8.23.2. Intermediate Frequencies and IF Amplifiers 470
8.23.3. Automatic Gain Control (AGC) or Automatic Volume Control (AVC) 472
8.23.4. Linear Diode Detector with Capacitor Filter and Simple AGC 472
8.24. Linear Diode Detector with π-Filter and Simple AGC 473
8.25. Delayed AGC 473
8.26. Amplified and Delayed AGC 474
8.27. AGC Characteristics 475
8.28. LIC Superheterodyne AM Receivers 475
8.29. AM Transmitter 476
Summary *481*
Short Questions with Answers *483*
Review Questions *484*

9. Sampling Theory and Pulse Modulation 485-533
9.1. Introduction 485
9.2. The Sampling Theorem 486
9.3. Proof of Sampling Theorem 486
9.3.1. Few points about sampling theorem 488
9.4. Nyquist Rate and Nyquist Interval 489
9.5. Reconstruction Filter (Low Pass Filter) 492
9.6. Signal Reconstruction: The Interpolation Formula 493
9.7. Effect of Under Sampling : Aliasing 495
9.8. Sampling of Bandpass Signals 495
9.9. Sampling Techniques 500
9.9.1. Ideal Sampling or Instantaneous Sampling or Impulse Sampling 500
9.9.2. Natural sampling 502
9.9.3. Flot Top Sampling or Rectangular Pulse Sampling 504
9.10. Aperture Effect 507
9.11. Comparison of Various Sampling Techniques 509
9.12. Analog Pulse Modulation Methods 511
9.13. Pulse Amplitude Modulation (PAM) 512
9.13.1. Naturally Sampled Pulse Amplitude Modulated Signal (PPM) 516
9.13.2. Instantaneous or Idealy Sampled Pulse Amplitude Modulated (PAM) signal 516
9.13.3. Transmission Bandwidth in Pulse Amplitude Modulation 517
9.13.4. Demodulation of PAM Signals 518
9.13.5. Transmission of PAM signals 519
9.13.6. Drawbacks of Pulse-amplitude modulated (PAM) signal 519
9.14. Pulse Time Modulation 519
9.14.1. Pulse Width Modulation 520
9.14.2. Frequency Spectrum for PWM Wave 520
9.14.3. Modulation of PWM Signal or PWM Generation 521
9.14.4. Demodulation of PWM Signal 522
9.14.5. Advantage of PWM 523
9.14.6. Disadvantages of PWM 523
9.15.7. Pulse Position Modulation 523
9.14.8. Generation of PPM Signal 524
9.14.9. Demodulation of PPM 525
9.14.10 Advantages of PPM 525
9.14.11 Disadvantages of PPM 527
9.15. Comparison of Various Pulse Analog Modulation Methods 526
Summary *526*
Short Questions with Answers *529*
Review Questions *531*
Numerical Problems *532*
Objective Type Questions *533*

10. Waveform Coding Techniques 534-586
10.1. Introduction 535
10.2. Discretization in Time and Amplitude 535
10.3. Concept of Quantization 536
10.4. Pulse Code Modulation (PCM) 536
10.5. A PCM Generator or Transmitter 538
10.6. PCM Transmission Path 539
10.7. PCM Receiver 540
10.8. Quantizer 541
10.8.1. Classification of Quantization Process 541
10.8.2. A Uniform Quantizer 541
10.9. Working Principle of Quantizer 542
10.10. A Uniform Quantizer with Incorrect Quantization Characteristics 543
10.11. Transmission Bandwith in a PCM System 544
10.12. Quantization Noise/Error in PCM 545
10.13. Signal to Quanitization Noise Ratio for Linear Quantization 548
10.14. Effect of Noise on the PCM System 549
10.15. Robust Quantization: Necessity of Non-uniform Quantization in a PCM System 558
10.16. Necessity of Non-uniform Quantization for Speech Signal 559
10.17. Nonuniform Quantization 561
10.18. Companding (*i.e.*, Companded PCM) 561
10.18.1. Compressor Characteristic 561
10.18.2. Expander Characteristics 562
10.19. Compander Characteristic 562
10.20. Different types of Compressor Characteristics 562
10.20.1. μ-law Companding 562
10.20.2. A-law Comanding 563
10.21. Applications of PCM 564
10.22. Advantages of PCM 564
10.23. Disadvantages of PCM 564
10.24. Delta Modulation 564
10.24.1. Advantages of Delta Modulation 566
10.24.2. Disadvantages of Delta Modulation: 566
10.25. Adaptive Delta Modulation 567
10.25.1. Advantages of Adaptive Delta Modulation 568
10.26. Differential Pulse Code Modulation (DPCM) 569
10.26.1. Reconstruction of DPCM Signal 571
10.27. Comparison of Digital Pulse Modulation Methods 571
Summary *580*
Short Questions with Answers *582*
Review Questions *584*
Numerical Problems *585*
Objective Type Questions *585*

11. Digital Multiplexers 587-607
11.1. Introduction 587
11.2. Multiplexing 587
11.2.1. Frequency Division Multiplexing (FDM) 587
11.2.2. Time Division Multiplexing (TDM) 588
11.3. A PAM/TDM system 588
11.3.1. Evaluation of Signaling Rate in a PAM/TDM System 590
11.3.2. Transmission Bandwidth of a PAM/TDM Channel 590
11.3.3. Synchronization in PAM/TDM System 591
11.3.4. Concept of Crosstalk in a PAM/TDM System 591
11.3.5. Concept of Guard Time 592
11.3.6. Advantages of TDM 592
11.3.7. Disadvantages of TDM 592
11.4. Introduction to Digital Multiplexers 595
11.4.1. Principle of Digital Multiplexing 596
11.5. Classification of Digital Multiplexers 596
11.6. Multiplexing Hierarchy for Digital Communication 596
11.6.1. Calculation of Bit Rate at each Level 597
11.6.2. Bit Rate and Transmission Channel Bandwidth 598
11.6.3. Bandwidth Efficiency 598
11.6.4. Advantages of Digital Multiplexing 598
11.7. North American Hierarchy 600
11.8. T Lines 600
11.9. A PCM-TDM System (T1 Carrier System) 601
11.10. E Lines 603
Summary 604
Short Questions with Answers 605
Review Questions 607

12. Digital Baseband Transmission 608-684
12.1. Introduction 609
12.2. A baseband digital communication system 609
12.3. Introduction to Discrete PAM signals (Digital Data Formats) 611
12.4. Line Coding and its Properties 612
12.5. Various PAM Formats of Line Codes 613
12.6. Unipolar RZ and NRZ 614
12.7. Polar RZ and NRZ 615
12.8. Bipolar NRZ 616
12.9. Split Phas Manchester Format 616
12.10. Polar Quaternary NRZ Format 616
12.11. High density bipolar (HDB) signalling 617
12.12. B8ZS Line Code 618

12.13. Power Spectra of Discrete PAM Signals (Various Line Codes) 618
12.14. Power Spectral Density (psd) of NRZ Unipolar Format 619
12.15. Power Spectral Density (psd) of NRZ Polar Format 621
12.16. Power Spectral Density (psd) of NRZ Bipolar Format 621
12.17. Power Spectral Density (psd) of Manchester Format 622
12.18. Comparison of Various Discrete PAM Formats on the Basis of Power Spectra 625
12.19. Difference between Source Coding and Line Coding 627
12.20. Introduction to Matched Filter and Intersymbol Interference (ISI) 627
12.21. Integrate and Dump Filter (*i.e.*, Receiver) 628
12.21.1. Calculation of Signal to Noise Radio for the Integrate and Dump Filter 629
12.22. The Optimum Filter (i.e., Optimum Receiver) 638
12.22.1. Calculation of Probability of Error (P_e) for Optimum Filter 638
12.22.2. Evaluation of Transfer Function for the Optimum Filter 640
12.23. Matched Filter 643
12.23.1. Calculation of Impulse Response for the Matched Filter 643
12.23.2. Calculation of Probability of Error (P_e) for the Matched Filter 646
12.24. The Correlator 653
12.25. Intersymbol Interference (ISI) 655
12.26. Cause of Intersymbol Interference (ISI) 656
12.27. Nyquist's Criterion for distortionless baseband binary transmission 657
12.27.1. Freqency Domain Representation 658
12.28. Ideal Solution 659
12.29. Concept of Eye Pattern 660
Summary *680*
Short Questions with Answers *682*
Review Questions *683*
Numerical Problems *684*

13. Introduction to Signal Space Analysis 685-696
13.1. Introduction 685
13.2. Concept of Additive White Gaussian Noise (AWGN) Channel 686
13.3. Concept of Optimum Receiver 686
13.4. Geometric Representation of Signals . 687
13.5. Schwarz Inequality 690
13.6. Gram-Schmidt Orthogonalization Procedure 691
Summary *694*
Short Questions with Answers *695*
Review Questions *696*

14. Digital Modulation Techniques 697-743
14.1. Introduction 697
14.2. Digital modulation formats 698
14.3. Types of digital modulation techniques 699
14.4. Coherent binary modulation techniques 699
14.5. Coherent binary amplitude shift keying or on-off keying 699
14.5.1. Signal space diagram of ASK 700
14.5.2. Generation of ASK signal 700
14.5.3. Coherent demodulator of binary ASK 701
14.6. Binary Phase Shift Keying 702
14.6.1. Generator of BPSK signal 703
14.6.2. Reception (*i.e.*, Detection) of BPSK Signals 703
14.6.3. The spectrum of BPSK signal 705
14.6.4. A geometrical representation for BPSK signals 707
14.6.5. Bandwidth for BPSK signal 708
14.6.6. Error probability of BPSK Signal Employing Coherent Reception or Matched Filter 708
14.6.7. Drawbacks of BPSK 709
14.7. Coherent binary frequency shift keying (BFSK) 710
14.7.1. Generation of BFSK 710
14.7.2. The spectrum of BFSK signal 711
14.7.3. Bandwidth of BFSK signal 712
14.7.4. Detection of BFSK 713
14.7.5. Geometrical representation of Orthogonal BFSK 713
14.7.6. Geometrical representation of Non-orthogonal BFSK signals 714
14.7.7. Probability of Error, P_e of BFSK Signals 715
14.7.8. Advantages and disadvantages of BFSK signals 715
14.8. Non-coherent binary modulation techniques 715
14.9. Non-coherent Binary Amplitude Shift Keying (ASK) 715
14.10. Non-coherent detection of FSK 716
14.11 Differential Phase Shift Keying (DPSK) 716
14.11.1. Generation of DPSK 716
14.11.2. Detection of DPSK 718
14.11.3. Evaluation of Bandwidth of DPSK signal 719
14.11.4. Error Probability of DPSK ... 719
14.11.5. Advantages and disadvantages of DPSK 720
14.12. Quadrature Phase Shift (*i.e.*, Detection of GPSK) Seying (QPSK) 721
14.13. Generation of QPSK 721
14.13.1. Reception of QPSK (*i.e.*, Detection of GPSK) 723
14.13.2. Concept of Carrier Synchronization in QPSK ... 725
14.13.3. Signal space representation of QPSK signals 725
14.13.4. Spectrum of QPSK signal 727
14.13.5. Bandwidth of QPSK signal 728
14.13.6. Probability of Error of QPSK System 729
14.13.7. Advantages of QPSK 730
14.14. Minimum shift keying (MSK) 730
14.14.1. Signal space representation of MSK and distance between the signal points (*i.e.*, Geometrical Representation of MSK) 733
14.14.2. Power spectral density and bandwidth of MSK 734
14.14.3. Generation of MSK 735
14.14.4. Reception of MSK 738
14.14.5. Advantages and Disadvantages of MSK as Compared to QPSK 739
14.15. Comparison of Digital Modulation Techniques 739
Summary *739*
Short Questions with Answers *741*
Review Questions *742*
Numerical Problems *743*

15. Information Theory 744-797
15.1. Introduction 744
15.2. What is information ? 745
15.3. Information Sources 745
15.4. Information content of a discrete memoryless source (DMS) 746
15.5. Information content of a symbol (*i.e*, Logarithmic Measure of Information) 746
15.6. Entropy (*i.e.* Average Information) 749
15.7. Information Rate 750
15.8. The Discrete Memoryless Channels (DMC) 757
15.8.1. Channel representation 757
15.8.2. The Channel matrix 757
15.9. Types of Channels 758
15.9.1. Lossless Channel 758
15.9.2. Deterministic Channel 759
15.9.3. Noiseless Channel 759
15.9.4. Binary Symmetric Channel 759
15.10. The Conditional and Joint Entropies . 762
15.11. The Mutual Information 763
15.12. The Channel Capacity 768
15.12.1. Channel Capacity Per Symbol C_s 768
15.12.2. Channel Capacity Per Second C 768
15.12.3. Capacities of Special Channel 768
15.13. Entropy Relations for a Continuous Channel 770
15.14. Capacity of an Additive White Gaussian Noise (AWGN) Channel: Shannon-Hartley Law 771
15.15. Channel Capacity 771
15.16. Transmission of Continuous Signals . 773
15.17. Uncertainty in the Transmission Process 774
15.18. Exchange of Bandwidth for Signal-to-Noise Ratio 775
15.19. The Source Coding 780
15.19.1. The Code Length and Code efficiency 780
15.19.2. The Source Coding Theorem 780

15.19.3. Classification of Codes 780
15.19.4. The Kraft Inequality 781
15.20. Entropy Coding 785
15.20.1. Shannon-Fano Coding 786
15.20.2. The Huffman Encoding 786
Summary *791*
Short Questions with Answers *794*
Review Questions *795*
Numerical Problems *795*

16. Error Control Coding 798-861
16.1. Introduction 798
16.2. Few Terms Related to Error Control Coding 799
16.3. Parity Coding 801
16.4. Vertical Redundancy Check (VRC) 801
16.4.1. Few Points 801
16.4.2. Longitudinal Redundancy check (LRC) 802
16.4.3. Few Points 803
16.4.4. Advantages of LRC Over VRC 804
16.5. Linear Block Codes 804
16.5.1. Steps for Determination of all the Codewords for a Linear Block Code 805
16.5.2. Concept of Parity Check Matrix (H) for Linear Block Code 807
16..6. Hamming Codes 808
16.6.1. Error Detection and Correction Capabilities of Hamming Codes 809
16.7. Encoder of (7, 4) Hammin Code 812
16.8. Syndrome Decoding: Method to Correct errors 812
16.9. Error Correcting using Syndrome Vector 816
16.9.1. What Happens if Double Error Occurs in Y 817
16.9.2. Hamming Bound 818
16.10. Syndrome Decoder for (n, k) Block Code 818
16.11. Other Linear Block Code 825
16.11.1. Single Parity Check Bit Codes 825
16.11.2. Repeated Codes 825
16.11.3. Hadamard Code 825
16.11.4. Extended Codes 827
16.12. Cyclic codes 829
16.13. Generator and Parity Check Matrices of Cyclic Codes 835
16.14. Encoders for Cyclic Codes 839
16.14.1. Syndrome Decoding for Cyclic Codes 841
16.14.2. Decoder for Cyclic Codes 843
16.14.3. Advantages and Disadvantages of Cyclic Codes 844
16.15. BCH Codes (Bose-Chaudhuri-Hocquenghem) Codes 844
16.16. Convolutional codes 844
16.16.1. Code Trees, Trellis, and State Diagram for a Convolution Encoder 846
16.16.2. Decoding methods of Convolutional Codes 849
16.16.3. Free Distance and Coding Gain 853
16.16.4. Advantages and Disadvantages of Convolutional Codes 855
Summary *856*
Short Questions with Answers *857*
Review Questions *860*
Numerical Problems *861*

17. Spread-Spectrum Techniques 862-878
17.1. Introduction 862
17.2. Principles of spread spectrum 862
17.2.1. Direct sequence psuedonoise (DSPN) Spreading 863
17.2.2. Frequency HOP Spreading 864
17.3. Time Hop spreading 864
17.4. Spread-spectrum modulation system 865
17.5. Generation of Psuedorandom Sequences 866
17.5.1. The sequence length 867
17.5.2. Independence of sequences 868
17.5.3. Number of 1's and 0's in a maximal sequence 868
17.5.4. Clustering in an PN sequence 868
17.5.5. The Correlation Properties 869
17.6. The Spread Spectrum Modulation 869
17.6.1. Continuous-Phase-Shift Modulation (CPSM) 870
17.7. The Reception of spread-spectrum signal 871
17.7.1. Synchronization 873
17.8. CDMA 874
Summary *875*
Short Questions with Answers *876*
Review Questions *878*

18. Cellular and Mobile Communications 879-904
18.1. Introduction 879
18.2. Main Methods of Radio Transmission 880
18.3. GSM Standards for Cellular Telephony 881
18.4. Architecture of GSM 881
18.5. The cellular mobile radio systems 884
18.6. Structure of Cellophane 885
18.7. Working 885
18.8. Other services of GSM 886
18.9. Performance criteria for cellular phones 886
18.9.1. Voice quality 887
18.9.2. Service quality 887
18.9.3. Special features 888
18.10. Operation of cellular systems 888
18.11. The concept of frequency reuse channels 889
18.11.1. Frequency reuse schemes 889
18.11.2. Frequency reuse distance 889
18.12. Consideration of the components in a cellular systems 890
18.13. The power control for cellular systems 891
18.14. Function of the MTSO 891
18.15. Cellular analog switching equipment 892
18.15.1. Modified analog switching equipment 892
18.15.2. Cell-site controllers and Hardware 894
18.16. The cellular digital switching equipment 894
18.16.1. General concept 894
18.16.2. Elements of switching 895
18.17. Digital Mobile Telephony 895
18.18. MTSO Interconnection 897
18.18.1. Connection to Wire-Line Network 897
18.18.2. Connection to a Cell site 898
18.19. Radio Paging Systems 898
18.19.1. Concept of Paging 898
18.19.2. The components of a Paging system 900
18.19.3. Operational features of a Paging system 901
18.19.4. Pager (Subscriber's Radio) 902
Summary *902*
Short Questions with Answers *903*
Review Questions *904*

Objective Type Questions 905-934
Appendices 935-944
Model Question Papers 945-952
References 953-956
Index 957-964

Exclusive by

www.dahiyastech.blogspot.com

@sTc(small team creation)

CHAPTER 1

Introduction

Inside this Chapter

1.1. Introduction 1
1.2. The Communication Process: Elements of a Communication System 2
1.3. Sources of Information 3
1.4. Communication Channels 6
1.5. Baseband and Passband Signals 9
1.6. Representations of Signals and Systems 9
1.7. Classification of Communication 11
1.8. The Modulation Process 11
1.9. Primary Communication Resources 13
1.10. Analog and Digital Signals 14
1.11. Conversion of Analog Signals to Digital Signals 15
1.12. Analog and Digital Communication 16
1.13. Advantages and Disadvantages of Digital Communication 21
1.14. Fundamental Limitations of Communication System 21
1.15. Historical Perspective: The Evolution of Modern Electronic Communication 23

1.1. Introduction

Communication is the process of establishing connection or link between two points for information exchange.

OR

Communication is simply the process of conveying message at a distance or communication is the basic process of exchanging information.

The electronic equipments which are used for communication purpose, are called communication equipments. Different communication equipments when assembled together form a communication system.

Typical examples of communication system are line telephony and line telegraphy, radio telephony and line telegraphy, radio broadcasting, point-to-point communication and mobile communication, computer communication, radar communication, television broadcasting, radio telemetry, radio aids to navigation, radio aids to aircraft landing etc.

The earliest communication system namely line-telegraphy originated in eighteen fourties (1840s). In addition to this, line telephony came a few decades later whereas radio-communication could become possible in the beginning of twentieth century on invention of triode valve. Radio commu-

nication was further greatly improved during World War II. It became more widely used through the invention of transistor, integrated circuits (ICs) and other semiconductor devices in the subsequent years. Also, in recent years, communication has become more widespread with the use of satellites and fibre optics. Today, there has been an increasing emphasis on the use of computers in communication.

1.2. The Communication Process: Elements of a Communication System

In the most fundamental sense, communication involves the transmission of information from one point to another through a succession of process as listed ahead:

(*i*) the generation of a thought pattern or image in the mind of an originator.
(*ii*) the description of that image, with a certain measure of precision, by a set of aural or visual symbols.
(*iii*) the encoding of these symbols in a form that is suitable for transmission over a physical medium of interest.
(*iv*) the transmission of the encoded symbols to the desired destination.
(*v*) the decoding and reproduction of the original symbols.
(*vi*) the recreation of the original thought pattern or image, with a definable degradation in quality, in the mind of a recipient.

Hence, as discussed above, the purpose of a communication system is to transmit an information-bearing signal, from a source, located at one point, to a user or destination, located at another point some distance away. Figure 1.1 shows the block diagram of a general communication system* in which the different functional elements are represented by blocks. The essential components of a communication system are information source, input transducer, transmitter, communication channel, receiver and destination.

Now, we shall briefly discuss the functioning of these blocks.

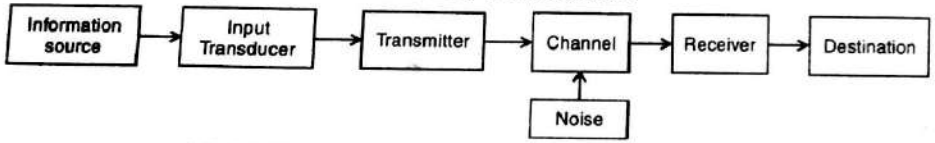


Fig. 1.1. *Block diagram of a communication system*

1. Information Source

We know that a communication system serves to communicate a message or information. This message or information originates in the information source. In general, there can be various messages in the form of words, groups of words, code, symbols, sound signal etc. However, out of these messages, only the desired message is selected and conveyed or communicated.

In short, we can say that the function of information source is to produce required message which has to be transmitted.

2. Input Transducer

A transducer is a device which converts one form of energy into another form. The message from the information source may or may not be electrical in nature. In a case when the message produced by the information source is not electrical in nature, an input transducer is used to convert it into a time-varying electrical signal.** For example, in case of radio-broadcasting, a microphone converts the information or message which is in the form of sound waves into corresponding electrical signal.

* It is also known as elements of a communication system.

** Whenever a message signal produced by information source is not electrical in nature, we have to convert it into an electrical signal because all the signal processing devices available are electrical in nature.

3. Transmitter

The function of the transmitter is to process the electrical signal from different aspects. For example in radio broadcasting the electrical signal obtained from sound signal, is processed to restrict its range of audio frequencies (upto 5 kHz in amplitude modulation radio broadcast) and is often amplified. In wire telephony, no real processing is needed. However, in long-distance radio communication or broadcast, signal amplification is necessary before modulation. Modulation is the main function of the transmitter. In modulation, the message signal is superimposed upon the high-frequency carrier signal. In short, we can say that inside the transmitter, signal processings such as restriction of range of audio frequencies, amplification and modulation are achieved. All these processings of the message signal are done just to ease the transmission of the signal through the channel.

4. The Channel and the Noise

With reference to the block diagram of a communication system in figure 1.1, the term channel means the medium through which the message travels from the transmitter to the receiver. In other words, we can say that the function of the channel is to provide a physical connection between the transmitter and the receiver.

There are two types of channels, namely point-to-point channels and broadcast channels. Examples of point-to-point channels are wirelines, microwave links and optical fibres. Wirelines operate by guided electromagnetic waves and they are used for local telephone transmission. In case of microwave links, the transmitted signal is radiated as an electromagnetic wave in free space. Microwave links are used in long-distance telephone transmission. An optical fiber is a low-loss, well-controlled, guided optical medium. Optical fibres are used in optical communications. Although these three channels operate differently, they all provide a physical medium for the transmission of signals from one point to another point. Therefore, for these channels, the term point-to-point is used.

On the other hand, the broadcast channels provide a capability where several receiving stations can be reached simultaneously from a single transmitter. An example of a broadcast channel is a satellite in geostationary orbit, which covers about one-third of the earth's surface.

During the process of transmission and reception the signal gets distorted due to noise introduced in the system.

Noise is an unwanted signal which tend to interfere with the required signal. Noise signal is always random in character. Noise may interfere with signal at any point in a communication system. However, the noise has its greatest effect on the signal in the channel.

5. Receiver

The main function of the receiver is to reproduce the message signal in electrical form from the distorted received signal. This reproduction of the original signal is accomplished by a process known as the demodulation or detection. Demodulation is the reverse process of modulation carried out in transmitter.

6. Destination

Destination is the final stage which is used to convert an electrical message signal into its original form. For example in radio broadcasting, the destination is a loudspeaker which works as a transducer *i.e.* it converts the electrical signal in the form of original sound signal.

1.3. Sources of Information

As a matter of fact, the telecommunication environment is dominated by the following four important sources of information:

(*i*) speech
(*ii*) television

(iii) facsimile, and

(iv) personal computers

A source of information may be characterised in terms of the signal which carries the information. Further, a signal is defined as a single-valued function of time that plays the role of the dependent variable. At every instant of time, the function has a unique value.

(i) Speech

We know that speech is the primary method for human communication. Basically, the speech communication process involves the transfer of information from a speaker to a listener, which takes place in three successive stages as under:

(a) *Production:* An intended message in the speaker's mind is represented by a speech signal that consists of sounds (*i.e.*, pressure waves) generated inside the speaker's mouth and whose arrangement is governed by the rules of language.

(b) *Propagation:* The sound waves propagate through the air, reaching the listener's ears.

(c) *Perception:* The incoming sounds are deciphered by the listener into a received message, and thus completing the chain of events that results in the transfer of information from the speaker to the listener.

In fact, the speech-production process may be viewed as a form of filtering, in which a sound source excites a vocal tract filter. The vocal tract consists of a tube of non-uniform cross-sectional are, beginning at the glottis (*i.e.*, the opening between the vocal cords) and ending at the lips. Now, as the sound propagates along the vocal tract, the spectrum (*i.e.*, the frequency content) is shaped by the frequency selectivity of the vocal tract. This effect is quite similar to the resonance phenomenon observed in organ pipes. At this stages, it may be noted that the hearing mechanism is highly sensitive to frequency. In addition to this, the type of communication system being considered has an important bearing on the bond of frequencies considered to be essential for the communication process. As an example, a frequency band from 300 to 3100 Hz is considered adequate for telephonic communication on a commercial basis.

(ii) Television

The second source of information, television (TV), refers to the transmission of pictures in motion by means of electrical signals. To accomplish this transmission, each complete picture has to be sequentially scanned. The scanning process is carried out in a TV camera. In a black-and-white TV, the camera contains optics designed to focus an image on a photocathode consisting of a large number of photosensitive elements.

The charge pattern so generated on the photosensitive surface is scanned by an electron beam, and thus producing an output current which varies temporally in accordance with the way in which the brightness of the original picture varies spatially from one point to another. This resulting output current is called a **video signal**.

The type of scanning used in television is a form of spatial sampling called **raster scanning**. The purpose of **raster scanning** is basically to convert a two-dimensional image intensity into a one-dimensional waveform.*

Further, in television, a picture is divided into 525 lines, which constitute a frame. Now, each frame is decomposed into two interlaced fields, each one of which consists of 262.5 lines. For simplicity of presentation, we can refer to the two fields as I and II. Figure 1.2 illustrates the scanning procedure.

The lines of field *I* are depicted as solid lines, and the lines of field *II* are depicted as dashed lines. The figure also includes the start and end of each field. Field *I* is scanned first. The scanning spot of the TV camera moves with constant velocity across each line of the field from left to right. The image intensity at the center of the spot is measured. Now, when the end of a particular line is

* In fact, it is quite similar to the manner in which we read a printed paper in that the scanning is performed from left to right on a line-by-line basis.

reached, the scanning spot quickly flies back (in a horizontal direction) to the start of the next line down in the field. This flyback is known as the **horizontal retrace**. The scanning process is continued until the whole field has been covered. When this condition is reached, the scanning spot moves quickly (in a vertical direction) from the end of field *I* to the start of field *II*.

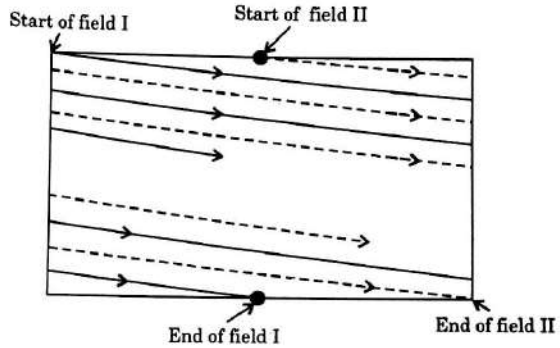


Fig. 1.2. *Illustration of interlaced raster scan*

This second flyback is known as the **vertical retrace**. Field *II* is treated in the same fashion as field *I*. The time taken for each field to be scanned is 1/60 second.

Correspondingly, the time taken for a frame or a complete picture to be scanned is 1/30 second. With 525 lines in a frame, the line-scanning frequency equals 15.75 kHz. Hence, by flashing 30 still pictures per second on the display tube of the TV receiver, the human eye perceives them to be moving pictures. This effect is due to a phenomenon known as the **persistence of vision**.

Also, the reproduction quality of a TV picture is limited by the following two basic factors:

(a) *the number of lines available in a raster scan, which limits resolution of the picture in the vertical direction.*

(b) *the channel bandwidth available for transmitting the video signal which limits resolution of the picture in the horizontal direction.*

(iii) Facsimile

The purpose of the third source of information, facsimile (fax) machine, is to transmit still pictures over a communication channel (most notably, a telephone channel). Such a machine provides a highly popular facility for the transmission of hand written or printed text from one point to another. Further, transmitting text by facsimile is treated simply like transmitting a picture.

The basic principle employed for signal generation in a facsimile machine is to scan an original document (*i.e.*, picture) and use an image sensor to convert the light to an electrical signal.

(iv) Personal Computers

As a matter of fact, personal computers (PCs) are becoming increasingly an important part of our daily lives. We use them for electronic mail, exchange of software, and sharing of resources. It is estimated that over 30 per cent of the personal computers in use today are already networked and the number is increasing rapidly. The text transmitted by a PC is usually encoded using the American Standard Code for Information Interchange (ASCII), which is the first code developed specifically for computer communications.

Each character in ASCII is represented by seven data bits constituting a unique binary pattern made up of 0s and 1s. Hence, a total of $2^7 = 128$ different characters can be represented in ASCII. The characters are various lowercase and uppercase letters, numbers, special control symbols, and punctuation symbols commonly used such as @, $, and %. Some of the special "control" symbols such

as BS (back space) and CR (carriage return), are used to control the printing of characters on a page. Other symbols, such as ENQ (enquiry) and ETB (end of transmission block), are used for communication purposes.

The seven data bits are ordered starting with the most significant bit b_7 down to the least significant bit b_1 as shown in figure 1.3.

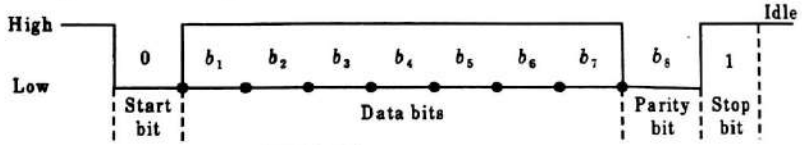


Fig. 1.3. *An ASCII format*

At the end of the data bits, an extra bit b_8 is appended for the purpose of error detection. This error-detection bit is called a **parity bit**. A sequence of eight bits is referred to as a byte or an octet. The parity bit is set in such a way that the total number of 1s in each byte is odd for odd parity, or even for even parity.

Personal computers are often connected via their RS (recommended standard)-232 ports. When ASCII data are transmitted through these ports, a start bit, set to 0, and one or more stop bits, set to 1, as depicted in figure 1.3, are added to provide character framing. When the transmission is idle, a long series of 1s is sent so as to keep the circuit connective alive.

Note: *Computer-generated data and television signals are both wideband signals in that their power content occupies a wide range of frequencies. Another important characteristics of data communication between personal computers is burstiness, which means that information is usually transmitted from one terminal to another in bursts with silent periods between bursts.*

1.4. Communication Channels

(BHU, 2003)

In last article, we have discussed several sources of information which are responsible for the generation of message signals. These message signals need to be transmitted to the destination point. Now, let us discuss various communication channel* that carryout the transmission.

Basically, we shall consider following channels:

(*i*) telephone channels
(*ii*) optical fibers
(*iii*) mobile radio channels, and
(*iv*) satellite channels.

These four different channels have been chosen because they play important roles in modern telecommunications environment.

(*i*) Telephone Channels

A telephone network makes use of a switching mechanism. This switching mechanism is known as circuit switching and it is used to establish an end-to-end communication link on a temporary basis. Infact, the primary purpose of the network is to ensure that the telephone transmission between a speaker at one end of the link and a listener at the other end, is an acceptable replacement for face-to-face conversation.

In this form of communication, the message source is the sound produced by the speaker's voice and the destination is the listener's ear. However, it may be noted that the telephone channel supports only the transmission of electrical signals.

* That is physical media.

Because of this reason, appropriate transducer are used at the transmitting and receiving ends of the system.

Further, the telephone channel is essentially a linear, bandwidth-limited channel. A speech signal (male or female) is essentially limited to a band from 300 to 3100 Hz in the sense that frequencies in lying outside this band do not contribute much to articulation efficiency. Hence, this particular frequency band may be viewed as a *rough guideline* for the passband of a telephone channel that provides a satisfactory service.

(*ii*) Optical Fibers

The second communication channel, an optical fiber, is a dielectric waveguide which transports light signals from one place to another just as a metallic wire pair or a co-coxial cable, transports electrical signals.

An optical fiber consists of a central core within which the propagating electromagnetic field is confined and which is surrounded by a cladding layer, which is itself surrounded by a thin protective jacket. Basically, the core and cladding are both made of pure silica glass, whereas, the jacket is made of plastic. Optical fibers have unique characteristics that make them highly attractive as a transmission medium. Optical fibers offer the following unique advantages:

(a) Enormous Potential Bandwidth

For optical fibers, enormous potential bandwidth results from the use of optical carrier frequencies around 2×10^{14} Hz. With such a high carrier frequency and a bandwidth roughly equal to 10 per cent of the carrier frequency, the theoretical bandwidth of a light wave system is around 2×10^{13} Hz, which is very large indeed.

(b) Low Transmission Losses

Optical fibers offer low transmission losses as low as 0.2 dB/km.

(c) Immunity to Electromagnetic Interference

In fact, immunity to electromagnetic interference is an inherent characteristic of an optical fiber which is viewed as a dielectric waveguide.

(d) Small Size and Weight

For optical fibers, small size and weight are characterized by a diameter which is no greater than that of a human air.

(e) Ruggedness and Flexibility

Optical fibers offer unique advantage of ruggedness and flexibility because of very high tensile strengths and the possibility of being bent or twisted without damage.

Note: *All these unique properties of optical fibers have fueled phenomenal advances in light-wave systems technology during the past two decades, which in turn, have revolutionized long-distance communications.*

(*iii*) Mobile Radio Channels

The third communication channel, mobile radio channel, extends the capability of the publi telecommunications network by introducing mobility into the network by virtue of its ability t broadcast.

The term 'mobile radio' is usually meant to encompass terrestrial situations where a radi transmitter or receiver is capable of being moved, regardless of whether it actually moves or no Basically, there is no 'line-of-sight' path for communication, rather, radio propagation takes plac mainly by way of scattering from the surfaces of the surrounding buildings and by diffraction ove and/or around them.

Therefore, the energy reaches the receiving antenna via more than one path. Therefore, in mobile radio environment, we face a problem of multipath phenomenon in the sense that the variou

incoming radio waves reach their destination from different directions and with different time delays. Because of this, the received signal strength varies with location in a very complicated manner and therefore, a mobile radio channel may be viewed as a linear time-varying channel that is statistical in nature.

(iv) Satellite Channel

A satellite channel provides broad-area coverage in a continental as well as intercontinental sense. Moreover, access to remove areas not covered by conventional cable or fiber communications is also a distinct feature of satellites.

In almost all satellite communication systems, the satellites are placed in geostationary orbit. For the orbit to be geostationary, it has to satisfy two requirements. First, the orbit is geosynchronous, which requires the satellite to be at an altitude of 22,300 miles, a geosynchronous satellite orbits the earth in exactly 24 hours. Second, the satellite is placed in orbit directly above the equator on an eastward heading.

Viewed from earth, a satellite in geostationary orbit appears to be stationary in the sky.

Communications satellites in geostationary orbit offer the following unique system capabilities:

(*a*) broad-area coverage

(*b*) reliable transmission links

(*c*) wide transmission bandwidth

In terms of services, satellites can provide fixed point-to-point links extending over long distances and into remote areas, communication to mobile platforms (e.g., aircraft, ships), or broadcast capabilities.

In a typical satellite communication system, a message signal is transmitted from an earth station via an uplink to a satellite, amplified in a transponder on board the satellite. It is then retransmitted from the satellite via a downlink to another earth station. In effect, the satellite acts as a powerful repeater in the sky. The most popular frequency band for satellite communications is 6 GHz for the uplink and 4 GHz for the downlink.

In fact, the use of this frequency band offers the following advantages:

(*a*) relatively inexpensive microwave equipment

(*b*) low attenuation due to rain fall*

(*c*) insignificant sky background noise.

In the 6/4 GHz band, a typical satellite is assigned a 500 MHz bandwidth that is divided among 12 transponders on board the satellite. Each transponder, using approximately 36 MHz of the satellite bandwidth, corresponds to a specific radio channel.

A single transponder can carry at least one colour television signal, 1200 voice circuits, or digital data at a rate of 50 Mb/s.

Note: *To summarize, we can conclude that a communication channel is central to the operation of a communication system. Its properties determine both that information-carrying capacity of the system and the quality of service offered by the system.*

As a matter of fact, communication channels may be classified in different ways as under:

(*a*) A channel may be linear or nonlinear. A telephone channel is linear, whereas a satellite channel is usually (but not always) nonlinear.

(*b*) A channel may be time invariant or time varying. An optical fiber is time invariant whereas a mobile radio channel is time varying.

(*c*) A channel may be bandwidth limited or power limited (*i.e.*, limited in the available transmitted power). A telephone channel is bandwidth limited whereas an optical fiber link and a satellite channel are both power limited.

* Rainfall is a primary atmospheric cause of signal loss.

1.5. Baseband and Passband Signals

(Pune University, 2000)

As discussed earlier, the message signal generated from the information source is known as baseband signal. This baseband signal may be a combination of two or more message signals. If the baseband signal is transmitted directly, then it is known as **baseband transmission**. This means that the baseband transmission does not used modulators and demodulators. The baseband signal may be both analog as well as digital.

The analog baseband signal varies continuously with time and has continuous amplitude. The digital baseband signal is discrete in both time and amplitude.

The baseband transmission is preferred at low frequencies and for short distances. However, Inter symbol Interference (ISI) is the major problem associated with the baseband transmission.

Now, if the modulated signal is transmitted over the channel, it is known as **bandpass** or **simply passband transmission**. Whenever a modulating signal is impressed upon a carrier, the modulated signal is produced. This modulated signal has fixed band of frequencies around carrier frequency. In fact, the nature of such a signal is bandpass type. Therefore, modulated signals are known as passband or bandpass signals. As an example, if a 1 kHz modulating signal is impressed upon a 100 kHz carrier signal by using amplitude modulation technique then two sidebands will be generated. Lower sideband will be at (100 – 1) *i.e.*, 99 kHz and upper sideband will be at (100 + 1) *i.e.*, 101 kHz.

Thus, the amplitude modulated signal will have frequencies from 99 kHz to 101 kHz. These frequencies are bandpass type. This means that the modulated signals are also bandpass *i.e.*, passband type.

It may be noted that the bandpass transmission is generally used at high frequencies and for long distances. All television and satellite transmission are bandpass type. The bandpass transmission may be both analog and digital. If the digital signal modulates the carrier, then it is known as digital passband transmission.

1.6. Representations of Signals and Systems*

In order to develop knowledge of the operation of a communication system, we require some mathematical tools for the representations of signals and systems. As a matter of fact, the most useful method of signal representation depends upon the particular type of signal under consideration.

Therefore, depending upon the feature of interest, we may distinguish following three different classes of signals which arise in the study of communication systems:

(*i*) Periodic and nonperiodic signals

(*ii*) Deterministic and random signals

(*iii*) Energy and power signals.

(i) Periodic and Nonperiodic Signals

A periodic signal $g(t)$ is a function of time which satisfies the following condition:

$$g(t) = g(t + T_0) \text{ for all values of } t$$

where, t denotes time, and

T_0 is a constant ...(1.1)

Now, the smallest value of T_0 for which this condition is satisfied is called the period of $g(t)$. Accordingly, the period T_0 defines the duration of one complete cycle of $g(t)$. Also, any signal for which equation (1.1) does not hold for any T_0 is called a nonperiodic or simply aperiodic signal.

(ii) Deterministic and Random Signals

A deterministic signal is a signal about which there is no uncertainty with respect to its value at any time, past, present or future.

Accordingly, deterministic signals may be modeled as completely specified functions of time.

* This particular topic has been covered in detail in next chapter "Signal Analysis and Transmission."

On the other hand, a random signal is a signal about which there is some degree of uncertainty before it actually occurs.

(iii) Energy and Power Signals

We already know that in electrical systems, a signal may represent a voltage or a current. Now, let us consider a voltage $v(t)$ developed across a resistor R, producing a current $i(t)$.

Now, the instantaneous power dissipated in this resistor may be defined as under:

$$p(t) = \frac{|v(t)|^2}{R} \quad ...(1.2)$$

above expression may also be modified as

$$p(t) = R\,|i(t)|^2 \quad ...(1.3)$$

It may be noted that in both the above cases, the instantaneous power $p(t)$ is proportional to the squared amplitude of the signal. Further, for a resistor R equal to 1 ohm, equations (1.2) and (1.3) both take on same mathematical form.

Accordingly, in signal analysis, it is customary to normalize the calculations by working with a 1 ohm resistor, so that, regardless of whether a given signal $g(t)$ represents a voltage or a current, we may express the instantaneous power associated with the signal simply as under:

$$p(t) = |g(t)|^2 \quad ...(1.4)$$

Now, based upon above convention, the total energy of a signal $g(t)$ may be defined as

$$E = \lim_{T\to\infty} \int_{-T}^{T} |g(t)|^2\, dt \quad ...(1.5)$$

or

$$F = \int_{-\infty}^{\infty} |g(t)|^2\, dt$$

Correspondingly, the average power of a signal $g(t)$ is defined as

$$P = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} |g(t)|^2\, dt \quad ...(1.6)$$

Now, we can say that the signal $g(t)$ is an energy signal if and only if the total energy of the signal satisfies the condition

$$0 < E < \infty$$

Also, it can be concluded that the signal $g(t)$ is a power signal if and only if the average power of the signal satisfies the condition

$$0 < P < \infty$$

Note: *It may be noted that the energy and power classifications of signals are mutually exclusive. In particular, an energy signal has zero average power, whereas a power signal has infinite energy. Also, it will be interesting to note that, usually, periodic signals and random signals are power signals and signal that are both deterministic and non-periodic are energy signals.*

Theoretically, there are several possible methods for the representation of signals. However, practically, Fourier analysis, involving the resolution of signals into sinusoidal components, is most important among all other methods. Basically, this is a consequence of the well-known fact that the response of a system to a sine-wave input is another sine-wave of the same frequency* under following two conditions:

(*i*) the system is linear in that it obeys the principle of superposition.

* But with a different phase and amplitude.

This means that if $y_1(t)$ and $y_2(t)$ denote the responses of a system to the inputs $x_1(t)$ and $x_2(t)$, respectively, the system is linear if the response to the composite input $a_1x_1(t) + a_2x_2(t)$ is equal to $a_1y_1(t) + a_2y_2(t)$, where a_1 and a_2 are arbitrary constants.

(*ii*) the system is time invariant

This means that if $y(t)$ is the response of a system to the input $x(t)$, the system is time invariant if the response to the time-shifted input $x(t - t_0)$ is equal to $y(t - t_0)$ where t_o is constant.

Note: *There are several methods of Fourier analysis available for the representation of signals. The particular method that is used, in practice, basically depends upon the type of signal under consideration. As an example, if the signal is periodic, then the logical choice is to use the Fourier series to represent the signal in terms of a set of harmonically related sine waves.*

On the other hand, if the signal is an energy signal, then it is customary to use the Fourier Transform to represent the signal. By using the Fourier series or the Fourier Transform, we obtain the frequency domain description or spectrum of the signal.

1.7. Classification of Communication

Regarding the mode of propagation, communication may be divided in the following two forms:

(*i*) Line Communication

(*ii*) Wireless or Radio Communication

1.7.1. Line Communication

In line communication, the medium of transmission is a pair of conductors called transmission line. This is also called as line channel. This means that in line communication, the transmitter and the receiver are connected through a wire or line. However, the installation and maintenance of a transmission line is not only costly and complex but also overcrowds the open space. Apart from this, its message transmission capability is also limited.

1.7.2. Wireless or Radio Communication

In wireless or radio communication, a message is transmitted through open space by electromagnetic waves called as *radio waves*. Radio waves are radiated from the transmitter in open space through a device called *antenna*. A receiving antenna intercepts the radio waves at the receiver. All the radio, TV and satellite broadcasting are wireless or radio communication. The advantages of wireless communication are cost effectiveness, possible long distance communication and simplicity.

1.8. The Modulation Process

As discussed earlier, the purpose of a communication system is to deliver a message signal from an information source in recognizable form to a user destination, with the source and the user being physically separated from each other.

To do this, the transmitter modifies the message signal into a form which is suitable for transmission over channel. This modification is achieved by means of a process known as modulation, which involves varying some parameter of a carrier wave in accordance with the message signal. Also the receiver recreates the original message signal from a degraded version of the transmitted signal after propagation through the channel.* This recreation is achieved by using a process called as **demodulation**. In fact, the demodulation is the reverse of the modulation process used in the transmitter. We can define modulation process in a more systematic form as under:

Modulation may be defined as the process by which some characteristic of a signal called carrier is varied in accordance with the instantaneous value of another signal called ***modulating signal.***

* The carrier performs certain functions in radio communication.

Signals containing information or intelligence are referred as modulating signals. This information bearing signal is also called ***baseband signal****. The carrier frequency is greater than the modulating frequency. The signal resulting from the process of modulation is called* ***modulated signal****.*

1.8.1. Types of Modulation

Modulation is basically of two types:

(*i*) Continuous Wave Modulation: When the carrier wave is continuous in nature, the modulation process is known as continuous wave (CW) modulation or analog modulation. Examples of continuous wave modulation are Amplitude Modulation and Angle Modulation. When the amplitude of the carrier is varied in accordance with the message signal, it is known as **amplitude modulation (AM)**. Also, when the angle of the carrier is varied according to the instantaneous value of the modulating signal, it is called **angle modulation**. Angle modulation may be further subdivided into Frequency modulation (FM) and Phase modulation (PM), in which the instantaneous frequency and phase of the carrier, respectively, are varied in accordance with the message signal.

(*ii*) Pulse Modulation: When the carrier wave is a pulse-type waveform, the modulation process is known as pulse modulation. In pulse modulation, the carrier consists of a periodic sequence of rectangular pulses. Pulse modulation can be of an analog or digital type.

In analog pulse modulation, the amplitude, duration or position of a pulse is varied in accordance with sample values of the message signal. The analog pulse modulation may be of following three types:

(*i*) pulse-amplitude modulation (PAM)

(*ii*) pulse-duration modulation (PDM).

(*iii*) pulse-position modulation (PPM).

On the other hand, the digital form of pulse modulation is known as **pulse-code modulation (PCM)**.

1.8.2. Need for Modulation or Benefits of Modulation

As discussed earlier, the message signal or baseband signal is used to modulate a high frequency carrier signal inside the transmitter. After modulation, the resulting modulated signal is transmitted with the help of an antenna which is connected at the output side of the transmitter. This modulated signal then travels down the channel to reach at the input of the receiver.

Now, one question can arise why we use modulation in communication system or what will happen if we transmit message signal or audio signal without modulation. The answer is that the modulation serves several purpose in communication system as discussed below:

(*i*) ***Practicality of antenna:*** We know that in case when free space is used as a transmitting medium (*i.e.* channel), messages are transmitted and received with the help of antennas. For efficient radiation and reception the transmitting and receiving antennas must have lengths comparable to a quarter-wavelength of the frequency used. For example, in AM broadcast systems, the maximum audio frequency transmitted from a radio station is of the order of 5 kHz. If this message audio signal were to be transmitted without modulation, then the height of the antenna required for an effective radiation and reception will be 1/4th of the wavelength given as

$$l = \frac{\lambda}{4} = \frac{c}{4f} = \frac{3\times10^8}{4\times5\times10^3} = 5 \text{ km}$$

Obviously, it will be totally impracticable to construct and install an antenna of such a height. However, this height of the antenna may be reduced by modulation technique and yet effective radiation and reception is achieved. In modulation process, low frequency or audio signal at radio stations are translated to higher frequency spectrum *i.e.*, radio frequency range. These higher radio frequencies with the small wavelength act as carrier for the audio frequencies (*i.e.* modulating signal). Thus, the height of the antenna required is much reduced and becomes practical.

As an example, if an audio frequency is translated to a radio frequency carrier of frequency 4 MHz, the antenna height required would be

$$l = \frac{\lambda}{4} = \frac{c}{4f} = \frac{3\times10^8}{4\times3\times10^6}$$

or

$$l = \frac{1}{4}\times10^2 = 0.25\times100 = 25 \text{ metres}$$

@sTc

This antenna height may be achieved practically.

(*ii*) ***To remove interference:*** Another reason for not radiating modulating signal itself is that the frequency range of audio signal is from 20 Hz to 20 kHz. In radio-broadcasting, there are several radio stations. In case, there is no modulation, all these stations transmit audio or sound signals in the range of 20 Hz to 20 kHz. Due to this transmission over same range, the programmes of different stations will get mixed up.

Hence, in order to keep the various signals separate, it is necessary to translate or shift them to different portions of the electromagnetic spectrum. Thus each station is allocated a band of frequency. This also overcomes the drawback of poor radiation efficiency at low frequency.

As an example, in Amplitude Modulation radio-broadcast, the maximum modulating signal frequency permitted is 5 kHz. Amplitude Modulation requires a bandwidth of 10 kHz for each station or channel. Therefore, broadcast channels can be placed adjacent to each other, each channel occupying 10 kHz bandwidth. Hence, different stations may be allotted bandwidths say from 790 to 800 kHz, 800 to 810 kHz and so on. In radio receiver, a tuned circuit at the input selects the desired station and rejects all other stations.

(*iii*) ***Reduction of noise:*** Noise is the major limitation of any communication. Although noise can not be eliminated completely, but with the help of several modulation schemes, the effect of noise can be minimized.

1.9. Primary Communication Resources *(WBTU, Kolkata, 2000)*

As a matter of fact, in a communication system, there are following two primary communication resources to be employed:

(*i*) transmitted power, and

(*ii*) channel bandwidth

The transmitted power refers to the average power of the transmitted signal.

The channel bandwidth is defined as the band of frequencies allocated for the transmission of the message signal.

Now, the most important system design objective is to use these two resources as efficiently as possible. In most communication channels, one resource may be considered more important than the other. Because of this, we may classify communication channels as power limited or band limited.

For example, the telephone circuit is a typical band limited channel, whereas a space communication link or a satellite channel is typically power-limited channel.

Let us consider a particular case as under:

When the spectrum of a message signal extends down to zero or low frequencies, we define the bandwidth of the signal as that upper frequency above which the spectral content of the signal is negligible and therefore unnecessary for transmitting information. For example, the average voice spectrum extends well beyond 10 kHz, though most of the average power is concentrated in the range of 100 to 600 Hz, and a band from 300 to 3100 Hz gives good articulation.

Accordingly, we find that telephone circuits that respond well to this latter range of frequencies give quite satisfactory commercial telephone service.

Another aspect of utmost importance is the unavoidable presence of noise in a communication system. A quantitative way of accounting for the effect of noise is to introduce *signal-to-noise ratio*

(SNR) as a system parameter. For example, we may define the SNR at the receiver input as the ratio of the *average signal power to the average noise power*, both being measured at the same point.*

A fundamental question that arises in the study of modulation schemes is the following:

With channel bandwidth and signal-to-noise ratio being the two principal parameters that are available to the designer of a communication system, which particular modulation scheme provides for their most efficient use? or in other words we can say for a specified channel bandwidth, which modulation scheme requires the smallest signal to nose ratio for a prescribed level of system performance.

The answer to this important question lies with information capacity theorem by Shannon.

Information Capacity Theorem

The information capacity theorem, given by Shannon, deals with a continuous channel. According to this theorem, channel bandwidth and signal-to-noise ratio are *exchangeable* in the sense that we may trade off one for the other for a prescribed system performance. The choice of one modulation scheme over another for the transmission of a message signal in the communication system is often dictated by the nature of this trade off. Infact, the interplay between channel bandwidth and signal-to-noise ratio, and the limitation that they impose on communication, is highlighted by the *information capacity theorem*.

Let B denote the channel bandwidth, and SNR denote the received signal-to-noise ratio. The information capacity theorem states that ideally these two parameters are related by

$$C = B \log_2 (1 + \text{SNR}) \ b/s$$

where C is the information capacity.

The *information capacity* is defined as the maximum rate at which information may be transmitted without error through the channel. It is measured in *bits per second* (b/s). Above equation clearly shows that for a prescribed information capacity, we may reduce the required SNR by increasing the channel bandwidth, hence the advantage of using a broad bandwidth to transmit messages. Moreover, above equation provides an idealized framework for comparing the noise performance of one modulation scheme against another.

1.10. Analog and Digital Signals

Before we discuss analog and digital communications, let us have a quick view of analog and digital signals.

The analog signal is that type of signal which varies smoothly and continuously with time. This means that analog signals are defined for every value of time and they take on continuous values in a given time interval. Thus, we can say that analog messages are characterized by data whose values vary over a continuous range. The signal depicted in figure 1.4 is an analog signal. In fact, for analog signal, the name derives from the fact that such a signal is analogous to the physical signal that it represents. The vast majority of signals in the world around us are analog. For example, the temperature or the atmospheric pressure of a certain location may vary over a continuous range and may assume an infinite number of possible values. Similarly, a speech waveform is an analog signal since it has amplitudes that vary over a continuous range.

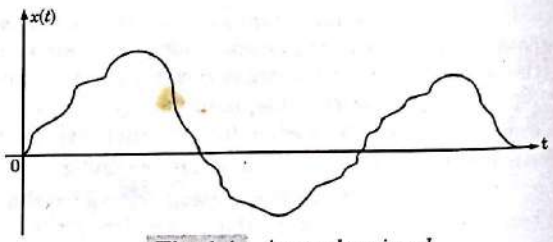


Fig. 1.4. *An analog signal.*

* The customary practice is to express the SNR in decibels (dBs), defined as 10 times the logarithm (to base 10) of the power ratio.

An alternative form of signal representation is that of a sequence of numbers, each number representing the signal magnitude at an instant of time. The resulting signal is called a digital signal. Digital messages are constructed with a finite number of symbols. For example, the printed language consists of 26 letters, 10 numbers, a space and several punctuation marks. Therefore, any text is a digital message constructed from about 50 symbols.

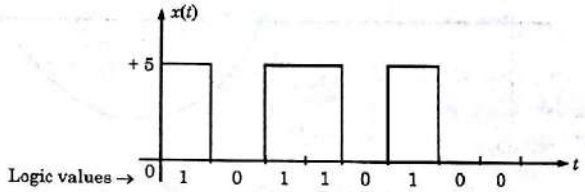


Fig. 1.5. *Variation of a binary digital signal with time.*

Now, since a digital signal is represented only by digits, therefore, we can use any number system to represent a digital signal. However, in practice, we generally use binary number system to represent a digital signal. In a binary system, each digit in the number takes on one of only two possible values, denoted 0 and 1. Correspondingly, the digital signals in binary systems need have only two voltage levels which may be labeled low and high.

Figure 1.5 shows a digital signal. Observe that the waveform is a pulse train with 0 V representing a '0' signal or logic '0' and + 5 V representing logic '1'.

1.11. Conversion of Analog Signals to Digital Signals

In communication systems, sometimes it happens that we are available with analog signal, however, we have to transmit a digital signal for a particular application. In such cases, we have to convert an analog signal into digital signal. This means that we have to canvert a continuous time signal in the form of digits. To see how a signal can be converted from analog to digital form, let us consider an analog signal as shown in figure 1.6(*a*). First of all, we get sample of this signal according to sampling theorem. For this purpose, we mark the time-instants t_0, t_1, t_2 and so on, at equal time-intevals along the time axis. At each of these time-instants, the magnitude of the signal is measured and thus samples of the signal are taken. Figure 1.6(*b*) shows a representation of the signal of figure 1.6(*a*) in terms of its samples.

Now, we can say that the signal in figure 1.6(*b*) is defined only at the sampling instants. This means that it no longer is a continuous function of time, but rather, it is a discrete-time signal. However, since the magnitude of each smaple can take any value in a continuous range, the signal in figure 1.6 (*b*) is still an analog signal.

This difficulty is neatly resolved by a process known as quantization. In quantization, the total, amplitude range which the signal may occupy is divided into a number of standard levels.

As shown in figure 1.6(*c*), amplitudes of the signal $x(t)$ lie in the range $(-m_p, m_p)$ which is partitioned into L intervals, each of magnitude $\Delta v = \dfrac{2m_p}{L}$. Now, each smaple is approximated or rounded off to the nearest quantized level as shown in figure. Since each sample is now approximated to one of the L numbers, therefore the information is digitized.

The quantized signal is an approximation of the original one. We can improve the accuracy of the quantized signal to any desired degree simply by increasing the number of levels L.

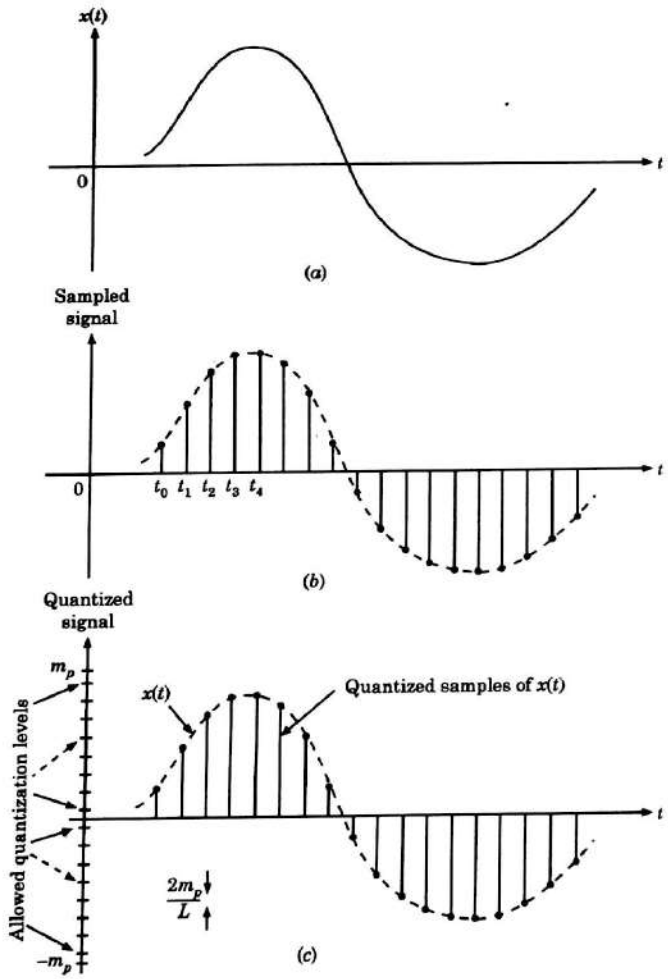


Fig. 1.6. *(a) An analog signal, (b) Samples of analog signal, (c) Quantization*

1.12. Analog and Digital Communication

Depending upon the message signal, communication may also be classified as under:

(*i*) Analog Communication,

(*ii*) Digital Communication.

1.12.1. Analog Communication

Analog communication is that type of communication in which the message or information signal to be transmitted is analog in nature. This means that in analog communication the modulating signal (*i.e.*, baseband signal) is an analog signal. This analog message signal may be obtained from sources such as speech, video shooting etc.

In analog communication, the analog message signal modulates some high carrier frequency inside the transmitter to produce modulated signal. This modulated signal is then transmitted with the help of a transmitting antenna to travel through the transmission channel. At the receiver, this modulated signal is received and processed to recover the original message signal. Figure 1.5 shows the block diagram of an analog communication system.

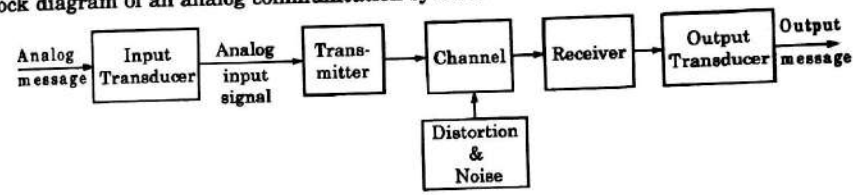


Fig. 1.7. *Basic Analog Communication System.*

Presently all the AM, FM radio transmission and TV transmission are examples of analog communication.

1.12.2. Digital Communicaton

In digital communication, the message signal to be transmitted is digital in nature. This means that digital communication involves the transmission of information in digital form.

1.12.3. Model of a Digital Communication System

Figure 1.8 shows the model of a digital communication system. The overall purpose of the system is to transmit the message or sequences of symbols coming out of a source to a destination point at as high a rate and accuracy as possible. The source and the destination point are physically separated in space and a communication channel connects the source to the destination point. The communication channel accepts electrical (*i.e.*, electromagnetic) signals and the output of the channel is usually a smeared or distorted version of the input due to the non-ideal nature of the communication channel. In addition to this, the information-bearing signal is also corrupted by unpredictable electrical signals (*i.e.*, noise) from both man-made and natural causes. Thus, the smearing and the noise introduce errors in the information being transmitted and limits the rate at which information can be communicated from the source to the destination.

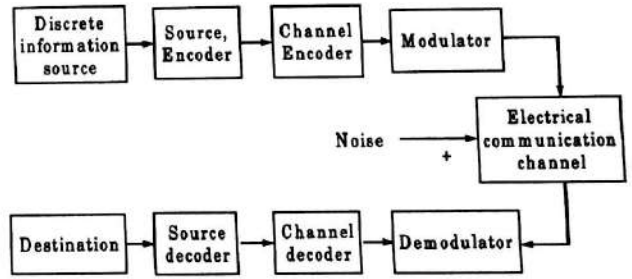


Fig. 1.8. *Model of a Digital Communication System.*

The probability of incorrectly decoding a message symbol at the receiver is often used as a measure of performance of a digital communication system.

Now, let us have a detailed look at each of the functional blocks in a digital communication system.

1.12.3.1. Discrete Information Source

Information source may be classified into two categories based upon the nature of their output *i.e.*, analog information sources and discrete information sources. In case of analog communication, the information source is analog. Analog information sources, such as microphone actuated by speech, emit one or more continuous amplitude signals.

In case of digital communication, the information source produces a message signal which is not continuously varying with time. Rather the message signal is intermittent with respect to time. The output of descrete information sources such as a teletype or the numerical output of a computer consists of a sequence of discrete symbols or letters. An analog information source may be transformed into a discrete information source through the process of sampling and quantizing. Discrete information sources are characterized by the following parameters:

(*i*) Source Alphabet: These are the letters, digits or special characters available from the information source.

(*ii*) Symbol Rate: It is the rate at which the information source generates source alphabets. It is generally repressented in symbols/sec unit.

(*iii*) Source Alphabet Probabilities: Each source alphabet from the source has independent occurrence rate in the sequence. As an example, letters A, E, I etc., occur frequently in the sequence. Hence, probability of the occurrence of each source alphabet can become one of the important property which is useful in digital communication.

(*iv*) Probabilistic Dependence of Symbols in a Sequence: The information carrying capacity of each source alphabet is different in a particular sequence. This parameter defines average information content of the symbols. The entropy of a source describes the average information content per symbol in long messages. Entropy may be defined in terms of bits per symbol. Bit is the abbreviation for a binary digit.

This means that the source information rate is the product of symbol rate and source entropy, *i.e.*,

$$\underset{\text{(Bits/sec.)}}{\text{Information rate}} = \underset{\text{(Symbols/sec.)}}{\text{Symbol rate}} \times \underset{\text{(Bits/Symbol)}}{\text{Source entropy}}$$

Thus, the information rate represents minimum average data rate required to transmit information from source to the destination.

1.12.3.2. Source Encoder and Decoder

The symbols produced by the information source are given to the source encoder. These symbols cannot be transmitted directly. They are first converted into digital form (*i.e.*, binary sequence of 1's and 0's) by the source encoder. Each binary '1' and '0' is known as a bit. The group of bits is called a codeword. The source encoder assigns codewords to the symbols. For each distinct symbol, there is an unique codeword.

The codeword can be of 4, 8, 16 or 32 bits length. As the number of bits are increased in each codeword, the symbols that may be represented are also increased.

As an example, 8 bits would have 2^8 *i.e.*, 256 distinct codewords. This means that 8 bits may be used to represent 256 symbols and similarly 16 bits may represent $2^{16} = 65536$ symbols and so on. Some typically source encoders are pulse code modulators, delta modulators, vector quantizers etc. Source encoders must have following important parameters:

(*i*) Block Size

Block size describes the maximum number of distinct codewords which can be represented by a source encoder. This depends on the number of bits in the codeword. As an example, the block size of 8 bits source encoder will be 2^8 *i.e.*, 256 codewords.

(*ii*) Codeword Length

Codeword length is the number of bits used to represent each codeword. As an example, if 8 bits are assigned to each codeword, then the codeword length will be 8 bits.

(*iii*) Average Data Rate

Average data rate is the output bits per second from the source encoder. In fact, the source encoder assigns multiple number of bits to each input symbol. Hence, the data rate is generally higher than the symbol rate. As an exmaple, if we consider that the symbols are given to the source encoder and the length of codeword is 8 bits, then the output data rate from the source encoder would be given as

$$\text{Data rate} = \text{Symbol rate} \times \text{Codeword length} = 10 \times 8$$

$$\text{Data rate} = 80 \text{ bits/seconds}$$

Also, since the information rate is the minimum number of bits per second needed to convey information from source to destination, therefore the optimum data rate is equal to the information rate. However, due to practical limitations, designing such type of source encoder is quite difficult. Thus, the average data rate is higher than the information rate and hence symbol rate also.

(*iv*) Efficiency of the Encoder

The efficiency of the encoder is the ratio of minimum source information rate to the actual output data rate of the source encoder.

In last, it may be noted that at the receiver end, some sort of decoder is used to perform the reverse operation to that of source encoder. It converts the binary output of the channel decoder into a symbol sequence. Some decoders also use memory to store codewords. The decoders and the encoders can be synchronous or asynchronous.

1.12.3.4. Channel Encoder and Decoder

After converting the message or information signal in the form of binary sequence by the source encoder, the signal is transmitted through the channel. The communication channel adds noise and interference to the signal being transmitted. Hence, errors are introduced in the binary sequence received at the receiver end.

Therefore, the errors are also introduced in the symbols generated from these binary codewords. Thus channel coding is done to avoid these types of errors. In fact, the channel encoder adds some redundant binary bits to the input sequence. Also, these redundant bits are always added with some properly defined logic. As an example, let us consider that the codeword from the source encoder to make it 4-bits long. This fourth bit is added (*i.e.*, 1 or 0) in such a manner that the number of 1's in the encoded word remain even (also known as even parity). Table 1.2 gives the output of a source encoder, the fourth bit depending on the parity and the output of channel encoder.

It may be observed from the table that each codeword at the output of channel encoder contains "even" number of 1's. Now, at the receiver end, if odd number of 1's are detected, then the receiver comes to know that there is an error in the received signal. The channel decoder at the receiver is thus able to reconstruct error free accurate bit sequence and reduce the effects of channel noise and distortion.

TABLE 1.1.

Output of source encoder			*Bit to be added by the channel encoder for an even parity*	*Output of a channel encoder*			
b_3	b_2	b_1	b_0	b_3	b_2	b_1	b_0
1	0	0	0	1	1	0	0
0	1	0	1	0	1	0	1
0	0	0	0	0	0	0	0
1	1	1	1	1	1	1	1
	⋮		⋮		⋮		

This means that the channel encoder and decoder serve to increase the reliability of a received signal. However the extra bits which are added by the channel encoders carry no information, rather, they are used by the channel decoder to detect and correct errors if any.

The coding and decoding operation at the encoder and decoder needs the memory and processing of binary data. However, in the modern time, due to use of microcontrollers and computers, the complexity of the encoders and decoders is much reduced.

A channel encoder must have the following important parameters:

(*i*) The coding rate that depends upon the redundant bits added by the channel encoder.

(*ii*) The coding method used.

(*iii*) Coding efficiency which is the ratio of data rate at the input to the data rate at the output of the encoder.

(*iv*) Error control capabilities.

(*v*) Feasibility of the encoder and decoder.

1.12.3.5. Digital Modulators and Demodulators

In article 1.8, we discussed why modulaton is needed in communication system. Now, if the modulating signal is digital (*i.e.* binary codewords), then digital modulation techniques are used. The carrier signal used by digital modulators is always continuous sinusoidal wave of high frequency. In fact, the digital modulators map the input binary sequence of 1's and 0's to the analog signal waveforms. For example, if one bit at a time is transmitted, then digital modulator signal is $s_1(t)$ to transmit binary '0' and $s_2(t)$ to transmit binary '1' as shown in figure 1.9.

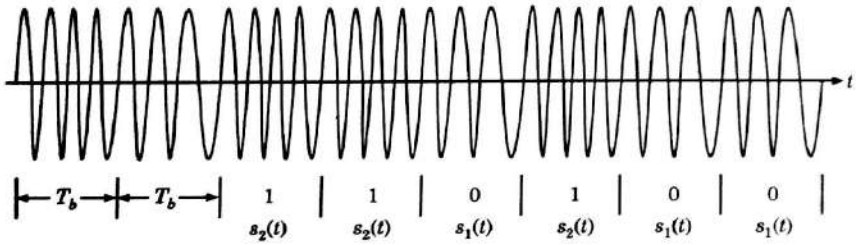


Fig. 1.9. *The output of a digital modulator.*

Here the signal $s_1(t)$ has low frequency compared to signal $s_2(t)$. Hence, here, even though the modulated signal seems to be continuous, the modulation is discrete (*i.e.*, in steps). This means that a signal carrier is converted into two waveforms $s_1(t)$ and $s_2(t)$ because of digital modulation.

Now, if the codeword consits of two bits and they are to be transmitted at a time, then there would be 2^2 *i.e.*, 4 distinct symbols *i.e.*, codewords. Thus, these codewords will require four distinct waveforms for transmission purpose. Such types of modulators are known as M-ary modulators. Amplitude shift keying (ASK), phase shift keying (PSK), frequency shift keying (FSK), differential phase shift keying (DPSK) and minimum shift keying (MSK) are the examples of various digital modulators.

However, since these modulators use a continuous carrier wave, therefore they are also known as digital *CW* modulators.

At the receiver end, the digital demodulator converts the input moodulated signal into the sequence of binary bits.

A digital modulation method must have following important parameters:

(*i*) Bandwidth needed to transmit the signal,

(*ii*) Probability of symbol or bit error,

(*iii*) Synchronous or asynchronous method of detection,

(*iv*) Complexity of implementation.

1.12.3.6. Communications Channel

As discussed earlier, the connection between transmitter and receiver is established through a communication channel. The communication can take place through wirelines, wireless or fiber optic channels. The other media such as optical disks, magnetic tapes and disks etc. may also be called as a communication channel since they can also carry data through them. However, it may be noted that each and every communication channel has some inherent problems. These are:

(*i*) *Signal Attenuation*: The signal attenuation in channel occurs due to the internal resistance of the channel and fading of the signal.

(*ii*) *Amplitude and Phase Distortion*: The transmitted signal is distorted in amplitude and phase due to the non-linear characteristics of the communication channel.

(*iii*) *Additive Noise Interference*: Additive noise interference is produced due to internal solid state devices and resistors etc., used to implement a communication system.

(*iv*) *Multipath Distortion*: The multipath distortion occurs mostly in wireless communication channels.

In fact, the signals coming from different paths tend to interfere with each other.

1.13. Advantages and Disadvantages of Digital Communication

(U.P.T.U., 2003)

In this section, we shall discuss the advantages and disadvantages of digital communication briefly.

Advantages

Following are the advantages of digital communication:

(*i*) The digital communication systems are simpler and cheaper compared to analog communication systems because of the advances made in the IC technologies.

(*ii*) In digital communication, the speech, video and other data may be merged and transmitted over a common channel using multiplexing.

(*iii*) Using data encryption, only permitted receivers may be allowed to detect the transmitted data. This property is of its most importance in military applications.

(*iv*) Since the transmission is digital and the channel encoding is used, therefore the noise does not accumulate from repeater to repeater in long distance communications.

(*v*) Since the transmitted signal is digital in nature, therefore a large amount of noise interference may be tolerated.

(*vi*) Since in digital communication, channel coding is used, therefore the errors may be detected and corrected in the receivers.

(*vii*) Digital communication is adaptive to other advanced branches of data processing such as digital signal processing, image processing and data compression etc.

Disadvantages

Although digital communication offers, so many advantages as discussed above, it has some drawbacks also. However, the advantages of digital communication outweight disadvantages.

The disadvantages are as under:

(*i*) Due to analog to digital conversion, the data rate becomes high. Therefore more transmission bandwidth is required for digital communication.

(*ii*) Digital communication needs synchronization in case of synchronous modulation.

1.14. Fundamental Limitations of Communication System

While designing a communication system, an engineer generally faces several constraints or limitations. These are noise limitation, bandwidth limitation and equipment limitation.

1.14.1. Noise Limitation

The noise may be defined as an unwanted form of energy which tend to interfere with the transmission and reception of the desired signals in a communication system. The noise cannot be eliminated completely. However, the effect of noise on desired signals can be minimised with the help of several techniques. Noise can be classified into two broad categories depending upon its source: external noise and internal noise. External noise is that type of noise whose sources are external to a communication system. The examples of external noise are atmospheric noise, galactic noise and industrial noise. On the other hand, internal noise is that type of noise whose sources are internal to a communication system. The examples of internal noise are thermal noise and shot noise. This type of noise is inevitable in every communication system and forms a basic limitation on transmission and reception of signals.

Infact, the noise limits our ability to identify the intended or desired message correctly and thus limits information transmission. Typical noise variations are measured in microvolts. If the signal variations are quite large in comparison, then the effect of noise is very small and may be ignored. Infact, in many practical communication systems, the effect of noise is generally unnoticeable. However, in long-distance communication systems operating with a limited amount of signal power, the signal may be as small as noise or even smaller than the noise. Thus, in such cases, the presence of noise severally limits the capabilities of a communication system.

1.14.2. Bandwidth Limitation

In addition to noise, the bandwidth limitation is another major constraint in a communication system. The frequency range or the band of frequency needed for a particular given transmission is known as **bandwidth**. This band of frequencies required for a particular transmission is also called *channel*. This band of frequencies or bandwidth for a particular transmission is always allocated by some International regulatory agencies. This type of regulation is essential to avoid interference among the signals having same frequency. But, for a given transmission, this allocated bandwidth may not be sufficient to convey the entire information.

Infact, the information theory states that the greater is the transmission bandwidth of a communication system, the more is the information that can be transmitted. Let us consider a simple example. Suppose one is listening to a music in an AM radio. The complete amount of information available to the human ear is contained in a frequency range upto 15 kHz, *i.e.*, musical information extends upto a frequency of 15 kHz. However, in AM radio the maximum modulating frequency is restricted upto 5 kHz and hence the maximum bandwidth of AM transmission is 10 kHz.

Therefore, an AM radio receiver cannot reproduce all the information contained in the music because this will require a bandwidth of 30 kHz. On the other hand, the bandwidth allocated to a FM transmission is about 200 kHz. Thus on FM receiver can easily reproduce the transmitted information without any distortion. This means that a FM system has a better fidelity than an AM system. Also, it may be observed in common life that one prefers to listen to a FM radio than an AM radio. Thus, we can conclude that bandwidth is a major fundamental limitation of a communication system.

1.14.3. Equipment Limitation

The noise and bandwidth limitations dictate theoretically what can or cannot be achieved in terms of performance in a communication system. However, this theoretical limit may not be realised in a practical system due to equipment limitations. For example, theory might require a bandpass filter with a quality factor of 100 at a centre frequency of 1 kHz. Such a filter cannot be realised in practice. Even if a filter with nearly identical characteristics is built, the cost may exceed what the user of the communication system wants to pay. Thus equipment limitation is another major problem in a communication system.

1.15. Historical Perspective
The Evolution of Modern Electronic Communication

The electronic communication has undergone through a chain of dramatic developments even since the first electronic communication invented by Morse in 1838. However, it may be noted that this development became possible due to the parallel development in the field of electronic devices, circuits etc. To provide historical perspective, the important developments in the field of electronic communication are listed ahead:

TABLE 1.2. A Chronology of Electronic Communication

Year	*Development*
1800-1837	Preliminary developments
1838	The invention of telegraphy
1845	The Kirchhoff's circuit laws enunciated
1864	Maxwell's equations predict electromagnetic radiation
1876	The invention of telephony
1887	The invention of wireless telegraphy
1904	Communication Electronics
1920	Transmission theory
1923	Invention of television
1927	Federal communications commission established
1931	Teletypewriter service initiated
1934	H.S. Black develops the negative-feedback amplifier
1936	Armstrong's paper states the case for FM radio
1937	Pulse code modulation
1938	Radar and microwave communication
1948	Information theory and coding
1950	Time-division multiplexing applied to telephony
1953	Colour TV Standards established in United States
1956	First transoceanic states telephone cable (36 voice channels)
1958	Long-distance data transmission system developed for military purposes
1960	Maiman demonstrated the first Laser
1961	Integrated circuits go into commercial production
1962	Satellite communication begins with Telestar I
1962-1966	High-speed digital communication
1963	Solid-state microwave oscillators
1964	Full electronic telephone switching system goes into service
1965	Mariner IV transmits pictures from Mars to earth
1966-1975	Wideband communication systems
1979	Intercity optical links
1988	Under sea fiber optical link

SUMMARY

1. Communication is the process of establishing connection or link between two points for information exchange.
2. The electronic equipments which are used for communication purpose, are called communication equipments. Different communication equipments when assembled together form a communication system.
3. Typical examples of communication system are line telephony and line telegraphy, radio telephony and line telegraphy, radio broadcasting, point-to-point communication and mobile communication, computer communication, radar communication, television broadcasting, radio telemetry, radio aids to navigation, radio aids to aircraft landing etc.
4. In the most fundamental sense, communication involves the transmission of information from one point to another through a succession of process.
5. That the function of information source is to produce required message which has to be transmitted.
6. A transducer is a device which converts one form of energy into another form. The message from the information source may or may not be electrical in nature. In a case when the message produced by the information source is not electrical in nature, an input transducer is used to convert it into a time-varying electrical signal.
7. The function of the transmitter is to process the electrical signal from different aspects. For example in radio broadcasting the electrical signal obtained from sound signal, is processed to restrict its range of audio frequencies (upto 5 kHz in amplitude modulation radio broadcast) and is often amplified.
8. That the function of the channel is to provide a physical connection between the transmitter and the receiver.
9. There are two types of channels, namely point-to-point channels and broadcast channels. Examples of point-to-point channels are wirelines, microwave links and optical fibres. Wirelines operate by guided electromagnetic waves and they are used for local telephone transmission. In case of microwave links, the transmitted signal is radiated as an electromagnetic wave in free space.
10. Noise is an unwanted signal which tend to interfere with the required signal. Noise signal is always random in character. Noise may interfere with signal at any point in a communication system. However, the noise has its greatest effect on the signal in the channel.
11. The main function of the receiver is to reproduce the message signal in electrical form from the distorted received signal. This reproduction of the original signal is accomplished by a process known as the demodulation or detection. Demodulation is the reverse process of modulation carried out in transmitter.
12. Destination is the final stage which is used to convert an electrical message signal into its original form. For example in radio broadcasting, the destination is a loudspeaker which works as a transducer *i.e.* it converts the electrical signal in the form of original sound signal.
13. The telecommunication environment is dominated by the following four important sources of information:
    (*i*) speech
    (*ii*) television
    (*iii*) facsimile, and
    (*iv*) personal computers
14. Several sources of information which are responsible for the generation of message signals. These message signals need to be transmitted to the destination point. Now, let us discuss various communication channel that carryout the transmission.
15. Basically, we shall consider following channels:
    (*i*) telephone channels
    (*ii*) optical fibers
    (*iii*) mobile radio channels, and
    (*iv*) satellite channels.

16. A telephone network makes use of a switching mechanism. This switching mechanism is known as circuit switching and it is used to establish an end-to-end communication link on a temporary basis.
17. The second communication channel, an optical fiber, is a dielectric waveguide which transports light signals from one place to another just as a metallic wire pair or a co-coxial cable, transports electrical signals.
18. An optical fiber consists of a central core within which the propagating electromagnetic field is confined and which is surrounded by a cladding layer, which is itself surrounded by a thin protective jacket.
19. The third communication channel, mobile radio channel, extends the capability of the public telecommunications network by introducing mobility into the network by virtue of its ability to broadcast.
20. A satellite channel provides broad-area coverage in a continental as well as intercontinental sense. Moreover, access to remove areas not covered by conventional cable or fiber communications is also a distinct feature of satellites.
21. Communications satellites in geostationary orbit offer the following unique system capabilities:
    (*a*) broad-area coverage
    (*b*) reliable transmission links
    (*c*) wide transmission bandwidth.
22. The message signal generated from the information source is known as baseband signal. This baseband signal may be a combination of two or more message signals. If the baseband signal is transmitted directly, then it is known as baseband transmission. This means that the baseband transmission does not used modulators and demodulators. The baseband signal may be both analog as well as digital.
23. The analog baseband signal varies continuously with time and has continuous amplitude. The digital baseband signal is discrete in both time and amplitude.
    The baseband transmission is preferred at low frequencies and for short distances. However, Inter Symbol Interference (ISI) is the major problem associated with the baseband transmission.
24. Regarding the mode of propagation, communication may be divided in the following two forms:
    (*i*) Line Communication
    (*ii*) Wireless or Radio Communication
25. In line communication, the medium of transmission is a pair of conductors called transmission line. This is also called as line channel.
26. In wireless or radio communication, a message is transmitted through open space by electromagnetic waves called as radio waves. Radio waves are radiated from the transmitter in open space through a device called antenna.
27. Modulation may be defined as the process by which some characteristic of a signal called carrier is varied in accordance with the instantaneous value of another signal called modulating signal. Signals containing information or intelligence are referred as modulating signals. This information bearing signal is also called baseband signal. The carrier frequency is greater than the modulating frequency. The signal resulting from the process of modulation is called modulated signal.
28. When the carrier wave is continuous in nature, the modulation process is known as continuous wave (CW) modulation or analog modulation. Examples of continuous wave modulation are Amplitude Modulation and Angle Modulation.
29. When the carrier wave is a pulse-type waveform, the modulation process is known as pulse modulation.
30. In a communication system, there are following two primary communication resources to be employed:
    (*i*) transmitted power, and
    (*ii*) channel bandwidth
31. The transmitted power refers to the average power of the transmitted signal.
    The channel bandwidth is defined as the band of frequencies allocated for the transmission of the message signal.
32. In communication systems, sometimes, it happens that we are available with analog signal, however we have to transmit a digital signal for a particular application. In such cases, we have to convert an analog signal into digital signal.

33. Depending upon the message signal, communication may also be classified as under:
(*i*) Analog Communication
(*ii*) Digital Communication

34. Analog communication is that type of communication in which the message or information signal to be transmitted is analog in nature. This means that in analog communication the modulating signal (*i.e.*, baseband signal) is an analog signal. This analog message signal may be obtained from sources such as speech, video shooting etc.

35. In digital communication, the message signal to be transmitted is digital in nature. This means that digital communication involves the transmission of information in digital form.

36. Following are the advantages of digital communication:
(*i*) The digital communication systems are simpler and cheaper compared to analog communication systems because of the advances made in the IC technologies.
(*ii*) In digital communication, the speech, video and other data may be merged and transmitted over a common channel using multiplexing.
(*iii*) Using data encryption, only permitted receivers may be allowed to detect the transmitted data. This property is of its most importance in military applications.
(*iv*) Since the transmission is digital and the channel encoding is used, therefore the noise does not accumulate from repeater to repeater in long distance communications.
(*v*) Since the transmitted signal is digital in nature, therefore a large amount of noise interference may be tolerated.
(*vi*) Since in digital communication, channel coding is used, therefore the errors may be detected and corrected in the receivers.
(*vii*) Digital communication is adaptive to other advanced branches of data processing such as digital signal processing, image processing and data compression etc.

REVIEW QUESTIONS

1. What are the basic constituents of a communication system?
2. Draw the block diagram of a communication system and explain the function of each block.
3. What is meant by the term 'Channel' as applied to a communication system.
4. Why is a high-frequency carrier needed in a communication system?
5. Explain the need for modulation in a communication system?
6. Explain the radio frequency spectrum used in communication system.
7. Explain the fundamental limitations of a communication system.
8. What are the basic constituents of a communication system?
9. Draw the block diagram of a communication system and explain the function of each block.
10. What is meant by the term 'Channel' as applied to a communication system?
11. Why is a high-frequency carrier needed in a communication system?
12. Explain the need for modulation in a communication system?
13. Explain the radio frequency spectrum used in communication system.
14. Explain the difference between Analog and Digital communications.
15. How will you convert an analog signal into a digital signal.
16. Write the advantages and disadvantages of a digital communication system.
17. Explain the difference between baseband transmission and bandpass transmission.
18. Draw the block diagram of a digital communication system and explain the function of each block.

OBJECTIVE TYPE QUESTIONS

Find the correct choice:

1. The highest modulating frequency used in AM broadcast system is
(*a*) 10 kHz (*b*) 15 kHz
(*c*) 5 kHz (*d*) 2 MHz
2. The high frequency (HF) range extends from
(*a*) 300-3000 kHz (*b*) 3-30 MHz
(*c*) 30-300 MHz (*d*) 300-3000 MHz
3. The very high frequency (VHF) range extends from
(*a*) 3-30 MHz (*b*) 30-300 MHz
(*c*) 300-3000 MHz (*d*) 3000-30,000 MHz
4. The ultra high frequency (UHF) range extends from
(*a*) 3-30 MHz (*b*) 30-300 MHz
(*c*) 300-3000 MHz (*d*) 3000-30,000 MHz
5. The function of the input transducer in a communication system is
(*a*) to transmit the message signal (*b*) to modulate the message signal
(*c*) to convert message sound signal into electrical signal
(*d*) none of the above
6. Communication is the process of
(*a*) keeping in touch (*b*) broadcasting
(*c*) exchanging information (*d*) entertainment by electronics
7. Two key barriers to human communication are
(*a*) distance (*b*) cost
(*c*) ignorance (*d*) language
8. Electronics communications was discovered in which century?
(*a*) sixteenth (*b*) eighteenth
(*c*) nineteenth (*d*) twentieth
9. Which of the following is not a major communications medium?
(*a*) free space (*b*) water
(*c*) wires (*d*) fiber-optic cable
10. The communications medium causes the signal to be
(*a*) amplified (*b*) modulated
(*c*) attenuated (*d*) interferred with
11. The process of transmitting two or more information signals simultaneously over the same channel is called
(*a*) multiplexing (*b*) telemetry
(*c*) detection (*d*) modulation
12. Recovering information from a carrier is known as
(*a*) demultiplexing (*b*) modulation
(*c*) detection (*d*) carrier recovery

ANSWERS

1. (*c*)	2. (*b*)	3. (*b*)	4. (*c*)	5. (*c*)	6. (*c*)	7. (*a* & *b*)
8. (*c*)	9. (*a*)	10. (*c*)	11. (*a*)	12. (*c*)		

□□□

@sTc

CHAPTER 2

Signal Analysis and Transmission

Inside this Chapter

2.1. Introduction 28
2.2. Signal and Signal Classification 29
2.3. Singularity Functions 32
2.4. Representation of Signals 34
2.5. Fourier series 35
2.6. Trigonometric Fourier Series 35
2.7. Properties of Continuous-time Fourier Series (CTFS) 47
2.8. Analysis of a Non-periodic Function over Entire Interval, the Fourier Transform 49
2.9. Properties of Fourier Transform 64
2.10. Fourier Transform of Periodic Functions 70
2.11. Convolution 70
2.12. Energy Signal 72
2.13. Power Signals 79
2.14. Cross-Correlation 82
2.15. Autocorrelation Functions 85
2.16. Application of Fourier Transform to the Analysis of LTI Networks 89

2.1. Introduction

As discussed earlier in Chapter 1, the purpose of a communication system is to transmit a message signal from an information source, located at one point, to a destination or user located at another point some distance away. The signal produced by information source may not be electrical in nature. However, for transmission purpose, we need an electrical signal because all the signal processing devices or equipments are electrical *i.e.*, they process electrical signals only. Therefore information signal is first converted into electrical form and then transmitted.

The present chapter is devoted to the signal classification and signal analysis. In the signal analysis, we shall consider that the signal is already in electrical form.

A signal is generally given in time-domain. Time-domain specifies that what is the amplitude of the signal at a particular time-instant.

But if we want to know the frequency components present in this signal, then we will have to take help of some types of mathematical tools.

Fourier series and Fourier transform are such mathematical tools which are used to analyse a signal *i.e.* to get frequency-domain information. This means that with the help of Fourier series and Fourier transform a time-domain signal is converted into frequency-domain signal.

In this chapter, we shall study Fourier series and Fourier transform and their properties.

2.2. Signal and Signal Classification

A signal may be a function of time, tempearature, position, pressure, distance etc. Some signals in our daily life are music, speech, picture and video signals. Systematically, we can define a signal as under:

"A function of one or more independent variables which contains some information is called a signal."

In electrical sense, the signal can be voltage or current. The voltage or current is the function of time as an independent variable.

In daily life, we come across several electric signals such as radio signal, TV signal, computer signal etc. Signals may be classified as under:

(*i*) Continuous-time and Discrete-time signals.
(*ii*) Real and complex signals.
(*iii*) Deterministic and random signals.
(*iv*) Periodic and non-periodic signals.
(*v*) Even and odd signals.
(*vi*) Energy and power signals.
(*vii*) Analog and digital signals.

2.2.1. Continuous-time and Discrete-time Signals

A signal $x(t)$ is a continuous-time signal if t is a continuous variable. This means that a continuous-time signal is defined continuously in the time-domain.

On the other hand, if time t is a discrete variable, *i.e.* $x(t)$ is defined at discrete times, then $x(t)$ is a discrete-time signal. Since a discrete-time signal is defined at discrete times, it is often identified as a sequence of numbers and is denoted by $x[n]$, where n = integer.

Figure 2.1 shows the graph of continuous-time and discrete-time signal.

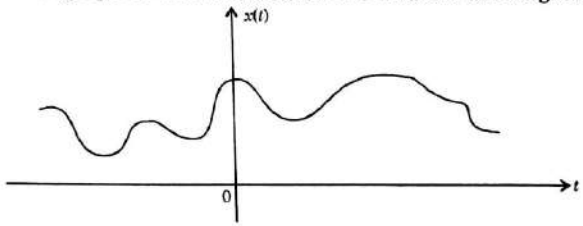


Fig. 2.1. *(a) A continuous-time signal*

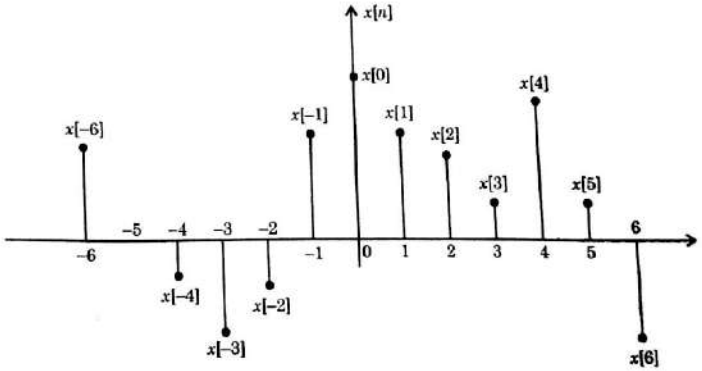


Fig. 2.1. *(b) A discrete-time signal*

2.2.2. Real and Complex Signals

A signal $x(t)$ is a real signal if its value is a real number. Similarly, a signal $x(t)$ is a complex signal if its value is a complex number.

2.2.3. Deterministic and Random Signals

Deterministic signals are those signals which can be completely specified in time. The pattern of this type of signal is regular and can be characterized mathematically. In addition to this, the nature and amplitude of such a signal at any time can be predicted.

Few examples of deterministic signals are:

(i) $x(t) = bt$

This is a ramp signal whose amplitude increases linearly with time and the slope is b.

(ii) $x(t) = a \sin \omega t$

For this signal, the amplitude varies sinusoidally with time and its maximum ampliture is A.

On the other hand, a random signal is one whose occurrence is always random in nature. The pattern of such a signal is quite irregular. Random signals are also called non-deterministic signals.

A typical example of a random signal is thermal noise, generated in an electric circuit. Such a noise signal has probabilistic behaviour.

2.2.4. Periodic and Aperiodic Signals

A periodic signal is that type of signal which has a definite pattern and repeats over and over with a repetition period of T. In other words, a signal is called periodic if it exhibits periodicity as follow:

$$x(t + T) = x(t), \quad -\infty < t < \infty$$

T is the period of the signal. The smallest value of period T which satisfies above equation is called the **fundamental period** T_0. Figure 2.2 shows a periodic signal.

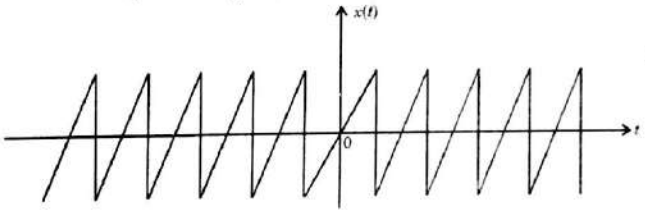


Fig. 2.2. *A periodic signal.*

A signal is said to be Aperiodic if it does not repeat. Sometimes aperiodic signals are said to have a period equal to infinity. Figure 2.3 shows an aperiodic signal. This is a decaying exponential pulse given by the equation:

$$x(t) = e^{-at}$$

Now, $\quad x(t + T_0) = e^{-a(t + T_0)} = e^{-a(t + \infty)}$

or $\quad x(t + T_0) = e^{-at}.e^{-\infty} = e^{-at}.0$

$$x(t + T_0) = 0$$

which is not equal to $x(t)$.

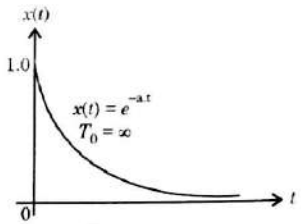


Fig. 2.3. *An aperiodic signal.*

This shows that the signal having period $T_0 = \infty$ is nothing but aperiodic signal.

Example 2.1. Determine whether the following signals are periodic or not:

(i) $x(t) = \sin 15\pi t$

(ii) $x(t) = \sin \sqrt{2}\, \pi t$

Solution: (i) $x(t) = \sin 15\, \pi t$ is a periodic signals.

The fundamental period T is given as

$$T = \frac{2\pi}{w} = \frac{2\pi}{15\pi} \qquad [\because \omega = 15\,\pi]$$

or $\quad T = \mathbf{0.133}$ seconds **Ans.**

(ii) $\quad x(t) = \sin \sqrt{2}\, \pi t$ is a periodic signal.

The fundamental period T is given as

$$T = \frac{2\pi}{w} = \frac{2\pi}{\sqrt{2}\,\pi} \qquad [\because \omega = \sqrt{2}\,\pi]$$

or $\quad T = \dfrac{2}{\sqrt{2}} = \sqrt{2} = \mathbf{1.41}$ seconds **Ans.**

Example 2.2. Assume $x_1(t)$ and $x_2(t)$ are periodic signals with periods T_1 and T_2 respectively. Under what conditions is the sum

$$x(t) = x_1(t) + x_2(t)$$

periodic and what will be the period of $x(t)$ if it is periodic?

Solution: Given that $x_1(t)$ and $x_2(t)$ are periodic with periods T_1 and T_2 respectively. Thus, $x_1(t)$ and $x_2(t)$ may be written as

$$x_1(t) = x_1(t + T_1) = x_1(t + mT_1) \qquad \text{where } m \text{ is an integer.}$$

and $\quad x_2(t) = x_2(t + T_2) = x_2(t + nT_2) \qquad$ where n is an integer

Now, if T_1 and T_2 are such that

$$mT_1 = nT_2 = T$$

Then $\quad x(t + T) = x_1(t + T) + x_2(t + T)$

$$x(t + T) = x_1(t) + x_2(t) = x(t)$$

i.e. $x(t)$ is periodic in this case.

Therefore, the condition for $x(t)$ to be periodic is

$$\frac{T_1}{T_2} = \frac{n}{m} = \text{a rational number.}$$

From above analysis, it is clear that the smallest common period is the least common multiple of T_1 and T_2 and is given by the equation

$$mT_1 = nT_2 = T$$

if the integers m and n are relatively prime. On the other hand, if the ratio T_1/T_2 is an irrational number, then the signals $x_1(t)$ and $x_2(t)$ will not have a common period and so $x(t)$ will not be periodic.

2.2.5. Even and Odd Signals

An even signal is that type of signal which exhibits symmetry in the time domain. This type of signal is identical about the origin. Mathematically, an even signal must satisfy the following condition:

$$x(t) = x(-t)$$

Figure 2.4 shows an even signals.

Similarly, an odd signal is that type of signal which exhibits anti-symmetry. This type of signal is not identical about the origin. Actually, the signal is identical to its negative. Mathematically, an odd signal must satisfy the following condition.

Figure 2.5 shows an odd signals.

@sTc(small team creation)

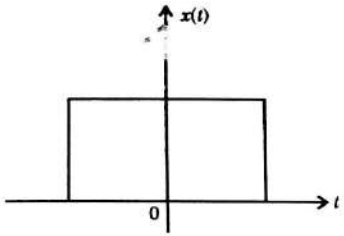


Fig. 2.4. *Even signal.*

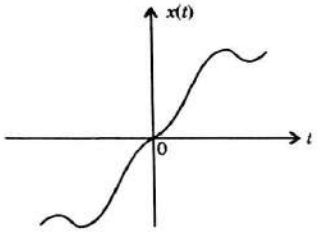


Fig. 2.5. *Odd signals.*

2.2.6. Energy and Power Signals

Signals may also be classified as energy and power signals. However, there are some signals which can neither be classified as energy signals nor power signals.

The energy signal is one which has finite energy and zero average power.

Hence, $x(t)$ is an energy signal, if:

$$0 < E < \infty \text{ and } P = 0$$

where E is the energy and P is the power of the signal $x(t)$.

The power signal, is one which has finite average power and infinite energy.

Hence, $x(t)$ is a power signal, if:

$$0 < P < \infty \text{ and } E = \infty$$

However, if the signal does not satisfy any of the above two conditions, then it is neither an energy signal nor a power signal.

2.3. Singularity Functions

(Anna University, Chennai 2000)

There are several singularity functions which play vital role in the study of communication system. These singularity functions serve as basic building blocks for the construction of more complex signals. Infact, these singularity functions may be used to model a large number of physical signals which occur in nature. These singularity functions may be listed as ahead:

1. Unit step function,
2. Unit impulse function,
3. Ramp function.

Now, we shall discuss all these elementary signals one-by-one.

2.3.1. Unit Step Function

The unit step function is that type of singularity function which exists only for positive side and is zero for negative side. Also the unit step function is discontinuous at $t = 0$.

The unit-step function is denoted by $u(t)$ and may be expressed as

$$u(t) = \begin{cases} 0 & \text{for } t < 0 \\ 1 & \text{for } t > 0 \end{cases} \qquad \text{...(2.1)}$$

Figure 2.6 shows the unit-step function $u(t)$.

The usefulness of the unit-step function lies in the fact that if we want a signal to start at $t = 0$, so that it may have a value of zero for $t < 0$, we only need to multiply the given signal with unit-step function $u(t)$.

For example, the signal e^{-at} represents an exponential signal which starts at $t = -\infty$. But if want this signal to start

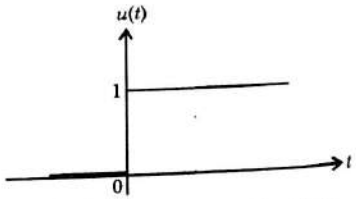
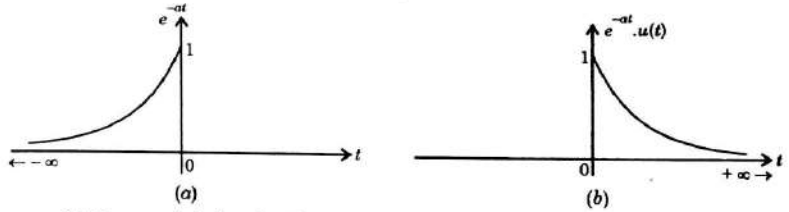


Fig. 2.6. *An unit-step function.*

at $t = 0$, it may be described as $e^{-at} \cdot u(t)$.

Figure 2.7 shows the signals e^{-at} and $e^{-at} \cdot u(t)$.



(a) Exponential signal e^{-at} *(b) Exponential signal $e^{-at}.u(t)$.*

Fig. 2.7.

2.3.2. Unit Impulse Function

The unit impulse function is the most widely used elementary function used in the analysis of communication system.

An unit-impulse function is denoted as $\delta(t)$ and may be expressed as

$$\delta(t) = 0 \qquad t \neq 0 \qquad \text{...(2.2)}$$

and

$$\int_{-\infty}^{\infty} \delta(t)\, dt = 1 \qquad \text{...(2.3)}$$

The unit-impulse function is also known as Dirac-delta function.

Figure 2.8 shows the graphical representation of an unit impulse function.

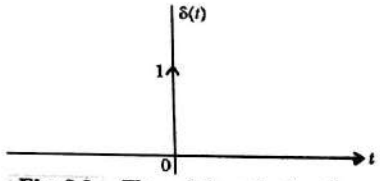


Fig. 2.8. *The unit-impulse function.*

If unit-impulses function is assumed in the form of a pulse, then following points may be observed about an unit-impulse function:

(*i*) the width of the pulse is zero. This means that the pulse exists only at $t = 0$.

(*ii*) The height of the pulse goes to infinity.

(*iii*) The area under the pulse-curve is always unity.

(*iv*) The upward pointing arrow indicates the actual value of infinity.

(*v*) The height of arrow indicates the total area under the impulse. As in figure 2.8, the height of arrow is 1, therefore the total area under the impulse is unity.

The total area under the pulse is also called the **strength** of the impulse.

Figure 2.9 shows the unit impulse function centred at $t = t_0$. It is denoted by $\delta(t - t_0)$.

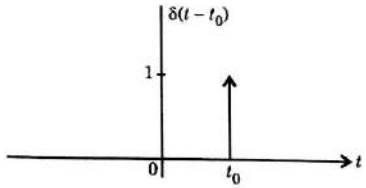


Fig. 2.9. *Unit-impulse function centered at $t = t_0$.*

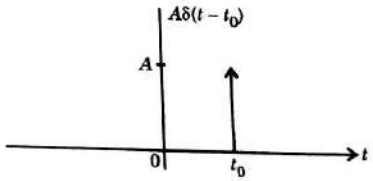


Fig. 2.10. *Impulse-function centered at $t = t_0$ and having area A..*

Figure 2.10 shows the impulse function centred at $t = t_0$ and having area equal to A.

Importance

The unit impulse function is generally used to represent point mass, point charge, point source and other physical entities used in science and engineering.

2.3.3. Unit Ramp Function

A continuous-time unit-ramp function is that type of function which starts at $t = 0$ and increases linearly with time t. It is denoted by $r(t)$.

Mathematically,

$$r(t) = \begin{cases} 0 \text{ for } t < 0 \\ t \text{ for } t < 0 \end{cases} \quad ...(2.4)$$

Figure 2.11 shows the graphical representation of an unit-ramp function.

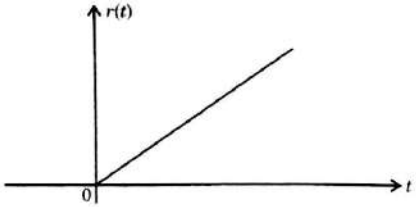


Fig. 2.11. *Continuous-time unit-ramp function.*

2.4. Representation of Signals

Before we discuss continuous-time Fourier series (CTFS) and continuous-time Fourier transform (CTFT), let us discuss how we can represent an electrical signal. Signals contain information about a variety of things and activities in physical world. As a matter of fact, electrical signal may be represented in two equivalent forms, *i.e.*, as a voltage signal or a current signal. This means that an electrical signal may be represented either in the form of a voltage source or in the form of a current source as shown in figure 2.12. These two forms are known as Thevenin's form and Norton's form of signal representation respectively.

Now, an electrical signal, either a voltage signal or a current signal, may further be represented in two forms:

These two types of representation are as under:

(*i*) Time domain representation

(*ii*) Frequency domain representation

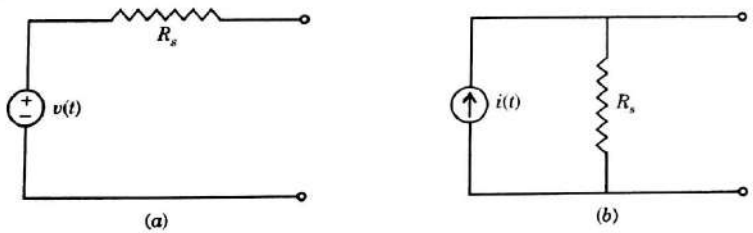


Fig. 2.12. *Representation of signals (a) Thevenin's form (b) Norton's form.*

(*i*) ***Time domain representation.*** In time-domain representation, a signal is a time-varying quantity.

(*ii*) ***Frequency domain representation.*** In frequency domain, a signal is represented by its frequency spectrum. To obtain frequency spectrum of a signal, Fourier series and Fourier transform are used.

Fourier series is used to get frequency spectrum of a time-domain signal, when the signal is a periodic function of time. With the help of Fourier series, a given periodic function of time may be expressed as the sum of an infinite number of sinusoids whose

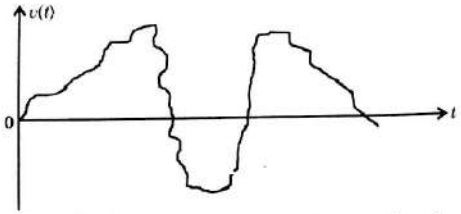


Fig. 2.13. *An arbitary time-domain signal.*

frequency are harmonically related.

Figure 2.13 shows an arbitary time-domain signal.

Also, figure 2.14 shows the time-domain representation of a square wave.

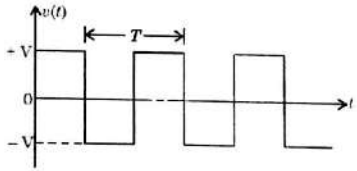


Fig. 2.14. *A symmetrical square-wave signal (time-domain representations).*

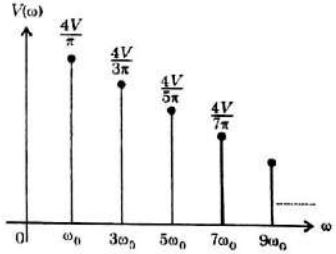


Fig. 2.15. *Frequency spectrum of square wave signal (Frequency domain representations).*

When Fourier series is applied, we get

$$v(t) = \frac{4v}{\pi}\left(\sin \omega_0 t + \frac{1}{3}\sin 3\omega_0 t + \frac{1}{5}\sin 5\omega_0 t + ...\right) \quad ...(2.5)$$

From above, it is clear that the frequency spectrum of a periodic signal is discrete.

Fourier transform is used to get frequency spectrum of a time-domain signal when signal waveform is an arbitary function of time or signal waveform is a nonperiodic function of time. Also the Fourier transform provides a continuous frequency spectrum of an arbitary time domain signal waveform.

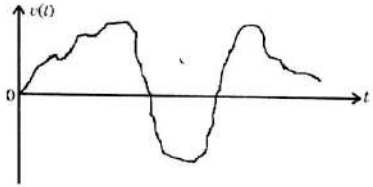


Fig. 2.16. *Time domain representation of an arbitary signal waveform.*

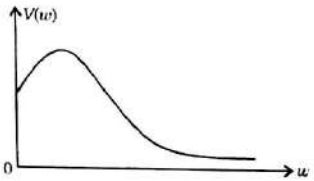


Fig. 2.17. *Frequency domain representation of time-domain signal.*

2.5. Fourier Series

The continuous-time Fourier series exists only when the function $x(t)$ satisfies the following three conditions called **Dirichlet's conditions.**

(*i*) $x(t)$ is well-defined and a single-valued function.

(*ii*) $x(t)$ must possess only a finite number of discontinuities in the period T.

(*iii*) $x(t)$ must have a finite number of positive and negative maxima in the period T.

2.6. Trigonometric Fourier Series

A periodic function $x(t)$ may be expressed in the form of trigonometric Fourier series comprising the following sine and cosine terms:

$$x(t) = a_0 + a_1 \cos \omega_0 t + a_2 \cos_2 \omega_0 t + \ ... + a_n \cos n\omega_0 t + ... + b_1 \sin \omega_0 t + b_2 \sin_2 \omega_0 t + ... + b_n \sin n\omega_0 t + ... \quad ...(2.6)$$

$$x(t) = a_0 + \sum_{n=1}^{\infty}(a_n \cos\omega_0 t + b_n \sin n\omega_0 t) \quad (t_0 \le t \le t_0 + T) \qquad ...(2.7)$$

Here $T = \dfrac{2\pi}{\omega_0}$ and a_n and b_n are the coefficients.

w_0 is called the fundamental frequency and $2\,\omega_0$, $3\,\omega_0$, $4\,\omega_0$ ... are called the harmonics of ω_0.

Equation (2.3) is the trigonometric Fourier series representation of function $x(t)$ over an interval $(t_0, t_0 + T)$, *i.e.*, over one period.

Now, the constant a_0 corresponds to $n = 0$.

It may be given by the expression,

$$a_0 = \frac{1}{T}\int_{t_0}^{t_0+T} x(t)\,dt \qquad ...(2.8)$$

From equation (2.8) it is clear that a_0 is the average value or d.c. component of $x(t)$ over the interval $(t_0, t_0 + T)$.

However, the constant $b_0 = 0$ because $\sin n\omega_0 t$ is zero for $n = 0$.

The constant a_n is given by

$$a_n = \frac{2}{T}\int_{t_0}^{t_0+T} x(t)\cos n\omega_0 t\,dt \qquad ...(2.9)$$

or equivalently

$$a_n = \frac{2}{T}\int_{0}^{T} x(t)\cos n\omega_0 t\,dt \qquad ...(2.10)$$

The constant b_n is given by

$$b_n = \frac{2}{T}\int_{t_0}^{t_0+T} x(t)\sin n\omega_0 t\,dt \qquad ...(2.11)$$

or equivalently

$$b_n = \frac{2}{T}\int_{0}^{T} x(t)\sin n\omega_0 t\,dt \qquad ...(2.12)$$

2.6.1. Fourier Series Representation of a Periodic Function over the Entire Interval ($-\infty < t < \infty$)

In last article we considered Fourier series representations of a periodic function over one period $(t_0, t_0 + T)$. A periodic function repeats after a particular period T. Hence, Fourier series for a periodic signal is same for entire interval $(-\infty, \infty)$ as for the interval $(t_0, t_0 + T)$.

Hence for a periodic function $x(t)$, we have

$$x(t) = a + \sum_{n=1}^{\infty}(a_n \cos n\omega_0 t + b_n \sin n\omega_0 t) \quad (-\infty < t < \infty) \qquad ...(2.13)$$

$$a_0 = \frac{1}{T}\int_{t_0}^{t_0+T} x(t)\,dt = \frac{1}{T}\int x(t)\,dt \qquad ...(2.14)$$

$$a_n = \frac{2}{T}\int_{t_0}^{t_0+T} x(t)\cos n\omega_0 t\,dt$$

$$a_n = \frac{2}{T}\int x(t)\cos n\omega_0 t\,dt \qquad ...(2.15)$$

and

$$b_n = \frac{2}{T}\int x(t)\sin n\omega_0 t\,dt \qquad ...(2.16)$$

2.6.2. Symmetry Conditions (Few Points)

(*i*) If the function $x(t)$ is even, then $x(-t) = x(t)$, $\cos t$, t^2, $t \sin t$ are examples of even function.

Figure 2.18 shows the even functions of t. The graph of an even function is symmetrical with respect to the y-axis and in Fourier series representation of an even function, only cosine terms are present (d.c. term is optional). Mathematically, if $x(t)$ is even

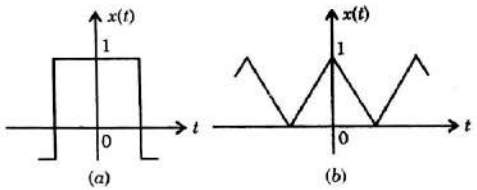


Fig. 2.18. *Graphs of even functions.*

$$\int_{-a}^{a} x(t)\,dt = 2\int_{0}^{a} x(t)\,dt \qquad ...(2.17)$$

(*ii*) If the function $x(t)$ is odd, then $x(-t) = -x(t)$ $\sin t$, t^3, $t \cos t$ are examples of odd function. Figure 2.19 shows the odd functions of t. The graph of an odd function is symmetrical about the origin and in Fourier series representation of an odd function, only sine terms are present (d.c. term is optional). Mathematically, if $x(t)$ is odd,

then, $\displaystyle\int_{-a}^{a} x(t)\,dt = 0.$

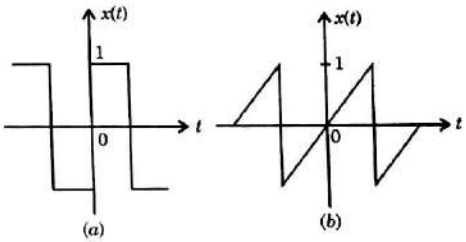


Fig. 2.19. *Graphs of odd functions.*

(*iii*) The sum or product of two or more even functions is an even function.

(*iv*) The sum of two or more odd functions is an odd function and the product of two odd functions is an even-function.

Example 2.3. Figure 2.20 shows a periodic square wave signal, which is symmetrical with respect to the vertical axis. Obtain its Fourier series representation.

Solution: The expression for Fourier series is given as

$$x(t) = a_0 + \sum_{n=1}^{\infty}(a_n \cos n\omega_0 t + b_n \sin n\omega_0 t) \qquad ...(i)$$

Since the given waveform is symmetrical about the vertical axis, *i.e.*, $x(t) = x(-t)$, so, in Fourier series representation only cosine terms are present. This means that $b_n = 0$.

In addition, since the given waveform is symmetrical about the horizontal axis, the average area is zero and hence the d.c. term $a_0 = 0$.

Putting the value of a_0 and b_n in equation (*i*)

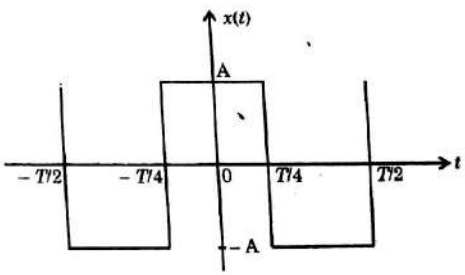


Fig. 2.20.

$$x(t) = \sum_{n=1}^{\infty} a_n \cos n\omega_0 t \qquad ...(ii)$$

Also, from figure, the function $x(t)$ is given by

$$x(t) = -A, \text{ for } -T/2 < t < -T/4 + A, \text{ for } -T/4 < t < +T/4 \qquad ...(iii)$$
$$-A, \text{ for } +T/4 < t < +T/2$$

Now, we shall find the value of a_n,

$$a_n = \frac{2}{T}\int_{-T/2}^{T/2} x(t)\cos n\omega_0 t\, dt$$

$$a_n = \frac{2}{T}\int_{-T/2}^{-T/4} -A\cos n\omega_0 t\, dt + \frac{2}{T}\int_{-T/4}^{T/4} A\cos n\omega_0 t\, dt + \frac{2}{T}\int_{T/4}^{T/2} -A\cos n\omega_0 t\, dt$$

or $$a_n = \frac{2A}{T}\int_{-T/2}^{-T/4} (-\cos n\omega_0 t)\, dt + \int_{-T/4}^{T/4} \cos n\omega_0 t\, dt + \int_{T/4}^{T/2} (-\cos n\omega_0 t)\, dt$$

$$a_n = \frac{2A}{T}\left[\left(\frac{-\sin n\omega_0 t}{n\omega_0}\right)_{-T/2}^{-T/4} + \left(\frac{\sin n\omega_0 t}{n\omega_0}\right)_{-T/4}^{T/4} + \left(\frac{-\sin n\omega_0 t}{n\omega_0}\right)_{T/4}^{T/2}\right]$$

or $$a_n = \frac{2A}{n\omega_0 T}\left[-\sin\left(\frac{-n\omega_0 T}{4}\right) + \sin\left(\frac{-n\omega_0 T}{2}\right) + \sin\left(\frac{n\omega_0 T}{4}\right)\right]$$
$$\left[-\sin\left(\frac{-n\omega_0 T}{4}\right) - \sin\left(\frac{n\omega_0 T}{2}\right) + \sin\left(\frac{n\omega_0 T}{4}\right)\right]$$

or $$a_n = \frac{8A}{n\omega_0 T}\sin\left(\frac{n\omega_0 T}{4}\right) - \frac{4A}{n\omega_0 T}\sin\left(\frac{n\omega_0 T}{2}\right)$$

We know that $\omega_0 = \dfrac{2\pi}{T}$ or $\omega_0 T = 2\pi$

So, that the second term in the above expression is zero for all integer values of n.

Therefore, $$a_n = \frac{8A}{2n\pi}\sin\left(\frac{n\pi}{2}\right) = \frac{4A}{n\pi}\sin\left(\frac{n\pi}{2}\right)$$

Putting the values of a_n in equation (ii)

$$x(t) = \sum_{n=1}^{\infty} \frac{4A}{n\pi}\sin\left(\frac{n\pi}{2}\right).\cos n\omega_0 t; \qquad n = 1, 2, 3 .....$$

$$x(t) = +\frac{4A}{3\pi}\sin\frac{3\pi}{2}\cos 3\omega_0 t + ...$$

$$x(t) = \frac{4A}{\pi}\cdot 1\cos\omega_0 t + 0 + \frac{4A}{3\pi}\cdot(-1)\cos 3\omega_0 t + ...$$

$$x(t) = \frac{4A}{\pi}\left[\cos\omega_0 t - \frac{1}{3}\cos 3\omega_0 t + \frac{1}{5}\cos(5\omega_0 t)...\right]$$

Example 2.4. Figure 2.21 shows the periodic rectangular waveform. Obtain its Fourier series representation.

Solution: The given waveform for one period may be written as

$$x(t) = \begin{cases} 0 & \text{for } -T/2 < t < -T/4 \\ A & \text{for } -T/4 < t < T/4 \\ 0 & \text{for } T/4 < t < T/2 \end{cases} \qquad ...(i)$$

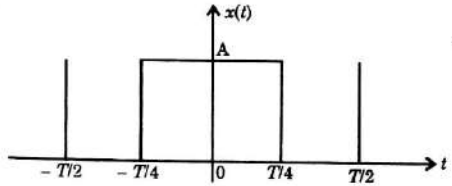


Fig. 2.21.

From figure it is clear that given waveform is an even function of time since $x(-t) = x(t)$ and, therefore, $b_n = 0$.

Fourier series expression is given by

$$x(t) = a_0 + \sum_{n=1}^{\infty}(a_n \cos n\omega_0 t + b_n \sin n\omega_0 t) \qquad (\because b_n = 0)$$

$$x(t) = a_0 + \sum_{n=1}^{\infty} a_n \cos n\omega_0 t \qquad ...(ii)$$

or $$x(t) = a_0 + a_1\cos\omega_0 t + a_2\cos 2\omega_0 t + a_3\cos 3\omega_0 t + ...$$

Now, will find the values of a_0 and a_n

The value of a_0 (d.c. term) is given as

$$a_0 = \frac{1}{T}\int_{-T/2}^{T/2} x(t)\, dt = \frac{1}{T}\int_{-T/4}^{T/4} A\cdot dt = \frac{1}{T}[At]_{-T/4}^{T/4}$$

$$a_0 = \frac{A}{T}[T/4 - (-T/4)] = \frac{A}{T}\cdot\frac{T}{2} = \frac{A}{2}$$

Now, $$a_n = \frac{2}{T}\int_{-T/2}^{T/2} x(t)\cos n\omega_0 t\, dt = \frac{2}{T}\int_{-T/4}^{T/4} A.\cos n\omega_0 t\, dt = \frac{2A}{T}\left[\frac{\sin n\omega_0 t}{n\omega_0}\right]_{-T/4}^{T/4}$$

$$a_n = \frac{2A}{n\omega_0 T}\left[\sin n\omega_0\frac{T}{4} - \sin\left(-n\omega_0\frac{T}{4}\right)\right] = \frac{2A}{n\omega_0 T}\cdot 2\sin\left(n\omega_0\frac{T}{4}\right)$$

$$a_n = \frac{4A}{n\omega_0 T}\sin\left(n\omega_0\frac{T}{4}\right)$$

But $$\omega_0 = \frac{2\pi}{T} \Rightarrow \omega_0 T = 2\pi$$

$$\therefore \quad a_n = \frac{4A}{2nA}\sin\left(\frac{2n\pi}{4}\right) = \frac{2A}{nA}\sin\left(\frac{n\pi}{2}\right)$$

Putting $n = 1, 2, 3, ...$

$$a_1 = \frac{2A}{\pi}\cdot\sin\left(\frac{\pi}{2}\right) = \frac{2A}{\pi}\cdot 1 = \frac{2A}{\pi}$$

$$a_2 = \frac{2A}{2\pi}\cdot\sin\left(\frac{2\pi}{2}\right) = 0$$

$$a_3 = \frac{2A}{3\pi}\cdot\sin\left(\frac{3\pi}{2}\right) = \frac{2A}{3\pi}\cdot(-1) = -\frac{2A}{3\pi}$$

Putting all the values in equation (*ii*)

$$x(t) = \frac{A}{2} + \frac{2A}{\pi}\left(\cos\omega_0 t - \frac{1}{3}\cos 3\omega_0 t + \frac{1}{5}\cos 5\omega_0 t + ...\right)$$

Example 2.5. Shows the graph for a half-wave rectified sine wave. Obtain its trigonometric Fourier Series representations.

Solution: From figure 2.22, it may be observed that the given waveform has no symmetry. Therefore, the Fourier series representation may contain sine and cosine terms.

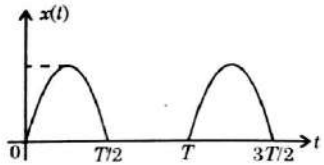


Fig. 2.22.

From figure, we have

$$x(t) = A\sin\omega_0 t$$

We know that the expression for trigonometric Fourier series is given as

$$x(t) = a_0 + \sum_{n=1}^{\infty}(a_n\cos n\omega_0 t + b_n\sin n\omega_0 t)$$

or

$$x(t) = a_0 + a_1\cos\omega_0 t + a_2\cos 2\omega_0 t + ... + a_n\cos n\omega_0 t + ... + b_1\sin\omega_0 t + b_2\sin 2\omega_0 t + ... + b_n\sin n\omega_0 t + ... \quad ...(i)$$

Now, we shall evaluate the values of a_0, a_n and b_n.

The value of a_0 is given as

$$a_0 = \frac{1}{T}\int_0^T x(t)\,dt$$

Putting the value of $x(t)$, we get

$$a_0 = \frac{1}{T}\int_0^T A\sin\omega_0 t\,dt$$

From figure we may write

$$a_0 = \frac{1}{T}\int_0^{T/2} A\sin\omega_0 t = \frac{A}{T}\int_0^{T/2}\sin\omega_0 t\,dt = \frac{A}{T}\left[-\frac{\cos\omega_0 t}{\omega_0}\right]_0^{T/2}$$

$$a_0 = -\frac{A}{T\omega_0}[\cos\omega_0 t]_0^{T/2} = -\frac{A}{T\omega_0}\left[\cos\frac{\omega_0 T}{2} - \cos 0\right]$$

since $\quad T = \dfrac{2\pi}{\omega_0}$

therefore, $\omega_0 T = 2\pi$

Hence $\quad a_0 = -\dfrac{A}{2\pi}[\cos\pi - \cos 0] = \dfrac{-A}{2\pi}[-1-1] \quad a_0 = \dfrac{2A}{2\pi} = \dfrac{A}{\pi} \quad ...(ii)$

The value of a_n is given as

$$a_n = \frac{2}{T}\int_0^T x(t)\cos n\omega_0 t\,dt = \frac{2}{T}\int_0^T A\sin\omega_0 t\cdot\cos n\omega_0 t\,dt$$

or

$$a_n = \frac{2A}{\omega_0 T}\left[\frac{-n\sin\omega_0 t\sin n\omega_0 t - \cos n\omega_0 t\cos\omega_0 t}{-n^2+1}\right]_0^{T/2}$$

since $\quad \omega_0 = \dfrac{2\pi}{T}$

So that $\quad \omega_0 T = 2\pi$

Therefore, $\quad a_n = \dfrac{A}{\pi(1-n^2)}(\cos n\pi + 1) = \begin{cases} \dfrac{2A}{\pi(1-n^2)} & \text{for even } n \\ 0 & \text{for odd } n \end{cases} \quad ...(iii)$

It may be observed from above expression that for $n = 1$, the value of a_n becomes infinite and therefore we have to integrate separately for a_1 as under:

$$a_1 = \frac{2}{T}\int_0^T x(t)\cos\omega_0 t\,dt = \frac{2}{T}\int_0^{T/2} A\sin\omega_0 t\cos\omega_0 t\,dt$$

or

$$a_1 = \frac{A}{T}\int_0^{T/2}\sin 2\omega_0 t\,dt = \frac{A}{2\omega_0 t}[-\cos 2\omega_0 t]_0^{T/2} = -\frac{A}{2\omega_0 t}[\cos\omega_0 T - \cos 0]$$

But $\quad \omega_0 T = 2\pi$

Therefore $a_1 = -\dfrac{A}{2(2\pi)}[-\cos 2\pi - \cos 0] = -\dfrac{A}{4\pi}[1-1] \quad a_1 = 0 \quad ...(iv)$

The value of b_n is expressed as

$$b_n = \frac{2}{T}\int_0^T x(t)\sin n\omega_0 t\,dt = \frac{2}{T}\int_0^{T/2} A\sin\omega_0 t\cdot\sin n\omega_0 t\,dt$$

or

$$b_n = \frac{2A}{\omega_0 T}\left[\frac{n\sin\omega_0 t\cos n\omega_0 t - \sin n\omega_0 t\cos\omega_0 t}{-n^2+1}\right]_0^{T/2} \quad ...(v)$$

Putting $\omega_0 T = 2\pi$, we have

$$b_n = 0$$

From expression in equation (*v*), it may be observed that for $n = 1$, the value of b_n is infinite. Therefore, we have to evaluate the value of b_1 separately as follow:

$$b_1 = \frac{2}{T}\int_0^T x(t)\sin\omega_0 t\,dt = \frac{2}{T}\int_0^{T/2} A\sin\omega_0 t\sin\omega_0 t\,dt$$

or

$$b_1 = \frac{2A}{T}\int_0^{T/2}\sin^2\omega_0 t\,dt = \frac{2A}{\omega_0 T}\left[\frac{\omega_0 t}{2} - \frac{\sin 2\omega_0 t}{4}\right]_0^{T/2}$$

Putting $\omega_0 T = 2\pi$, we have

$$b_1 = \frac{A}{2} \quad ...(vi)$$

Putting the values of all the coefficients in equation (*i*), we get the required expression as

$$x(t) = \frac{A}{\pi}\left[1+\frac{\pi}{2}\sin\omega_0 t\,\frac{-2}{3}\cos 2\omega_0 t\,\frac{-2}{15}\cos 4\omega_0 t\ldots\right] \quad \textbf{Ans.}$$

2.6.3. Polar Fourier Series Representation

The polar Fourier series is the modified form of the trigonometric Fourier series. The polar Fourier series may be derived from trigonometric Fourier series as follow:

Trigonometric Fourier series is given as

$$x(t) = a_0 + \sum_{n=1}^{\infty}(a_n \cos n\omega_0 t + b_n \sin n\omega_0 t)$$

or
$$x(t) = a_0 + \sum_{n=1}^{\infty}\sqrt{a_n^2+b_n^2}\left[\frac{a_n}{\sqrt{a_n^2+b_n^2}}\cos n\omega_0 t + \frac{b_n}{\sqrt{a_n^2+b_n^2}}\sin n\omega_0 t\right]$$

Putting $\dfrac{a_n}{\sqrt{a_n^2+b_n^2}} = \cos\phi_n$ and $\dfrac{b_n}{\sqrt{a_n^2+b_n^2}} = \sin\phi_n$

So that $\tan\phi_n = \dfrac{b_n}{a_n}$

or
$$\phi_n = \tan^{-1}\left(\frac{b_n}{a_n}\right)$$

$$x(t) = a_0 + \sum_{n=1}^{\infty}\sqrt{a_n^2+b_n^2}\left[\cos n\omega_0 t\cos\phi_n + \sin n\omega_0 t\sin\phi_n\right]$$

$$x(t) = a_0 + \sum_{n=1}^{\infty}\sqrt{a_n^2+b_n^2}\cdot\cos(n\omega_0 t - \phi_n)$$

Now, let $D_0 = a_0$

and $D_n = \sqrt{a_n^2+b_n^2}$

So that
$$x(t) = D_0 + \sum_{n=1}^{\infty} D_n \cos(n\omega_0 t - \phi_n) \qquad \text{...(2.18)}$$

$$\phi_n = \tan^{-1}\left(\frac{b_n}{a_n}\right) \qquad \text{...(2.19)}$$

The coefficients D_n are called spectral amplitudes. This means that D_n is the amplitude of the spectral component $D_n \cos(n\omega_0 t - \phi_n)$ having a frequency $n\omega_0$. Equivalently, D_n is also called as amplitude spectrum of function $x(t)$. ϕ_n specifies the phase information of the spectral component $n\omega_0$. Equivalently, ϕ_n is also called as phase spectrum of function $x(t)$.

2.6.4. Complex Fourier Exponential Series

The exponential form of Fourier series is simpler and more compact and hence this is most widely used in signal analysis.

We will derive an expression for complex Fourier Exponential series with the help of trigonometric Fourier series.

The trigonometric Fourier series is given as

$$x(t) = a_0 + \sum_{n=1}^{\infty} a_n \cos n\omega_0 t + b_n \sin n\omega_0 t \qquad \text{...(2.20)}$$

We know, Euler's identity

$$e^{j\theta} = \cos\theta + j\sin\theta \quad \text{and} \quad e^{-j\theta} = \cos\theta - j\sin\theta$$

Adding $2\cos\theta = e^{j\theta} + e^{-j\theta}$

or $\cos\theta = \dfrac{e^{j\theta}+e^{-j\theta}}{2}$

Subtracting $2j\sin\theta = e^{j\theta} - e^{-j\theta}$

or $\sin\theta = \dfrac{e^{j\theta}-e^{-j\theta}}{2j}$

Hence, $\cos n\omega_0 t = \dfrac{e^{jn\omega_0 t}+e^{-jn\omega_0 t}}{2}$

and $\sin n\omega_0 t = \dfrac{e^{jn\omega_0 t}-e^{-jn\omega_0 t}}{2}$

Substituting the values of $\cos n\omega_0 t$ and $\sin n\omega_0 t$ in equation (2.20).

$$x(t) = a_0 + \sum_{n=1}^{\infty} a_n\left(\frac{e^{jn\omega_0 t}+e^{-jn\omega_0 t}}{2}\right) + b_n\left(\frac{e^{jn\omega_0 t}-e^{-jn\omega_0 t}}{2j}\right)$$

or
$$x(t) = a_0 + \sum_{n=1}^{\infty}\left[\frac{(a_n - jb_n)e^{jn\omega_0 t}}{2} + \frac{(a_n + jb_n)e^{-jn\omega_0 t}}{2}\right] \qquad \text{...(2.21)}$$

Let $C_n = \dfrac{1}{2}(a_n - jb_n)$

$C_{-n} = \dfrac{1}{2}(a_n + jb_n)$

$C_0 = a_0$

where C_{-n} is the complex conjugate of C_n.

$$C_n = C_{-n}$$

Putting these values in equation (2.21)

$$x(t) = C_0 + \sum_{n=1}^{\infty} C_n \cdot e^{jn\omega_0 t} + \sum_{n=1}^{\infty} C_{-n}\cdot e^{-jn\omega_0 t}$$

or
$$x(t) = C_0 + \sum_{n=1}^{\infty} C_n \cdot e^{jn\omega_0 t} + \sum_{n=-\infty}^{-1} C_n \cdot e^{jn\omega_0 t}$$

or
$$x(t) = \sum_{n=-\infty}^{\infty} C_n \cdot e^{jn\omega_0 t}$$

C_0 has been included under the Σ sign by using the values of $n = 0$.

Now, substituting expression for coefficients a_n and b_n in the expression

$$C_n = \frac{1}{2}(a_n - jb_n)$$

or
$$C_n = \frac{1}{T}\int_{-\tau/2}^{\tau/2} x(t)\left[\cos n\omega_0 t - j\sin n\omega_0 t\right]dt$$

or
$$C_n = \frac{1}{T}\int_{-\tau/2}^{\tau/2} x(t)e^{-jn\omega_0 t}\,dt \qquad ...(2.22)$$

Similarly,
$$C_{-n} = \frac{1}{T}\int_{-\tau/2}^{\tau/2} x(t)e^{jn\omega_0 t}\,dt \qquad ...(2.23)$$

2.6.5. Concept of Negative Frequency

Complex exponential Fourier series represents a periodic function $x(t)$ as a linear combination of complex exponential functions of frequencies $0, \pm\omega_0, \pm 2\omega_0, \pm 3\omega_0$ ... How can we interpret the negative frequencies.

The negative frequency is present in complex exponential Fourier series because of the fact that the mathematical model of a signal requires the use of negative frequencies. Infact, a real function $x(t)$ cannot be represented by the complex exponential function $e^{jn\omega_0 t}$ without associating them with their opposite rotating counterpart $e^{-jn\omega_0 t}$. The function $e^{-jn\omega_0 t}$ of frequency $-n\omega_0$ (negative frequency) represents the counterpart of the signal $e^{jn\omega_0 t}$ of frequency $n\omega_0$ (positive frequency) and when these two exponential signals are combined, they provide a real function of time.

Actually, the complex-valued basic or elementary functions like $e^{jn\omega_0 t}$ and $e^{-jn\omega_0 t}$ have no physical meaning separately, however they provide real functions like $\cos n\omega_0 t$ or $\sin n\omega_0 t$ when combined together.

This means that negative frequency signals are not physical signals. They are only a mathematical tool or concept required to provide a real signal by a combination of complex exponentials of positive and negative frequencies.

2.6.6. Conversion of Trigonometric Fourier Series into Complex Exponential Fourier Series

With the help of trigonometric Fourier series, we can find the complex exponential Fourier series as follow:

$$C_0 = a_0$$

$$C_n = \frac{1}{2}(a_n - jb_n)$$

$$C_{-n} = \frac{1}{2}(a_n + jb_n)$$

2.6.7. Conversion of Complex Exponential Fourier Series into Trigonometric Fourier Series

With the help of complex exponential Fourier series, we can find the trigonometric Fourier series as follow:

$$a_0 = C_0 \quad a_n = C_n + C_{-n}$$
$$b_n = j(C_n - C_{-n})$$

Example 2.6. Figure 2.23 shows an impulse train. Determine the exponential Fourier series for this impulse train. Also plot its magnitude and phase specturm.

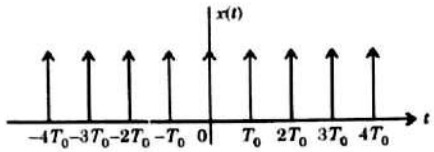


Fig. 2.23. *Unit Impulse train.*

Solution: Let us consider a rectangular pulse train of amplitude A, width τ and period T_0 as shown in figure 2.24.

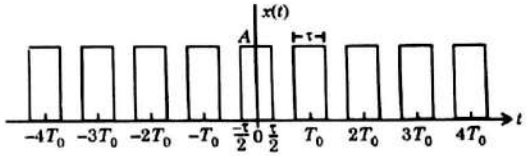


Fig. 2.24.

Now, we shall find exponential Fourier series for this rectangular pulse train.

Exponential Fourier series is expressed as

$$x(t) = \sum_{n=-\infty}^{\infty} C_n . e^{jn\omega_0 t} = \sum_{n=-\infty}^{\infty} C_n . e^{j2\pi nt/T_0}$$

where $\omega_0 = \dfrac{2\pi}{T_0}$

Here
$$C_n = \frac{1}{T_0}\int_{-T/2}^{T/2} x(t)\, e^{-jn\omega_0 t}dt = \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} x(t)\, e^{-j2\pi nt/T_0}dt$$

where $\omega_0 = \dfrac{2\pi}{T_0}$

or
$$C_n = \frac{1}{T_0}\int_{-\tau/2}^{\tau/2} A.e^{-j2\pi nt/T_0}dt = \frac{A}{n\pi}\sin\left(\frac{\pi n\tau}{T_0}\right)$$

$$C_n = \frac{A\tau}{T_0}\cdot\frac{\sin(\pi n\tau/T_0)}{(\pi n\tau/T_0)} = \frac{A\tau}{T_0}\cdot\sin c\left(\frac{\tau}{T_0}\right) \qquad ...(i)$$

It may be observed that the rectangtular pulse of figure 2.25 would approach unit imulse if its area $A\tau \to 1$ and $\tau \to 0$ as shown in figure 2.25.

An unit impulse as a limiting case of rectangular pulse.

Hence, it is obvious that we can obtain C_n for an unit impulse train simply by applying limits to equation (*i*), *i.e.*,

$$C_n = \lim_{\substack{A\tau \to 1 \\ \tau \to 0}} \frac{A\tau}{T_0}\sin c\left(\frac{n\tau}{T_0}\right)$$

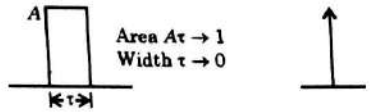


Fig. 2.25. *An unit impulse as a limiting case of rectangular pulse.*

www.dahiyastech.blogspot.com

or $$C_n = \lim_{\tau \to 0} \frac{1}{T_0} \operatorname{sinc}\left(\frac{n\tau}{T_0}\right) \qquad \left[\because \lim_{A\tau \to 1} (A\tau) = 1\right]$$

or $$C_n = \frac{1}{T_0} \lim_{\tau \to 0} \operatorname{sinc}\left(\frac{n\tau}{T_0}\right) = \frac{1}{T_0} \operatorname{sinc}(0) = \frac{1}{T_0} \cdot 1$$

or $$C_n = \frac{1}{T_0} \qquad \text{for unit impulse train} \qquad \text{...(ii)}$$

Thus, the exponential Fourier series will be

$$x(t) = \sum_{n=-\infty}^{\infty} C_n \cdot e^{j2\pi nt/T_0}$$

Putting the value of C_n, we get

$$x(t) = \frac{1}{T_0} \sum_{n=-\infty}^{\infty} e^{j2\pi nt/T_0}$$

This is the required Fourier series for unit impulse train.

The amplitude and phase spectrum have been shown in figures 2.26 and 2.27.

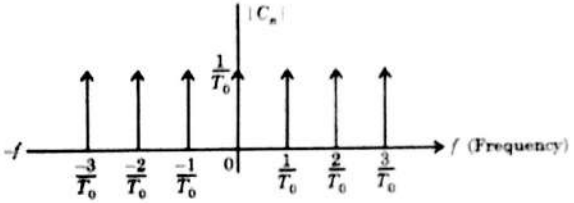


Fig. 2.26. *Amplitude spectrum of unit imulse train.*

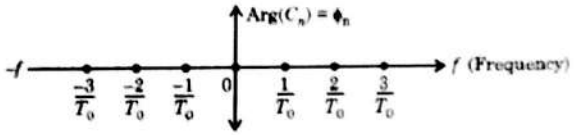


Fig. 2.27. *Phase spectrum of unit imulse train.*

Example 2.7. Figure 2.28 shows the sawtooth waveform. Find the exponential Fourier series for this sawtooth waveform. Also plot the magnitude and phase spectrum.

Solution: Over one period, the signal $x(t)$ may be represented as

$x(t) = t \qquad$ for $0 < t < 1$

Period $T_0 = 1$

Fourier series coefficient C_n is expressed as

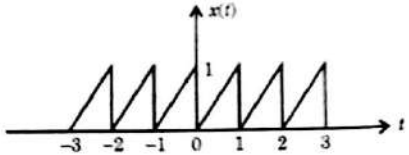


Fig. 2.28. *A sawtooth waveform.*

$$C_n = \frac{1}{T_0} \int_0^{T_0} x(t)\, e^{-j2\pi nt/T_0} dt$$

or $$C_n = \frac{1}{1} \int_0^1 t.e^{-j2\pi nt} dt \qquad [\because T_0 = 1]$$

To simplify above expression, we must use

$$\int xe^{ax} dx = e^{ax}\left(\frac{x}{a} - \frac{1}{a^2}\right)$$

Here, taking $\quad a = -j2\pi n$

Then C_n will be $C_n = \dfrac{-j}{2\pi n} \qquad$ (for $n \neq 0$ on simplification) $\qquad$ ...(i)

with $n = 0$, C_0 would be

$$C_0 = \frac{1}{T_0} \int_0^{T_0} x(t)\, dt = \int_0^1 t\, dt = \frac{1}{2}$$

From equation (i), the magnitude of C_n will be

$$|C_n| = \sqrt{0 + \frac{1}{(2\pi n)^2}} = \frac{1}{2\pi n}$$

Also $$\phi_n = \arg\,(C_n) = -\tan^{-1}\left(\frac{-1/2\pi n}{0}\right)$$

or $$\phi_n = \pm 90°$$

Since phase spectrum is an odd function, therefore

$$\arg\,(C_n) = -90° \text{ for } n > 0 + 90° \quad \text{for } n < 0 \quad \textbf{Ans.}$$

2.7. Properties of Continuous-Time Fourier Series (CTFS)

After the introduction to Continuous-Time Fourier Series (CTFS), let us now discuss its properties. The continuous-time Fourier series (CTFS) representation consists of a number of properties which may be used to develop conceptual fundamentals and in reducing the complexity of Fourier series of serval signals.

Following are the important properties of continuous-time Fourier series (CTFS):

1. Linearity
2. Time-shifting
3. Frequency-shifting
4. Time scaling
5. Time-differentiation
6. Convolution
7. Multiplication
8. Symmetry property

2.7.1. Linearity

We know that the complex exponential continuous-time Fourier series (CTFS) is expressed as

$$x(t) = \sum_{n=-\infty}^{\infty} C_n . e^{jn\omega_0 t}$$

where $x(t)$ is any continuous-time periodic signal and C_n is Fourier series coefficient.

Thus, in above equation, $x(t)$ has coefficients C_n.

For simplicity, let us denote these coefficients as C_{xn}. This means that we can represent the signal $x(t)$ and its coefficients C_{xn} as a continuous-time Fourier series pair *i.e.*,

$$x(t) \xleftrightarrow{CTFS} C_{xn}$$

Then the linearity property may be stated as:

If $\qquad x(t) \xleftrightarrow{CTFS} C_{xn}$

and $\qquad y(t) \xleftrightarrow{CTFS} C_{yn}$

Then $z(t) = a\,x(t) + by(t) \xleftrightarrow{CTFS} C_{zn} = aC_{xn} + bC_{yn}$

2.7.2. Time Shifting Property

Time shifting property states that

If $$x(t) \xleftrightarrow{CTFS} C_{xn}$$

Then $$x(t - t_0) \xleftrightarrow{CTFS} e^{-jn\omega_0 t_0} \cdot C_{xn}$$

Thus, magnitude of Fourier series coefficients remain unchanged when a periodic signal is shifted in time.

2.7.3. Frequency Shifting Property

Frequency shifting property states that

If $$x(t) \xleftrightarrow{CTFS} C_{xn}$$

Then $$e^{jn_0\omega_0 t}x(t) \xleftrightarrow{CTFS} C_{x(n-n_0)}$$

Thus, from above, it is clear that shifting the frequency components by n_0 is equivalent to multiplying $x(t)$ by $e^{jn_0\omega_0 t}$.

2.7.4. Time Scaling

Time scaling property states that if $x(t)$ is a periodic signal with period T_0 then $x(t)$ will be a periodic signal with peroid T_0/α, where α is a positive real number. Thus, with the application of time-scaling property, the period of a signal changes.

Now, it may be observed that since the time-scaling property applies to each of the harmonic components of a signal $x(t)$, therefore the Fourier series coefficients for each of these harmonic components remain the same.

Mathematically,

If $$x(t) \xleftrightarrow{CTFS} C_{xn}$$

then $$x(at) \xleftrightarrow{CTFS} C_{xn}$$

Thus, after time-scaling, the Fourier series coefficients are same. However, the spacing between the frequency components changes from $\frac{1}{T_0}$ to $\frac{a}{T_0}$.

2.7.5. Time-Differentiation

Time-differentiation property states that

If $$x(t) \xleftrightarrow{CTFS} C_{xn}$$

then $$\frac{d}{dx}x(t) \xleftrightarrow{CTFS} j.\frac{2\pi n}{T_0}C_n$$

Thus, differentiating a signal $x(t)$ is equivalent to multiplying its coefficients by $\frac{j2\pi n}{T_0}$.

2.7.6. Time-Convolution

The time-convolution property states that

If $$x(t) \xleftrightarrow{CTFS} C_{xn}$$

and $$y(t) \xleftrightarrow{CTFS} C_{yn}$$

then $$z(t) = x(t) \otimes y(t) \xleftrightarrow{CTFS} C_{zn} = T_0 \cdot C_{xn}\, C_{yn}$$

Thus, convolution of two periodic signals results in multiplication of their coefficients and period T_0.

2.7.7. Multiplication

The multiplication property states that

If $$x(t) \xleftrightarrow{CTFS} C_{xn}$$

and $$y(t) \xleftrightarrow{CTFS} C_{yn}$$

then $$z(t) = x(t) \cdot y(t) \xleftrightarrow{CTFS} C_{zn} = T_0 \cdot C_{xn} \otimes C_{yn}$$

Thus, it is clear that the multiplication of the two signals in time-domain results in convolution of their coefficients.

2.7.8. Symmetry Properties

Symmetry properties state that

If $x(t)$ is real then

$$C_n^* = C_{-n}$$

If $x(t)$ is imaginary then

$$C_n^* = -C_{-n}$$

2.8. Analysis of a Non-periodic Function Over Entire Interval; The Continuous-Time Fourier Transform

A non-periodic signal may be assumed as a limiting case of a periodic signal where the period of the signal approaches infinity. We can use this approach to develop the frequency domain representation of a non-periodic signal over an entire interval.

Let us consider a periodic function $x(t)$ having period T. The complex Exponential Fourier Series representation of this function may be written as,

$$x(t) = \frac{1}{T}\sum_{n=-\infty}^{\infty} C_n \cdot e^{jn\omega_0 t} \qquad ...(2.24)$$

where $$\omega_0 = \frac{2\pi}{T}$$

and $$C_n = \int_{-T/2}^{T/2} x(t)e^{-jn\omega_0 t}\,dt \qquad ...(2.25)$$

Figure 2.29 shows a typical discrete spectrum.

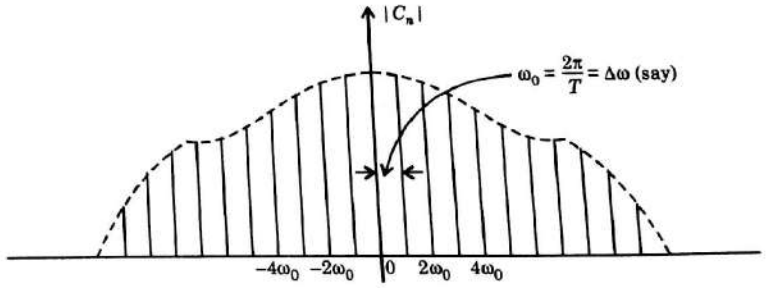


Fig. 2.29. *A typical discrete spectrum.*

From figure 2.29, the spacing between successive harmonics will be

$$\omega_0 = \frac{2\pi}{T} = \Delta\omega \text{ (say)}$$

$$\frac{1}{T} = \frac{\Delta\omega}{2\pi}$$

Putting this value of $\frac{1}{T}$ in the equation (2.24)

$$x(t) = \frac{1}{2\pi}\sum_{n=-\infty}^{\infty} C_n \cdot e^{jn\omega_0 t}(\Delta\omega) \qquad ...(2.26)$$

Now, if we consider the limiting case as $T \to \infty$, then $\Delta\omega \to 0$. This implies that the discrete lines in the spectrum will merge to form a continuous spectrum. Therefore, the infinite sum in equation becomes the ordinary integral.

C_n, now, is defined for all the frequencies rather only for integral multipler of $\omega_0 = \frac{2\pi}{T}$. Hence, the integer n has no significance now and in the limit $T \to \infty$, C_n becomes a continuous function $X(\omega)$ given by

$$X(\omega) = \lim_{T\to\infty} C_n = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}\,dt$$

Hence, the Fourier series in equation (2.26) for a periodic function now becomes the Fourier integral for a non-periodic function over entire integral given as

$$X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}\,dt \qquad ...(2.27)$$

Equation 2.27 is popularly known as **Fourier transform** of non-periodic function $x(t)$ denoted as

$$F[x(t)] = X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}\,dt$$

2.8.1. Fourier Transform Pair

Fourier transform may be expressed as

$$X(\omega) = F[x(t)] = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}\,dt \qquad ...(2.28)$$

In above equation, $X(\omega)$ is called the Fourier transform of $x(t)$. In other words, $X(\omega)$ is the frequency domain representation of time domain function $x(t)$. This means that we are converting a time domain signal into its frequency domain representation with the help of Fourier transform.

Conversely, if we want to convert a frequency domain signal into corresponding time domain signal, we will have to take inverse Fourier transform of the frequency domain signal. Mathematically,

Inverse Fourier Transform

$$F^{-1}[X(\omega)] = x(t)$$

$$x(t) = F^{-1}[X(\omega)] = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(\omega)e^{j\omega t}\,d\omega \qquad ...(2.29)$$

From equations (2.28) and (2.29), it is clear that $X(\omega)$ is the Fourier transform of $x(t)$ and $x(t)$ is the inverse Fourier transform of $X(\omega)$.

$$X(\omega) = F[x(t)]$$

and $$x(t) = F^{-1}[X(\omega)]$$

Hence, $x(t)$ and $X(\omega)$ are a **Fourier transform** pair.

Symbolically, this may be expressed as

$$x(t) \longleftrightarrow X(\omega)$$

Here, double ended arrow means that Fourier transform of $x(t)$ is $X(\omega)$ and inverse Fourier transform of $X(\omega)$ is $x(t)$.

In general, Fourier transform $X(\omega)$ is a complex function of ω and may be expressed as,

$$X(\omega) = |X(\omega)|\, e^{j\angle X(\omega)}$$

or $$X(\omega) = |X(\omega)|\, e^{j\theta(\omega)} \qquad ...(2.30)$$

where $|X(\omega)|$ is called the amplitude spectrum and $\angle X(\omega)$ is called the phase spectrum.

2.8.2. Existence of Fourier Transform

For a function $x(t)$ to be Fourier transformable, it is sufficient that $x(t)$ satisfies Dirichlet's conditions given below:

(*i*) The function $x(t)$ is a single valued function with a finite number of maxima and minima and a finite, number of discontinuities in any finite time interval.

(*ii*) The function $x(t)$ is absolutely integrable. Mathematically,

$$\int_{-\infty}^{\infty} |x(t)|\,dt < \infty \qquad ...(2.31)$$

Example 2.8. Obtain the Fourier transform of a single-sided exponential function $e^{-at} \cdot u(t)$ as shown in figure 2.30. Also draw the spectrum. Here $u(t)$ is the unit step function.

Solution: $e^{-at}.u(t)$ is a single-sided function because here the main function e^{-at} is multiplied by unit step function $u(t)$ which exists only for $t > 0$.

$$u(t) = \begin{cases} 1 & \text{for } t > 0 \\ 0 & \text{elsewhere} \end{cases}$$

Hence, when any function is multiplied by a unit-step function u(t), then the resulting function will exist for positive side only.

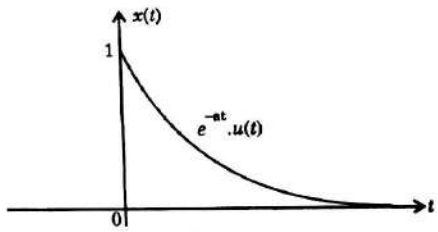


Fig. 2.30.

Now, given that $x(t) = e^{-at}.u(t)$

$$\therefore \quad X(\omega) = F[x(t)] = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}\,dt \; X(\omega)$$

or $$X(\omega) = \int_{-\infty}^{\infty} e^{-a(t)} \cdot u(t) \cdot e^{-j\omega t}\,dt$$

or $$X(\omega) = \int_{-\infty}^{\infty} e^{-a(t)} \cdot 0 \cdot e^{-at}\,dt + \int_{0}^{\infty} e^{-at} \cdot 1 \cdot e^{-j\omega t}\,dt = 0 + \int_{0}^{\infty} e^{-t(a+j\omega)}\,dt$$

or $$X(\omega) = -\frac{1}{a+j\omega}\left[e^{-t(a+j\omega)}\right]_0^{\infty} = -\frac{1}{a+j\omega}\left[e^{-\infty} - e^{0}\right]$$

or $$X(\omega) = -\frac{1}{a+j\omega}[0-1] = \frac{1}{a+j\omega}$$

To obtain $X(\omega)$ in proper form, we write

$$X(\omega) = \frac{1}{a+j\omega}\cdot\frac{(a-j\omega)}{(a-j\omega)}$$

or $$X(\omega) = \frac{a-j\omega}{a^2+\omega^2} = \frac{a}{a^2+\omega^2} - j\cdot\frac{\omega}{a^2+\omega^2}$$

Obtaining the above expression in polar-form

$$X(\omega) = \frac{1}{\sqrt{a^2+\omega^2}}\cdot e^{-j\tan^{-1}}\frac{1}{\sqrt{a^2+\omega^2}}$$

But we know that

$$X(\omega) = |X(\omega)|\, e^{j\phi(j\omega)}$$

On comparison amplitude spectrum

$$|X(\omega)| = \frac{1}{\sqrt{a^2+\omega^2}}$$

and phase spectrum

$$\phi(\omega) = -\tan^{-1}\left(\frac{\omega}{a}\right)$$

The amplitude spectrum $|X(\omega)|$ and the phase spectrum $\phi(\omega)$ are shown in figure 2.31.

The magnitude spectrum has a value $\frac{1}{a}$ at $\omega = 0$ and this value decreases as ω increases. Also the phase spectrum varies between $-\frac{\pi}{2}$ to $\frac{\pi}{2}$ depending upon the values of $-\tan^{-1}\left(\frac{\omega}{a}\right)$ at different frequencies. Both amplitude and phase plots are necessary since Fourier transform is complex.

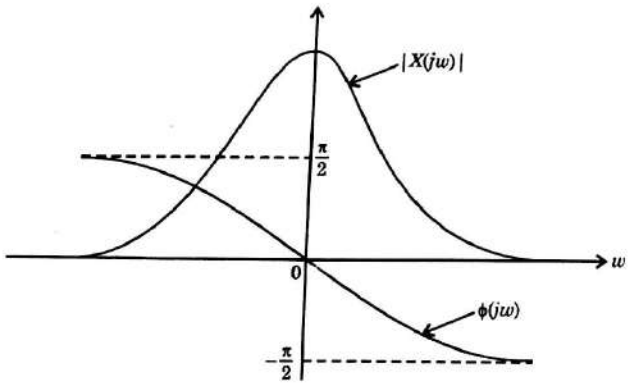


Fig. 2.31. *Amplitude and phase spectrum.*

2.8.3. Some Useful Functions and their Fourier Transforms

In this section, we shall discuss some useful functions such as gate pulse function and Interpolation function.

2.8.3.1. Gate Function

A gate function is a rectangular pulse. Figure 2.32 shows a gate function.

The gate function or rectangular pulse shown in figure 2.32 is written as rect(t).

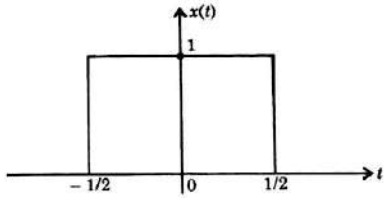


Fig. 2.32. *A Gate function.*

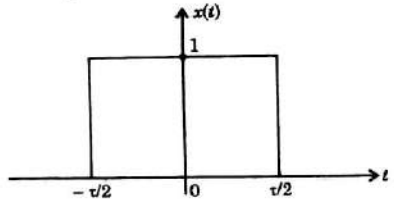


Fig. 2.33.

Similarly, the gate function shown in figure 2.33 is written as rect $\left(\frac{t}{\tau}\right)$.

From above two figures, it is clear that rect $\left(\frac{t}{\tau}\right)$ represents a gate pulse of height or amplitude unity and width τ.

Similarly, rect (t) represents a gate pulse of unit amplitude and unit width.

Hence, we may write the equation for gate pulse shown in figure 2.33 as

$$x(t) = \text{rect}\left(\frac{t}{\tau}\right) = \begin{cases} 1 & \text{for } -\frac{\tau}{2} < t < \frac{\tau}{2} \\ 0 & \text{otherwise} \end{cases} \quad \text{...(2.32)}$$

2.8.3.2. Sampling Function or Interpolating Function or sin c Function

The functions $\frac{\sin x}{x}$ is the "sine over argument" function and is denoted by sin $c(x)$. This function plays an important role in signal processing. It is also known as the filtering or interpolating function. Mathematically,

$$\sin c(x) = \frac{\sin x}{x} \quad \text{...(2.33)}$$

or $$S\,a(x) = \frac{\sin x}{x}$$

Figure 2.34 shows a sin $c(x)$ or sampling function.

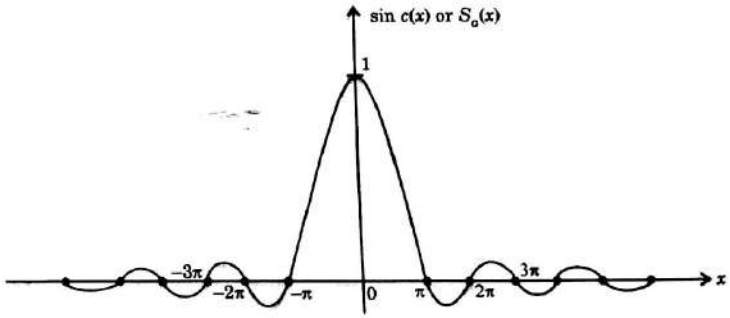


Fig. 2.34.

From the figure, following points may be observed about a sampling function:

(i) $S\,a\,(x)$ or $\sin c\,(x)$ is an even function of x.

(ii) $\sin c\,(x) = 0$ when $\sin x = 0$ except at $x = 0$, where it is indeterminate. This means that $\sin c(x) = 0$ for $x = \pm\, n\pi$ or $x = \pm\,\pi, \pm\,2\,\pi, \pm\,3\,\pi$ .....

(iii) Using L'H hospital's rule, $\sin c(0) = 1$

(iv) $\sin c(x)$ is the product of an oscillating signal $\sin x$ of period 2π and a decreasing function $\frac{1}{x}$. Therefore, $\sin c(x)$ exhibits sinusoidal oscillations of period 2π with amplitude decreasing continuously as $\frac{1}{x}$.

Example 2.9. Find the Fourier transform of the gate function shown in the figure 2.35.

Solution: $x\,(t) = \text{rect}\left(\frac{t}{\tau}\right) = \begin{cases} 1 & \text{for } -\frac{\tau}{2} < t < \frac{\tau}{2} \\ 0 & \text{otherwise} \end{cases}$

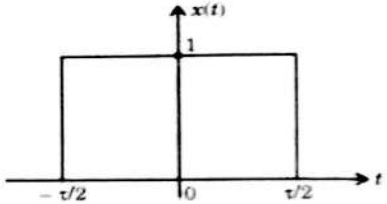


Fig. 2.35.

We know that Fourier transform is given by

$$X\,(\omega) = F\,[x\,(t)] = \int_{-\infty}^{\infty} x\,(t)\,e^{-j\omega t}\,dt$$

$$X\,(\omega) = \int_{-\infty}^{\infty} \text{rect}\left(\frac{t}{\tau}\right) e^{-j\omega t}\,dt = \int_{-\tau/2}^{\tau/2} 1\;e^{-j\omega t}\,dt = \left[\frac{e^{-j\omega t}}{-j\,\omega}\right]_{-\tau/2}^{\tau/2}$$

We know that $e^{j\theta} = \cos\theta + j\sin\theta$

and $e^{-j\theta} = \cos\theta - j\sin\theta$

Hence, $2\cos\theta = e^{j\theta} + e^{-j\theta}$

and $2\,j\sin\theta = e^{j\theta} - e^{-j\theta}$

Putting $\theta = \omega\tau/2$, we get

$$2\,j\sin\frac{\omega\tau}{2} = e^{j\omega\tau/2} - e^{-j\omega\tau/2}$$

Using above equations, we have

$$X\,(\omega) = \frac{1}{j\omega}\left(2\,j\sin\frac{\omega\tau}{2}\right) = \frac{2\sin\omega\tau/2}{\omega} = \frac{\tau\sin(\omega\tau/2)}{\omega\tau/2}$$

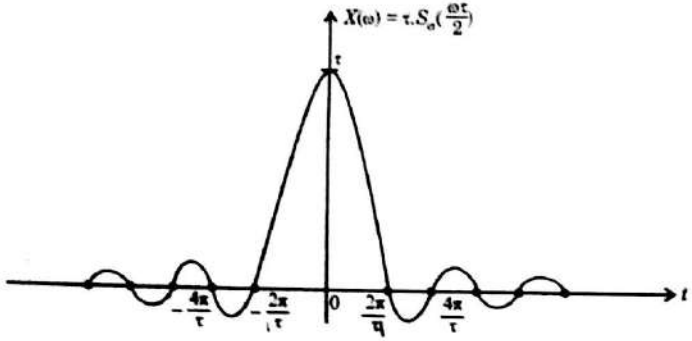


Fig. 2.36.

or $$X\,(\omega) = \tau\cdot\sin c\left(\frac{\omega\tau}{2}\right) = \tau\times Sa\left(\frac{\omega\tau}{2}\right)$$

Now, since $\sin c\,(x) = 0$, when $x = \pm\, n\pi$

Therefore, $\sin c\left(\frac{\omega\tau}{2}\right) = 0$, when $\frac{\omega\tau}{2} = \pm\, n\pi$

or $$\omega = \pm\frac{2n\pi}{\tau} \qquad (n = 1, 2, 3, \ldots)$$

Figure 2.36 shows the plot of $X(\omega)$.

2.8.3.3. *Unit Impulse Functions*

An Unit Impulse Function is one of the most important functions in the analysis of signals and system. This function was invented by P.A.M. Dirac and so it is also called as Delta function. Mathematically,

$$\delta(t) = 0\;t \neq 0$$

and $$\int_{-\infty}^{\infty}\delta(t)\,dt = 1$$

Unit impulse function has already been discussed in Chapter 1.

2.8.3.4. *Shifting Property of the Impulse Function*

If we take the product of unit impulse function $\delta(t)$ and any given function $x(t)$ which is continuous at $t = 0$, then this product will provide the function $x(t)$ existing only at $t = 0$ since $\delta(t)$ exists only at $t = 0$. Mathemtically,

$$\int_{-\infty}^{\infty} x(t)\,\delta(t)\,dt = x\,(0)\int_{-\infty}^{\infty}\delta(t)\,dt = x\,(0)\cdot 1 = x\,(0) \qquad \ldots(2.34)$$

where $x\,(0)$ is the value of the function $x\,(t)$ at $t = 0$.

The equation is known as shifting or sampling property of the impulse function because the impulse shifts the value of $x\,(t)$ at $t = 0$. This means that the value of $x\,(t)$ has been sampled at $t = 0$.

The shifting or sampling may also be done at any, instant $t = t_0$, if we define the impulse function at the instant to. Mathematically,

$$\int_{-\infty}^{\infty} x\,(t)\,\delta(t - t_0)\,dt = x(t_0) \qquad \ldots(2.35)$$

The above equation states that the product of a continuous function $x(t)$ with an impulse function $\delta\,(t - t_0)$ provides the sampled value of $x\,(t)$ at $t = t_0$.

Example 2.10. Solve the following integral:

$$\int_{-\infty}^{\infty}[t^2 + 1]\,\delta(t)\,dt.$$

Solution: The given integral may be written as

$$\int_{-\infty}^{\infty}[t^2 + 1]\,\delta(t)\,dt = \int_{-\infty}^{\infty} t^2\,\delta(t)\,dt + \int_{-\infty}^{\infty}\delta(t)\,dt$$

Using shifting property of impulse function for first factor

$$\int_{-\infty}^{\infty}[t^2+1]\,\delta(t)\,dt = [t^2]_{\text{at } t=0} + 1 = 0 + 1 = 1$$

Example 2.11. Solve the integral: $\int_{-1}^{2}\left[t^4+1\right]\delta(t-1)\,dt.$

Solution: The given integral may be written as

$$\int_{-1}^{2}\left[t^4+1\right]\delta(t-1)\,dt = \int_{-1}^{2} t^4\,\delta(t-1)\,dt + \int_{-1}^{2}\delta(t-1)\,dt$$

Using shifting property of the impulse function for first factor

$$\int_{-1}^{2}\left[t^4+1\right]\delta(t-1)\,dt = [t^4]_{\text{at } t=1} + 1 = 1 + 1 = 2$$

Example 2.12. Find the Fourier transform of an impulse function $x(t) = \delta(t)$. Also draw spectrum.

Solution: The expression for Fourier transform is given by

$$X(\omega) = F[x(t)] = \int_{-\infty}^{\infty} x(t)\,e^{-j\omega t}\,dt = \int_{-\infty}^{\infty}\delta(t)\,e^{-j\omega t}\,dt$$

Using shifting property of impulse function

$$X(\omega) = \left[e^{-j\omega t}\right]_{\text{at } t=0} \qquad [\because \delta(t) \text{ exists at } t = 0]$$

$$X(\omega) = 1$$

or $\qquad \delta(t) \longleftrightarrow 1$

Hence, the Fourier transform of an unit impulse function is unity.

Figure 2.37 shows an unit impulse function and its Fourier transform or spectrum. From figure it is clear that an unit impulse contains the entire frequency components having idential magnitude. This means that the bandwidth of the unit impulse function is infinite.

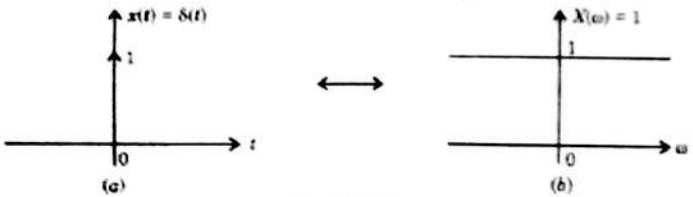


Fig. 2.37.

Also, since the spectrum is real, only magnitude spectrum is required. The phase spectrum $\theta(\omega) = 0$, which means that all the frequency components are in the same phase.

Example 2.13. Find the inverse Fourier transform of $\delta(\omega)$.

Solution: The inverse Fourier transform is expressed as

$$F^{-1}[X(\omega)] = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(\omega)\,e^{j\omega t}\,d\omega$$

Therefore, we have

$$F^{-1}[\delta(\omega)] = \frac{1}{2\pi}\int_{-\infty}^{\infty}\delta(\omega)\,e^{j\omega t}\,d\omega = \frac{1}{2\pi}\left[e^{j\omega t}\right]_{\text{at } \omega=0}$$

$$F^{-1}[\delta(\omega)] = \frac{1}{2\pi}\times e^0 = \frac{1}{2\pi}\cdot 1 = \frac{1}{2\pi}$$

$$F\left[\frac{1}{2\pi}\right] = \delta(\omega)$$

or $\qquad \dfrac{1}{2\pi} \longleftrightarrow \delta(\omega)$

or $\qquad 1 \longleftrightarrow 2\pi\,\delta(\omega)$

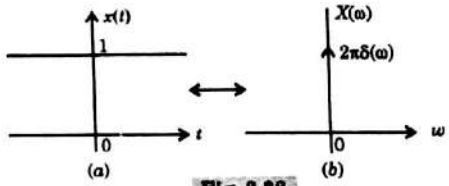


Fig. 2.38.

This shows that the spectrum of a constant signal $x(t) = 1$ an impulse $2\pi\delta(\omega)$.

This can also be interpreted as that $x(t) = 1$ is a d.c. signal which has a single frequency $\omega = 0$ (*dc*).

Example 2.14. Find the inverse Fourier transform of $\delta(\omega - \omega_0)$.

Solution: Inverse Fourier transform is expressed as

$$F^{-1}[X(\omega)] = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(\omega)\,e^{j\omega t}\,d\omega$$

or $\qquad F^{-1}[\delta(\omega-\omega_0)] = \dfrac{1}{2\pi}\displaystyle\int_{-\infty}^{\infty}\delta(\omega-\omega_0)\,e^{j\omega t}\,d\omega$

Using shifting or sampling property of impulse function, we get

$$F^{-1}[\delta(\omega-\omega_0)] = \frac{1}{2\pi}\left[e^{j\omega t}\right]_{\text{at } \omega=\omega_0} = \frac{1}{2\pi}e^{j\omega_0 t}$$

$$F\left[\frac{1}{2\pi}e^{j\omega_0 t}\right] = \delta(\omega-\omega_0)$$

or $\qquad \dfrac{1}{2\pi}e^{j\omega_0 t} \longleftrightarrow \delta(\omega-\omega_0)$

or $\qquad e^{j\omega_0 t} \longleftrightarrow 2\pi\,\delta(\omega-\omega_0)$

The above expression shows that the spectrum of an everlasting exponential $e^{j\omega_0 t}$ is single impulse at $\omega = 0$.

Similarly, $e^{-j\omega_0 t} \longleftrightarrow 2\pi\,\delta(\omega+\omega_0)$

Example 2.15. Find the Fourier transform of the everlasting sinusoid $\cos\omega_0 t$.

Solution: We know that Euler's identity is given by

$$e^{j\theta} = \cos\theta + j\sin\theta \text{ and } e^{-j\theta} = \cos\theta - j\sin\theta$$

$$\cos\theta = \frac{e^{j\theta}+e^{-j\theta}}{2}$$

and $\qquad \sin\theta = \dfrac{e^{j\theta}-e^{-j\theta}}{2j}$

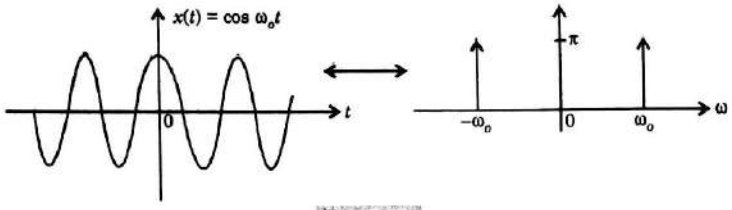


Fig. 2.39.

Hence, $\cos \omega_0 t = \frac{1}{2}\left(e^{j\omega_0 t} + e^{-j\omega_0 t}\right)$

We know that $e^{j\omega_0 t} \longleftrightarrow 2\pi\delta(\omega - \omega_0)$

and $e^{-j\omega_0 t} \longleftrightarrow 2\pi\delta(\omega + \omega_0)$

So that $\cos \omega_0 t \longleftrightarrow \frac{1}{2}[2\pi\delta(\omega - \omega_0) + 2\pi\delta(\omega + \omega_0)]$

or $\cos \omega_0 t \longleftrightarrow [\delta(\omega - \omega_0) + \delta(\omega + \omega_0)]$

Example 2.16. If $\delta(t)$ denotes the delta function, then find the value of the integral

$$\int_{-\infty}^{\infty} \delta(t) \cos\left(\frac{3t}{2}\right) dt$$

(GATE Examination, 2001)

Solution: We know that for a delta function $\delta(t)$, following expression is satisfied.

$$\int_{-\infty}^{\infty} \delta(at)\, \phi(t)\, dt = \frac{\phi(0)}{|a|} \text{ for any } a < 0$$

Therefore, for given integral, we have

$$\int_{-\infty}^{\infty} \delta(t).\cos\left(\frac{3t}{2}\right) dt = \frac{\cos 0}{1} = \frac{1}{1} = 1 \quad \textbf{Ans.}$$

2.8.3.5. Rectangular Pulse

Let us consider a rectangular pulse as shown in figure 2.40.

The expression for this rectangular pulse may be written as

$$x(t) = \begin{cases} 1 & \text{for } 0 < t < T \\ 0 & \text{elsewhere} \end{cases} \qquad ...(2.36)$$

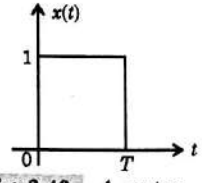


Fig. 2.40. *A rectangular pulse*

We know that Fourier transform is expressed as

$$F[x(t)] = X(\omega) = \int_{-\infty}^{\infty} x(t) e^{-j\omega t}\, dt$$

According to equation (2.36), we write

$$X(\omega) = \int_0^T 1 \cdot e^{-j\omega t}\, dt = \int_0^T e^{-j\omega t}\, dt = \left[\frac{e^{-j\omega t}}{-j\omega}\right]_0^T$$

$$X(\omega) = -\frac{1}{j\omega}\left[e^{-j\omega T} - e^{0}\right] = -\frac{1}{j\omega}\left[e^{-j\omega T} - 1\right]$$

or $$X(\omega) = -\frac{1}{j\omega}\left[\frac{e^{-j\omega T/2}}{e^{j\omega T/2}} - 1\right] = -\frac{1}{j\omega}\left[\frac{e^{-j\omega T/2} - e^{j\omega T/2}}{e^{j\omega T/2}}\right]$$

So that $e^{j\theta} - e^{-j\theta} = 2\sin\theta$

or $$X(\omega) = \frac{1}{j\omega} \cdot \frac{1}{e^{j\omega T/2}}\left[e^{j\omega T/2} - e^{-j\omega T/2}\right] = \frac{2e^{-j\omega T/2}}{\omega}\left[\frac{e^{j\omega T/2} - e^{-j\omega T/2}}{2j}\right]$$

or $$X(\omega) = Te^{-j\omega T/2}\left[\frac{\sin\left(\frac{\omega T}{2}\right)}{\left(\frac{\omega T}{2}\right)}\right] \qquad \left[\begin{array}{l} \because e^{j\theta} = \cos\theta + j\sin\theta \\ e^{-j\theta} = \cos\theta - j\sin\theta \end{array}\right]$$

$$X(\omega) = Te^{-j\omega T/2} \sin c\left(\frac{\omega T}{2}\right)$$

Figure 2.41 shows its amplitude spectrum.

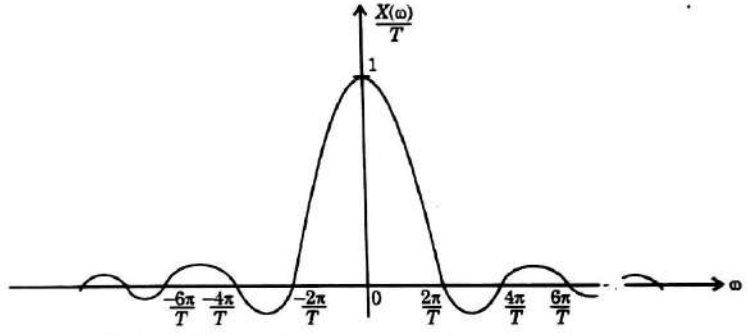


Fig. 2.41. *Amplitude spectrum of a rectangular pulse.*

Example 2.17. Obtain the Fourier transform of a rectangular pulse of duration 2 seconds and having a magnitude of 10 volts as shown in figure 2.42. *(A.M.I.E. Examination, 1997)*

Solution: We know that Fourier transform is expressed as

$$F[x(t)] = X(\omega) = \int_{-\infty}^{\infty} x(t) e^{-j\omega t}\, dt$$

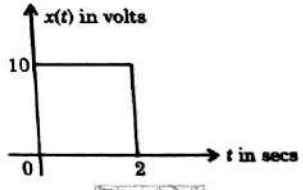


Fig. 2.42.

According to figure 2.42, we may write

$$X(\omega) = \int_0^2 10 \cdot e^{-j\omega t} dt = 10\int_0^2 e^{-j\omega t}\, dt = 10\left[\frac{e^{-j\omega t}}{-j\omega}\right]_0^2$$

$$X(\omega) = \frac{10}{-j\omega}\left[e^{-j\omega t}\right]_0^2 = \frac{10}{-j\omega}\left[e^{-j2\omega} - e^{0}\right] = \frac{10}{-j\omega}\left[e^{-j2\omega} - 1\right]$$

or $$X(\omega) = \frac{10}{-j\omega}\left[\frac{e^{-j\omega}}{e^{j\omega}} - 1\right] = \frac{10}{-j\omega}\left[\frac{e^{-j\omega} - e^{j\omega}}{e^{j\omega}}\right]$$

or $$X(\omega) = \frac{10\,e^{-j\omega}}{j\omega}\left[e^{j\omega} - e^{-j\omega}\right] = \frac{20\,e^{-j\omega}}{\omega}\left[\frac{e^{j\omega} - e^{-j\omega}}{2j}\right]$$

or $$X(\omega) = 20\frac{e^{-j\omega}}{\omega}.\sin\omega = 20\,e^{-j\omega}\left[\frac{\sin\omega}{\omega}\right] = 20\,e^{-jw}\,\text{sinc}\,(\omega) \quad \textbf{Ans.}$$

2.8.3.6. Gaussian Pulse

A pulse which is defined by the expression

$$x(t) = e^{-\pi t^2} \qquad ...(2.37)$$

is known as Gaussian Pulse. Figure 2.43 shows the graphical representation of Gaussian Pulse.

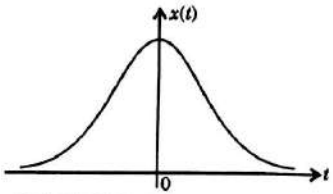


Fig. 2.43. *A Gaussian Pulse.*

Fourier Transform of Gaussian Pulse

We know that Fourier transform is expressed as

$$F[x(t)] = X(\omega) = \int_{-\infty}^{\infty} x(t)\,e^{-j\omega t}\,dt$$

$$F[e^{-\pi t^2}] = X(\omega) = \int_{-\infty}^{\infty} e^{-\pi t^2}.e^{-j\omega t}\,dt$$

or $$X(\omega) = \int_{-\infty}^{\infty} e^{-(\pi t^2 + j\omega t)}\,dt$$

Putting $\pi t^2 + j\omega t = \left(\sqrt{\pi}\,t + \frac{j\omega}{2\sqrt{\pi}}\right)^2 + \frac{\omega^2}{4\pi}$, we have

$$X(\omega) = \int_{-\infty}^{\infty} e^{-\left(\sqrt{\pi}\,t + \frac{j\omega}{2\sqrt{\pi}}\right)^2}.e^{-\omega^2/4\pi}\,dt = e^{-\omega^2/4\pi}\int_{-\infty}^{\infty} e^{-\left(\sqrt{\pi}\,t + \frac{j\omega}{2\sqrt{\pi}}\right)^2}\,dt$$

Again putting $u = \sqrt{\pi}\,t + \frac{j\omega}{2\sqrt{\pi}}$

So that $du = \sqrt{\pi}\,dt$

and hence $dt = \frac{du}{\sqrt{\pi}}$

Therefore $$X(\omega) = e^{-\omega^2/4\pi}.2\int_0^{\infty} e^{-u^2}\frac{du}{\sqrt{\pi}} = e^{-\omega^2/4\pi}.\frac{2}{\sqrt{\pi}}.\int_0^{\infty} e^{-u^2}\,du$$

or $$X(\omega) = e^{-\omega^2/4\pi}.\frac{2}{\sqrt{\pi}}.\frac{\sqrt{\pi}}{2} \qquad \left[\because \int_0^{\infty} e^{-u^2}\,du = \frac{\sqrt{\pi}}{2}\right]$$

or $$X(\omega) = e^{-w^2/4\pi}\,X(j\omega) = e^{-\pi f^2}$$

or $$e^{-\pi t^2} \longleftrightarrow e^{-\pi f^2} \qquad ...(2.37a)$$

Example 2.18. Determine the Fourier transform of the Gaussian pulse

$$x(t) = e^{-b^2 t^2}$$ *(Indraprastha University, Delhi, 2004)*

Solution: The given function is

$$x(t) = e^{-b^2 t^2}$$

Fourier transform is expressed as

$$F[x(t)] = X(\omega) = \int_{-\infty}^{\infty} x(t)\,e^{-j\omega t}\,dt$$

$$F\left[e^{-b^2 t^2}\right] = X(\omega) = \int_{-\infty}^{\infty} e^{-b^2 t^2}\,e^{-j\omega t}\,dt$$

or $$X(\omega) = \int_{-\infty}^{\infty} e^{-(b^2 t^2 + j\omega t)}\,dt$$

Putting $b^2 t^2 + j\omega t = \left(bt + \frac{j\omega}{2b}\right)^2 + \frac{\omega^2}{4b^2}$, we have

$$X(\omega) = \int_{-\infty}^{\infty} e^{-\left(bt + \frac{j\omega}{2b}\right)^2}.e^{-\frac{\omega^2}{4b^2}}\,dt = e^{-\frac{\omega^2}{4b^2}}\int_{-\infty}^{\infty} e^{-\left(bt + \frac{j\omega}{2b}\right)^2}\,dt$$

Again putting $u = bt + \frac{j\omega}{2b}$

So that $du = bdt$ and hence $dt = \frac{du}{b}$

Therefore, $$X(\omega) = e^{-\frac{\omega^2}{4b^2}}.2\int_0^{\infty} e^{-u^2}\frac{du}{b} = e^{-\frac{\omega^2}{4b^2}}.\frac{2}{b}\int_0^{\infty} e^{-u^2}\,du$$

or $$X(\omega) = e^{-\frac{\omega^2}{4b^2}}.\frac{2}{b}.\frac{\sqrt{\pi}}{2} \qquad \left[\because \int_0^{\infty} e^{-u^2}\,du = \frac{\sqrt{\pi}}{2}\right]$$

or $$X(\omega) = \frac{\sqrt{\pi}}{b}.e^{-\frac{\omega^2}{4b^2}} = \frac{\sqrt{\pi}}{b}.e^{-\left(\frac{\omega}{2b}\right)^2}$$

Hence $$e^{-b^2 t^2} \longleftrightarrow \frac{\sqrt{\pi}}{b}.e^{-\left(\frac{\pi f}{b}\right)^2}$$

2.8.3.7. Signum Function

The signum function is represented by sgn(t). Mathematically, it is defined as

$$\text{sgn}\,(t) = \begin{cases} -1 & \text{for } t < 0 \\ 0 & \text{for } t = 0 \\ 1 & \text{for } t > 0 \end{cases} \qquad ...(2.38)$$

or $$\text{sgn}(t) = u(t) - u(-t) \qquad ...(2.38a)$$

Figure 2.44 shows the graphical representation of signum function.

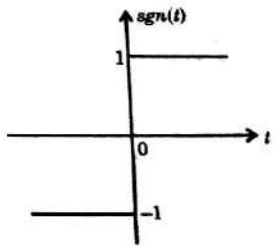


Fig. 2.44. *Signum function.*

Fourier Transform of Signum function

We know that Fourier transform is given as

$$F[x(t)] = X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}\,dt$$

$$F[\text{sgn}(t)] = X(\omega) = \int_{-\infty}^{\infty} \text{sgn}(t)e^{-j\omega t}\,dt$$

According to equation (2.38), we write

$$X(\omega) = \int_{-\infty}^{0} (-1)e^{-j\omega t}\,dt + \int_{0}^{\infty} 1\cdot e^{-j\omega t}\,dt = -\int_{-\infty}^{0} e^{-j\omega t}\,dt + \int_{0}^{\infty} e^{-j\omega t}\,dt$$

or $$X(\omega) = -\left[\frac{e^{-j\omega t}}{-j\omega}\right]_{-\infty}^{0} + \left[\frac{e^{-j\omega t}}{-j\omega}\right]_{0}^{\infty} = \frac{1}{j\omega}\left[e^{\circ} - e^{-\infty}\right] - \frac{1}{j\omega}\left[e^{-\infty} - e^{\circ}\right]$$

or $$X(\omega) = \frac{1}{j\omega}[1-0] - \frac{1}{j\omega}[0-1] = \frac{1}{j\omega} + \frac{1}{j\omega} = \frac{2}{j\omega} \qquad \text{...(2.39)}$$

Hence $\text{sgn}(t) \longleftrightarrow \dfrac{2}{j\omega}$

Figure 2.45 shows the amplitude and phase spectra of signum function.

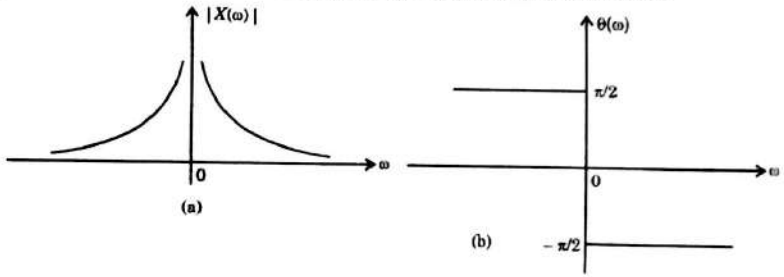


Fig. 2.45. *(a) Amplitude spectrum, (b) Phase spectrum.*

Triangular Pulse

Let us consider the triangular pulse shown in figure 2.46.

From figure 2.46, the equation of line P_1P_2 is given as

$$x(t)_{P_1P_2} = \frac{A}{T/2}t + A = A\left(1 + \frac{2}{T}t\right)$$

and the equation of line P_2P_3 is given as

$$x(t)_{P_2P_3} = \frac{-A}{T/2}t + A = A\left(1 - \frac{2}{T}t\right)$$

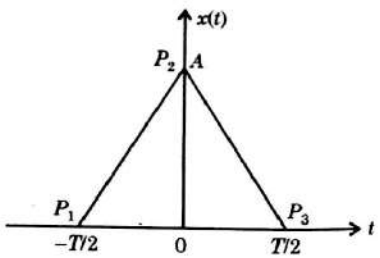


Fig. 2.46. *A triangular pulse.*

Hence, we have $$x(t) = \begin{cases} A\left(1 + \dfrac{2}{T}t\right) \text{ for } \dfrac{-T}{2} < t \le 0 \\ A\left(1 - \dfrac{2}{T}t\right) \text{ for } 0 \le t \le \dfrac{T}{2} \end{cases}$$

Now, the Fourier transform is expressed as

$$X(\omega) = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}dt = \int_{-T/2}^{T/2} x(t)e^{-j\omega t}dt = \int_{-T/2}^{0} x(t)e^{-j\omega t}dt + \int_{0}^{T/2} x(t)e^{-j\omega t}dt$$

or $$X(\omega) = \int_{-T/2}^{0} A\left(1 + \frac{2}{T}t\right)e^{-j\omega t}dt + \int_{0}^{T/2} A\left(1 - \frac{2}{T}t\right)e^{-j\omega t}dt$$

or $$X(\omega) = A\int_{-T/2}^{0} e^{-j\omega t}dt + A\int_{0}^{T/2} e^{-j\omega t}dt + \frac{2A}{T}\int_{-T/2}^{0} t.e^{-j\omega t}dt - \frac{2A}{T}\int_{0}^{T/2} t.e^{-j\omega t}dt$$

Now, changing t-to-t in the first and third intergrals in above equation, we get

$$X(\omega) = A\left[\int_{0}^{T/2} e^{j\omega t}dt + \int_{0}^{T/2} e^{-j\omega t}dt\right] - \frac{2A}{T}\left[\int_{0}^{T/2} t.e^{j\omega t}dt + \int_{0}^{T/2} t\cdot e^{-j\omega t}dt\right]$$

$$X(\omega) = A\int_{0}^{T/2} 2\cos\omega t\,dt - \frac{2A}{T}\int_{0}^{T/2} 2t\cos\omega t\,dt$$

or $$X(\omega) = 2A\frac{\sin\frac{\omega T}{2}}{\omega} - \frac{4A}{T}\left[\frac{T}{2}\cdot\frac{\sin\frac{\omega T}{2}}{\omega} - \int_{0}^{T/2}\frac{\sin\omega t}{\omega}dt\right]$$

or $$X(\omega) = 2A\frac{\sin\frac{\omega T}{2}}{\omega} - \frac{4A}{T}\left[\frac{T}{2}\cdot\frac{\sin\frac{\omega T}{2}}{\omega} + \left\{\frac{\sin\omega t}{\omega^2}\right\}_{0}^{T/2}\right]$$

or $$X(\omega) = 2A\cdot\frac{\sin\frac{\omega T}{2}}{\omega} - \frac{4A}{T}\left[\frac{T}{2}\cdot\frac{\sin\frac{\omega T}{2}}{\omega} + \frac{\cos\frac{\omega T}{2}}{\omega^2} - \frac{1}{\omega^2}\right]$$

or $$X(\omega) = \frac{2A}{\omega}\cdot\sin\frac{\omega T}{2} - \frac{2A}{\omega}\sin\frac{\omega T}{2} - \frac{4A}{T\omega^2}\left(\cos\frac{\omega T}{2} - 1\right) = \frac{4A}{T\omega^2}\left(1 - \cos\frac{\omega T}{2}\right)$$

or $$X(\omega) = \frac{4A}{T\omega^2}\left[2\sin^2\frac{\omega T}{4}\right] = \frac{8A}{T\omega^2}\sin^2\frac{\omega T}{4}$$

or $$X(\omega) = \frac{8A}{T\omega^2}\left(\frac{\omega T}{4}\right)^2\cdot\frac{\sin^2\frac{\omega T}{4}}{\left(\frac{\omega T}{4}\right)^2} = \frac{AT}{2}\sin c^2\left(\frac{\omega T}{4}\right)$$

Figure 2.47 shows the frequency spectrum of the single triangular pulse.

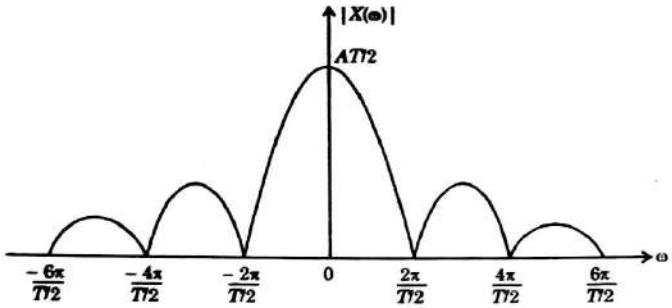


Fig. 2.47. *The frequency spectrum of a single triangular pulse.*

2.9. Properties of Fourier Transform

In this section, we shall discuss various properties of Fourier transform.

2.9.1. Time-Scaling Property

Time-scaling property states that the time compression of a signal results in its spectral expansion and time expansion of the signal results in its spectral compressions. Mathematically,

If $\quad x(t) \longleftrightarrow X(\omega)$

Then, for any real constant a,

$$x(at) \longleftrightarrow \frac{1}{|a|} \cdot X\left(\frac{\omega}{a}\right)$$

Proof: The general expression for Fourier transform is

$$X(\omega) = F[x(t)] = \int_{-\infty}^{\infty} x(t)\, e^{-j\omega t}\, dt$$

Now $\quad F[x(at)] = \int_{-\infty}^{\infty} x(at) e^{-j\omega t}\, dt$

Putting $\quad at = y,$

we have $\quad dt = \frac{dy}{a}$

Case (*i*): When a is a positive real constant

$$F[x(at)] = \int_{-\infty}^{\infty} x(y)\, e^{-j\left(\frac{\omega}{a}\right)y} \frac{dy}{a} = \frac{1}{a} \int_{-\infty}^{\infty} x(y)\, e^{-j\left(\frac{\omega}{a}\right)y}\, dy = \frac{1}{a} \cdot X\left(\frac{\omega}{a}\right)$$

Case (*ii*): When a is a negative real constant

$$F[x(at)] = -\frac{1}{a} \cdot X\left(\frac{\omega}{a}\right)$$

Combining two cases, we have

$$F[x(at)] = \frac{1}{|a|} \cdot X\left(\frac{\omega}{a}\right) \quad \text{or} \quad x(at) \longleftrightarrow \frac{1}{|a|} \cdot X\left(\frac{\omega}{a}\right)$$

Significance: *As stated earlier, according to scaling property, the time compression of a signal results in its spectral expansion and time expansion of the signal results in its spectral compressions.*

The function $x(at)$ represents the function $x(t)$ compressed in time-domain by a factor a. Similarly, a function $X\left(\frac{\omega}{a}\right)$ represents the function $X(\omega)$ expanded in frequency domain by the same factor a.

The compression in time by a factor a means that the signal is varying rapidly by the same factor a. Hence in frequency domain, the frequencies of its sinusoidal components must be increased by the factor a. This means that its frequency spectrum is expanded by the factor a.

Similarly, a signal expanded in time-domain varies more slowly and so the frequencies of its components are lowered. This means that its frequency spectrum is compressed.

As an example, consider a function $\sin \omega_0 t$ of frequency ω_0. When this function is compressed in the time-domain by a factor 2, it becomes $\sin 2\,\omega_0 t$. The frequency now becomes $2\,\omega_0$, which is doubled as compared to previous case, ω_0.

Therefore, the compression of a sine wave in the time domain is equivalent to expansion in the frequency domain.

2.9.2. Linearity Property

Linearity property states that the Fourier transform is linear. This means that

If $\quad x_1(t) \longleftrightarrow X_1(\omega)$

and $\quad x_2(t) \longleftrightarrow X_2(\omega)$

Then $a_1 x_1(t)\, a_2 x_2(t) \longleftrightarrow a_1 X_1(\omega) + a_2 X_2(\omega)$

2.9.3. Duality or Symmetry Property

Duality or symmetry, property states thats

If $\quad x(t) \longleftrightarrow X(\omega)$

Then $\quad X(t) \longleftrightarrow 2\pi x(-\omega)$

Hence, according to the duality or symmetry property:

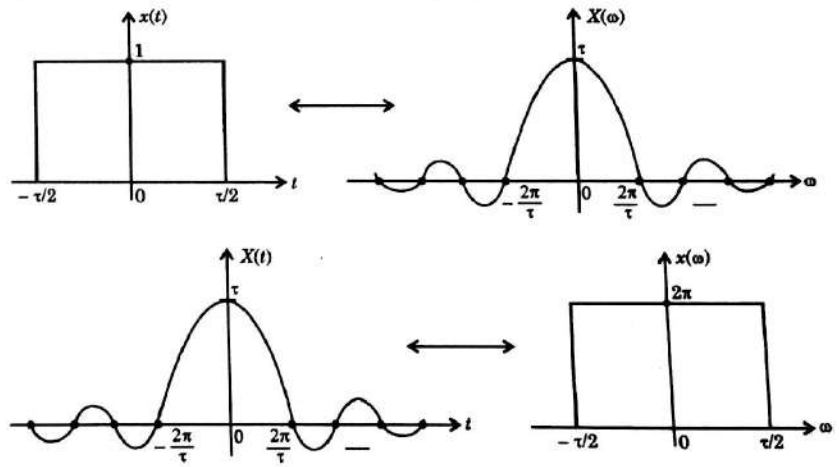


Fig. 2.48.

Proof: The general expression for inverse Fourier transform is

$$F^{-1}[X(\omega)] = x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(\omega)\, e^{j\omega t}\, d\omega$$

Therefore, $$x(-t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(\omega)\, e^{-j\omega t}\, d\omega$$

or $$2\pi x(-t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(\omega)\, e^{-j\omega t}\, d\omega$$

Since ω is a dummy variable, interchanging the variables t and ω, we have

$$2\pi x(-\omega) = \int_{-\infty}^{\infty} X(t)\, e^{-j\omega t}\, d\omega = F[X(t)]$$

or $$F[X(t)] = 2\pi x(-\omega)$$

or $$X(t) \longleftrightarrow 2\pi x(-\omega)$$

For an even function $x(-\omega) = x(\omega)$

Therefore, $$X(t) \longleftrightarrow 2\pi x(\omega)$$

Example 2.19. The Fourier transform $F[e^{-t}.u(t)]$ is equal to $\frac{1}{1+j2\pi f}$. Therefore, $F\left[\frac{1}{1+j2\pi t}\right]$ is equal to

(a) $e^{f}.u(f)$; (b) $e^{-f}.u(f)$; (c) $e^{f}.u(-f)$; (d) $e^{-f}, u(-f)$. *(GATE Examination, 2002)*

Solution: Using Duality property of Fourier transform, we have

If $$g(t) \leftrightarrow G(f)$$

then $$G(t) \leftrightarrow g(-f)$$

Therefore,

If $$e^{-t}\, u(t) \leftrightarrow \frac{1}{1+j2\pi f}$$

then $$\frac{1}{1+j2\pi t} \leftrightarrow e^{f}.u(-f)$$

Hence, option (c) is correct. **Ans.**

2.9.4. Time Shifting Property

Time shifting property states that a shift in the time-domain by an amount b is equivalent to multiplication by $e^{-j\omega b}$ in the frequency domain. This means that magnitude sepctrum $|X(\omega)|$ remains inchanged but phase spectrum $\theta(\omega)$ is changed by $-\omega b$. A shift in time domain does not change the magnitude of a frequency component. A shift of b at a frequency ω is equivalent to a phase shift of ωb. Mathematically,

If $$x(t) \longleftrightarrow X(\omega)$$

Then $$x(t-b) \longleftrightarrow X(\omega)\cdot e^{-j\omega b}$$

Proof: The general expression for Fourier transform is

$$X(\omega) = F[x(t)] = \int_{-\infty}^{\infty} x(t)\, e^{-j\omega t}\, dt$$

And $$F[x(t-b)] = \int_{-\infty}^{\infty} x(t-b)\, e^{-j\omega t}\, dt$$

Putting $t - b = y$, so that $dt = dy$

Hence $$F[x(t-b)] = \int_{-\infty}^{\infty} x(y)\, e^{-j\omega(b+y)}\, dy = \int_{-\infty}^{\infty} x(y)\, e^{-j\omega b}\cdot e^{-j\omega y}\, dy$$

or $$F[x(t-b)] = e^{-j\omega b}\int_{-\infty}^{\infty} x(y)\, e^{-j\omega y}\, dy$$

Since y is a dummy variable, we have

$$F[x(t-b)] = e^{-j\omega b}\cdot X(\omega) = X(\omega)\cdot e^{-j\omega b}$$

or $$x(t-b) \longleftrightarrow X(\omega)\cdot e^{-j\omega b}$$

2.9.5. Frequency Shifting Property

Frequency shifting property states that the multiplication of a function $x(t)$ by $e^{j\omega_0 t}$ is equivalent to shifting its Fourier treansform $X(\omega)$ in the positive direction by an amount ω_0. This means that the spectrum $X(\omega)$ is translated by an amount ω_0. Hence, this property is often called as **frequency translated theorem**. Mathematically,

If $$x(t) \longleftrightarrow X(\omega)$$

Then $$e^{j\omega_0 t}\cdot x(t) \longleftrightarrow X(\omega-\omega_0)$$

Proof: The general expression for Fourier transform is

$$X(\omega) = F[x(t)] = \int_{-\infty}^{\infty} x(t)\, e^{-j\omega t}\, dt$$

Now, $$F[e^{j\omega_0 t}\cdot x(t)] = \int_{-\infty}^{\infty} e^{j\omega_0 t}\cdot x(t)\cdot e^{-j\omega t}\, dt$$

or $$F[e^{j\omega_0 t}\cdot x(t)] = \int_{-\infty}^{\infty} x(t)\cdot e^{-j(\omega-\omega_0)t}\, dt$$

or $$F[e^{j\omega_0 t}\cdot x(t)] = X(\omega-\omega_0)$$

or $$e^{j\omega_0 t}\cdot x(t) \longleftrightarrow X(\omega-\omega_0)$$

2.9.6. Time Differentiation Property

Time differentiation property states that the differentiation of a function $x(t)$ in the time domain is equivalent to multiplication of its Fourier transform by a factor $j\omega$. Mathematically,

If $$x(t) \longleftrightarrow X(\omega)$$

then $$\frac{dx(t)}{dt} \longleftrightarrow j\omega X(\omega)$$

Proof: The general expression for inverse Fourier transform is

$$F^{-1}[X(\omega)] = x(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(\omega)\, e^{j\omega t}\, d\omega$$

Taking differentiation, we have

$$\frac{dx(t)}{dt} = \frac{1}{2\pi}\cdot\frac{d}{dt}\left[\int_{-\infty}^{\infty} X(\omega)\, e^{j\omega t}\, d\omega\right]$$

Interchanging the order of differentiation and integration, we have

$$\frac{dx(t)}{dt} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\frac{d}{dt}[X(\omega)\, e^{j\omega t}]\, d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} j\omega\cdot X(\omega)\cdot e^{j\omega t} d\omega$$

or $$\frac{dx(t)}{dt} = F^{-1}[j\omega\, X(\omega)]$$

or $$F\left[\frac{dx(t)}{dt}\right] = j\omega\cdot X(\omega)$$

or $$\frac{dx(t)}{dt} \longleftrightarrow j\omega\cdot X(\omega)$$ **Hence Proved.**

Example 2.20. The Fourier transform $G(\omega)$ of the a signal $g(t)$ in figure 2.49(a), is given as $G(\omega) = \frac{1}{\omega^2}(e^{j\omega} - j\omega e^{j\omega} - 1)$. Using this information, and the time-shifting and time-scaling properties, determine the Fourier transform of signals in figure 2.49(b), (c) and (d).

(GATE Examination-2001)

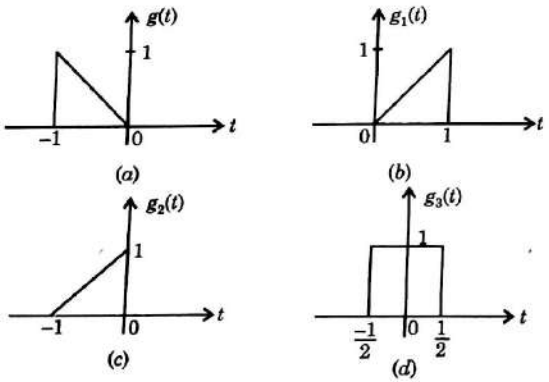


Fig. 2.49.

Solution: We know that the Fourier transform of a signal $x(t)$ is given by

$$X(\omega) = \int_{-\infty}^{\infty} x(t)\, e^{-j\omega t} dt$$

For signal $x_1(t) = x(-t)$, we have

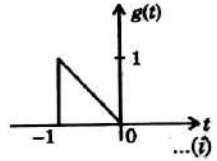
@sTc(small team creation)

$$X_1(\omega) = \int_{-\infty}^{\infty} x(-t).e^{-j\omega t} dt$$

Substituting $p = -t$, we have

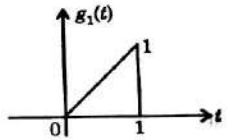
$$X_1(\omega) = -\int_{+\infty}^{-\infty} x(p)\, e^{+j\omega p} dp$$

or $$X_1(\omega) = \int_{-\infty}^{\infty} x(p)\, e^{j\omega p}\, dp = X(-\omega) = X^*(\omega) \quad \text{...(i)}$$



For time-shifting signal, $x_2(t) = x_1(t+1)$, we have

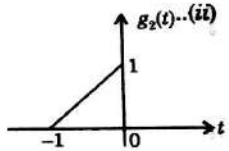
$$X_2(\omega) = \int_{-\infty}^{\infty} x_1(t+1)\, e^{-j\omega t} dt$$



Substituting $t + 1 = p$, we get

$$X_2(\omega) = +\int_{-\infty}^{\infty} x_1(p)\, e^{-j\omega p}.e^{j\omega}\, dp$$

or $$X_2(\omega) = e^{j\omega}\int_{-\infty}^{\infty} x_1(p)\, e^{-j\omega p}\, dp = X_1(j\omega)\, e^{j\omega} \quad \text{...(ii)}$$



Now, given that

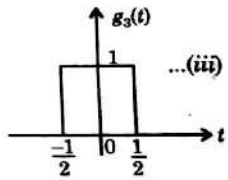
For signal $g(t)$, the Fourier transform

$$G(\omega) = \frac{1}{\omega^2}(e^{j\omega} - j\omega\, e^{j\omega} - 1)$$

$$g_1(t) = g(-t)$$

Therefore, using equation (i) $G_1(\omega) = G(-\omega)$

or $$G_1(\omega) = \frac{1}{\omega^2}(e^{-j\omega} + j\omega\, e^{-j\omega} - 1) \quad \text{...(iii)}$$



Then $$G_2(\omega) = G_1(\omega)\, e^{j\omega}$$

Now, using equations (ii) and (iii), we have

$$G_2(\omega) = \frac{1}{\omega^2}(1 + j\omega - e^{j\omega})$$

and $$G_3(\omega) = \int_{-\infty}^{\infty} g_3(t)\, e^{-j\omega t}\, dt$$

$$G_3(\omega) = +\int_{-1/2}^{1/2} e^{-j\omega t}\, dt = \text{sinc}\,(f)$$

or $$G_3(\omega) = \frac{\sin \pi f}{\pi f}$$ **Ans.**

2.10. Fourier Transform of Periodic Functions

Till now, we have discussed that a given function may be analyzed over an entire interval using Fourier series if the function is periodic and by using Fourier transform if the function is non-periodic.

But we can also find the transform of periodic function. This means that Fourier transform may be used as a universal mathematical tool to analyzed both periodic and non-periodic waveforms over the entire interval. Fourier trnasform of periodic functions may be found using the concept of impulse function.

Let us consider a periodic function $x(t)$. We have to find the Fourier transform of $x(t)$.

The periodic function $x(t)$ may be expressed in terms of complex exponential Fourier series as

$$x(t) = \sum_{n=-\infty}^{\infty} C_n \cdot e^{jn\omega_0 t} \qquad ...(2.40)$$

where $\omega_0 = \dfrac{2\pi}{T}$ = fundamental frequency

Taking Fourier transform of both sides

$$F[x(t)] = F\left[\sum_{n=-\infty}^{\infty} C_n \cdot e^{jn\omega_0 t}\right] = \sum_{n=-\infty}^{\infty} C_n \cdot F\left[1 \cdot e^{jn\omega_0 t}\right] \qquad ...(2.41)$$

Using frequency shifting theorem, we can write

$$F[1 \cdot e^{jn\omega_0 t}] = 2\pi\,\delta(\omega - n\omega_0)$$

Hence, $$F[x(t)] = \sum_{n=-\infty}^{\infty} C_n \cdot 2\pi\,\delta(\omega - n\omega_0) = 2\pi \sum_{n=-\infty}^{\infty} C_n \cdot \delta(\omega - n\omega_0) \qquad ...(2.42)$$

Hence, the Fourier transform of a periodic function consists of a train of equally spaced impulses. These impulses are located at the harmonic frequencies of the signal and the strength or area of each impulse is given by $2\pi C_n$. This seems logical because Fourier series analysis of periodic functions provides the discrete lines (discrete frequency spectrum) located at harmonic frequencies. Here, also we have got the discrete spectrum by Fourier transform which generally provides a continuous spectrum.

2.11. Convolution

Convolution is a mathematical operation which is used to express the input/output relationship in a linear time invariant system.

Mathematically, the convolution of two time functions $x_1(t)$ and $x_2(t)$ may be defined by the following expression

$$x_1(t) \otimes x_2(t) = \int_{-\infty}^{\infty} x_1(\tau)\, x_2(t-\tau)\, d\tau^*$$

2.11.1. The Convolution Theorems

In this section, we will discuss the convolution theorems associated with Fourier transform. There are two convolution theorems:

(*i*) Time convolution theorem

(*ii*) Frequency convolution theorem

* An encircles multiplication notation ⊗ is used to represent the convolution of two time functions.

2.11.2. Time Convolution Theorem *(M.D. University, Haryana 2001)*

The time convolution theorem states that convolution in time domain is equivalent to multiplication of their spectra in freaquency domain. Mathematically,

If $x_1(t) \longleftrightarrow X_1(\omega)$

and $x_2(t) \longleftrightarrow X_2(\omega)$

then $x_1(t) \otimes x_2(t) \longleftrightarrow X_1(\omega)\, X_2(\omega)$

Proof: We know that Fourier transform is expressed as

$$F[x(t)] = X(\omega) = \int_{-\infty}^{\infty} x(t)\, e^{-j\omega t}\, dt$$

Therefore, $$F[x_1(t) \otimes x_2(t)] = \int_{-\infty}^{\infty} [x_1(t) \otimes x_2(t)]\, e^{-j\omega t}\, dt$$

But we know that $$x_1(t) \otimes x_2(t) = \int_{-\infty}^{\infty} [x_1(\tau)\, x_2(t-\tau)]\, dt$$

Therefore, $$F[x_1(t) \otimes x_2(t)] = \int_{-\infty}^{\infty}\left[\int_{-\infty}^{\infty} x_1(\tau)\, x_2(t-\tau)\, d\tau\right] e^{-j\omega t}\, dt$$

Rearranging the terms, we get

$$F[x_1(t) \otimes x_2(t)] = \int_{-\infty}^{\infty} x_1(\tau)\left[\int_{-\infty}^{\infty} e^{-j\omega t}\, x_2(t-\tau)\, dt\right] d\tau$$

Using time-shifting property, the term inside the bracket may be written as

$$\int_{-\infty}^{\infty} e^{-j\omega t}\, x_2(t-\tau)\, dt = X_2(\omega)\, e^{-j\omega\tau}$$

Hence $$F[x_1(t) \otimes x_2(t)] = \int_{-\infty}^{\infty} x_1(\tau) \cdot X_2(\omega)\, e^{-j\omega\tau}\, d\tau$$

$$F[x_1(t) \otimes x_2(t)] = X_2(\omega) \int_{-\infty}^{\infty} x_1(\tau)\, e^{-j\omega\tau}\, d\tau$$

or $$F[x_1(t) \otimes x_2(t)] = X_2(\omega) \cdot X_1(\omega)$$

or $$[x_1(t) \otimes x_2(t)] \longleftrightarrow X_1(\omega)\, X_2(\omega)$$

2.11.3. Frequency Convolution Theorem

Frequency convolution theorem states that multiplication of two functions in time domain is equivalent to convolution of their spectra in the frequency domain. Mathematically,

If $x_1(t) \longleftrightarrow X_1(\omega)$

and $x_2(t) \longleftrightarrow X_2(\omega)$

then $2\pi\, x_1(t)\, x_2(t) \longleftrightarrow X_1(\omega) \otimes X_2(\omega)$

Example 2.21. Prove that the convolution of a function $x(t)$ with an unit impulse function results the function itself.

Solution: Time-convolution theorem states that

If $x_1(t) \longleftrightarrow X_1(\omega)$

and $x_2(t) \longleftrightarrow X_2(\omega)$

then $x_1(t) \otimes x_2(t) \longleftrightarrow X_1(\omega)\, X_2(\omega)$

or $F[x_1(t) \otimes x_2(t)] = X_1(\omega)\, X_2(\omega)$

or $x_1(t) \otimes x_2(t) = F^{-1}[X_1(\omega)\, X_2(\omega)]$ ...(i)

Now let $x_1(t) = x(t)$ and $x_2(t) = \delta(t)$

So that $X_1(\omega) = X(\omega)$ and $X_2(\omega) = 1$

Substituting, all these values in equation (i), we get

$$x(t) \otimes \delta(t) = F^{-1}[X(\omega) \cdot 1]$$

$$x(t) \otimes \delta(t) = F^{-1}[X(\omega)] = F^{-1}[F\{x(t)\}] = x(t) \quad \textbf{Hence Proved.}$$

2.12. Energy Signal

As stated earlier, an energy signal is one which has finite energy and zero average power.

Hence, $x(t)$ is an energy signal if

$$0 < E < \infty \quad \text{and} \quad P = 0$$

where, E is the energy and P is the power of the signal $x(t)$.

The signal energy must be finite for it to be meaningful. For the energy to be finite, the signal amplitude $x(t)$ must go to zero $[x(t) \to 0]$ as $|t| \to \infty$.

Therefore, almost all the practical non-periodic signals which are defined over finite-time (also called time-limited signals) are energy signals.

For continuous-time signal, the energy of a single is expressed as

$$E = \int_{-\infty}^{\infty} x^2(t)\, dt \qquad ...(2.43)$$

2.12.1. Parseval's Theorem for Energy Signals

The Parseval's theorem states that the energy of a signal may be obtained with the help of its Fourier transform. This means that according to Parseval's theorem, we can find the energy of signal without knowing its time-domain.

Proof: Let there be a signal $x(t)$ with its Fourier transform $X(\omega)$.

The energy E of signal $x(t)$ is expressed as

$$E = \int_{-\infty}^{\infty} x^2(t)\, dt = \int_{-\infty}^{\infty} x(t) \cdot x(t)\, dt \qquad ...(2.44)$$

We know that inverse Fourier transform $x(t)$ is wirtten as

$$x(t) = F^{-1}[X(\omega)] = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega)\, e^{j\omega t}\, d\omega \qquad ...(2.45)$$

Replacing one $x(t)$ in equation (2.44) from equation (2.45), we get

$$E = \int_{-\infty}^{\infty} x(t) \left[\frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega)\, e^{j\omega t}\, d\omega \right] dt$$

Interchanging the order of integration, we get

$$E = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega) \left[\int_{-\infty}^{\infty} x(t)\, e^{j\omega t}\, dt \right] d\omega$$

The integral under the bracket is equal to $X(-j\omega)$

Since $$X(\omega) = \int_{-\infty}^{\infty} x(t)\, e^{-j\omega t}\, dt$$

So $$X(-\omega) = \int_{-\infty}^{\infty} x(t)\, e^{j\omega t}\, dt$$

Therefore $$E = \frac{1}{2\pi} \int_{-\infty}^{\infty} X(\omega)\, X(-\omega)\, d\omega \qquad ...(2.46)$$

We know that for a real signal $x(t)$, the Fourier transform $X(\omega)$ and $X(-\omega)$ are complex conjugates, *i.e.*,

$$X(-\omega) = X^*(\omega)$$

So that $$X(\omega)\, X(-\omega) = X(\omega)\, X^*(\omega)$$

$$X(\omega)\, X(-\omega) = |X(\omega)|^2$$

Putting this values of $\{X(\omega)\, X(-\omega)\}$ in equation (2.46), we get

$$E = \frac{1}{2\pi} \int_{-\infty}^{\infty} |X(\omega)|^2\, d\omega = \int_{-\infty}^{\infty} |X(\omega)|^2\, df$$

Above equation is known as Parseval's theorem for energy signals.

Example 2.22. Verify Parseval's theorem for the energy signal $x(t) = e^{-at} \cdot u(t)$ $(a > 0)$.

Solution: We know that energy of a signal is expressed as

$$E = \int_{-\infty}^{\infty} x^2(t)\, dt = \int_{-\infty}^{0} \left[e^{-at} \cdot u(t)\right]^2 dt + \int_{0}^{\infty} \left[e^{-at} \cdot u(t)\right]^2 dt$$

But we know that $u(t)$ exists only for positive values of t.

Therefore, we have

$$E = \int_{-\infty}^{0} \left[e^{-at} \cdot 0\right]^2 dt + \int_{0}^{\infty} \left[e^{-at} \cdot 1\right]^2 dt = 0 + \int_{0}^{\infty} e^{-2at}\, dt = \left[\frac{e^{-2at}}{-2a}\right]_0^{\infty}$$

or $$E = -\frac{1}{2a}\left[e^{-\infty} - e^{0}\right] = -\frac{1}{2a}[0 - 1] = \frac{1}{2a} \qquad ...(i)$$

Now according to Parseval's theorem, the energy of a signal is expressed as

$$E = \frac{1}{2\pi} \int_{-\infty}^{\infty} |X(\omega)|^2\, d\omega \qquad ...(ii)$$

Also $$X(\omega) = F[x(t)] = \int_{-\infty}^{\infty} x(t)\, e^{-j\omega t}\, dt = \int_{-\infty}^{\infty} e^{-at} \cdot u(t) \cdot e^{-j\omega t}\, dt$$

or $$X(\omega) = \int_{-\infty}^{0} e^{-at} \cdot 0 \cdot e^{-j\omega t}\, dt + \int_{0}^{\infty} e^{-at} \cdot 1 \cdot e^{-j\omega t}\, dt$$

or $$X(\omega) = 0 + \int_0^{\infty} e^{-(a+j\omega)t}\,dt = \left[\frac{e^{-(a+j\omega)t}}{-(a+j\omega)}\right]_0^{\infty} = -\frac{1}{a+j\omega}\left[e^{-\infty} - e^0\right] = -\frac{1}{a+j\omega}[0-1]$$

or $$X(\omega) = \frac{1}{a+j\omega} = \frac{1}{a+j\omega} \times \frac{(a-j\omega)}{(a-j\omega)} = \frac{a-j\omega}{a^2+\omega^2} = \frac{a}{a^2+\omega^2} - j\cdot\frac{\omega}{a^2+\omega^2}$$

or $$X(\omega) = \frac{1}{\sqrt{a^2+\omega^2}}$$

Putting the value of $|X(\omega)|$ in equation (*ii*), we get

$$E = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{1}{a^2+\omega^2}\,d\omega = \frac{1}{2\pi}\cdot\frac{1}{a}\tan^{-1}\left[\frac{\omega}{a}\right]_{-\infty}^{\infty}$$

or $$E = \frac{1}{2\pi a}\cdot\left[\frac{\pi}{2} - \left(-\frac{\pi}{2}\right)\right] = \frac{1}{2\pi a}\cdot\left[\frac{\pi}{2} + \frac{\pi}{2}\right] = \frac{1}{2\pi a}\cdot\pi = \frac{1}{2a} \qquad ...(iii)$$

Hence, from equation (i) and equation (iii), the energy is same in both cases. Therefore, the Parseval's theorem is verified. **Hence Proved.**

Example 2.23. Determine the signal energy and signal power for the following signals:

(*i*) $x(t) = e^{-3|t|}$

(*ii*) $x(t) = e^{-3t}$.

Solution: (*i*) The given signal is

$$x(t) = e^{-3|t|}$$

We know that the energy of the signal $x(t)$ is given by

$$E = \int_{-\infty}^{\infty} |x(t)|^2\,dt = \int_{-\infty}^{\infty} [e^{-3|t|}]^2\,dt = \int_{-\infty} e^{6t}dt + \int_0^{\infty} e^{-6t}dt$$

Simplifying, we get

$$E = 2\int_0^{\infty} e^{-6t}dt = \frac{-2}{6}[e^{-6t}]_0^{\infty} = \frac{-2}{6}[0-1] = \frac{1}{3}$$

The signal power $P = 0$ since E is a finite value.

Therefore, the given signal $x(t)$ is an energy signal. Ans.

(*ii*) The given signal is

$$x(t) = e^{-3t}$$

We know that the energy of the signal $x(t)$ is given by

$$E = \int_{-T}^{T} (e^{-3t})^2\,dt = \int_{-T}^{T} e^{-6t}dt = -\frac{1}{6}[e^{-6T} - e^{6T}]$$

Now, as $T \to \infty$, E approaches infinity, its average power is

$$P = \lim_{T\to\infty}\frac{1}{2T}\cdot E = \lim_{T\to\infty}\cdot\frac{e^{6T} - e^{-6T}}{8T}$$

or $$P = \infty$$

[Using L'Hospital's rule]

Therefore, it is clear that e^{-3t} is neither an energy signal nor a power signal. **Ans.**

Example 2.24. Calculate the signal energy for the signal expressed as under:

$$x(t) = e^{-4t}.u(t)$$

where $$u(t) = \begin{cases} 1 & \text{for } t \geq 0 \\ 0 & \text{for } t < 0 \end{cases}$$ *(Anna University, Chennai-1999)*

Solution: The given signal is

$$x(t) = e^{-4t}.u(t)$$

Since the value of the given signal is zero between $-T$ to 0. Therefore, the limits must be applied between 0 to T instead of $-T$ to T.

Thus, we have

$$E_x = \lim_{T\to\infty}\int_{-T}^{T} |x(t)|^2\,dt = \lim_{T\to\infty}\int_0^{T} |e^{-4t}|^2\,dt$$

or $$E = \lim_{T\to\infty}\int_0^{T} e^{-8t}dt = \lim_{T\to\infty}\left[\frac{e^{-8t}}{-8}\right]_0^{T} = \lim_{T\to\infty}\left[\frac{e^{-8T}}{-8} + \frac{1}{8}\right]$$

Simplifying, we get

$$E = \frac{1}{8}$$

and $$P_x = \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} |x(t)|^2\,dt$$

or $$P_x = \lim_{T\to\infty}\frac{1}{2T}\left[\frac{e^{-8t}}{-8}\right]_{-T}^{T} = \lim_{T\to\infty}\frac{(e^{-8T}-1)}{-16T} = 0$$

Here, it may be noted that since the signal energy E_x is finite and the signal power $P_x = 0$, therefore, the given signal is an energy signal. **Ans.**

Example 2.25. Determine the signal energy and signal power for the following complex-valued signal and determine whether it is an energy signal or a power signal

$$x(t) = A\,e^{j2\pi\alpha t}$$

Solution: Here, the signal period is $T = \frac{1}{\alpha}$.

Since $x(t)$ is periodic signal, therefore it cannot be an energy signal. Hence, the power of a power signal is evaluated as

$$P_x = \frac{1}{T}\int_{t_1}^{t_1+T} |x(t)|^2\,dt = \alpha\int_{t_1}^{t_1+\left(\frac{1}{\alpha}\right)} |Ae^{2\pi\alpha t}|^2\,dt$$

$$P_x = \alpha\int_{t_1}^{t_1+\left(\frac{1}{\alpha}\right)} A^2dt = \alpha\cdot[A^2t]_{t_1}^{t_1+\frac{1}{\alpha}}$$

Simplifying, we get $P_x = A^2$

Since, the given signal has finite power, therefore, it is a power signal and $E_x = \infty$.

Example 2.26. A signal $f(t)$ has energy E. Calculate the energy of the signal $f(2t)$.
(GATE examination-2001)

Solution: We know that energy of a signal $f(t)$ is given by

$$E = \int_{-\infty}^{\infty} |f(t)|^2 \, dt$$

Now, for signal $f(2t)$, we have

$$E' = \int_{-\infty}^{\infty} |f(2t)|^2 \, dt$$

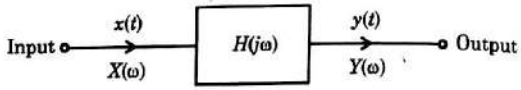


Fig. 2.50. *Low pass filter (LPF).*

Substituting $2t = z$, we get

$$E' = \frac{1}{2} \int_{-\infty}^{\infty} |f(z)|^2 \, dz = \frac{1}{2} E$$

Thus, we have $\quad E' = \frac{1}{2} E \quad$ **Ans.**

2.12.2. Energy Spectral Density

Let us consider a signal $x(t)$ which is applied to an ideal low pass filter as shown in figure 2.50.

Figure 2.51 shows the graph of transfer function $H(\omega)$ of the ideal low pass filter.

The response or output of a system is expressed as

$$Y(\omega) = X(\omega) \, H(\omega) \qquad ...(2.47)$$

Here $\quad X(\omega)$ = Fourier transform of $x(t)$

and $\quad Y(\omega)$ = Fourier transform of $y(t)$

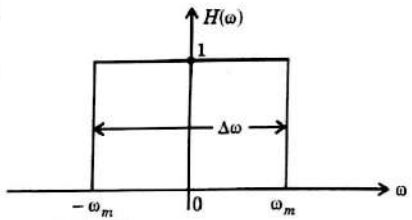


Fig. 2.51. *Transfer function of ideal low-pass filter.*

Also, the energy E_0 of the output signal $y(t)$ may be expressed as (using Parseval's theorem)

$$E_0 = \frac{1}{2\pi} \int_{-\infty}^{\infty} |Y(\omega)|^2 \, d\omega$$

or

$$E_0 = \frac{1}{2\pi} \int_{-\infty}^{\infty} |H(\omega) \cdot X(\omega)|^2 \, d\omega \qquad [\because Y(\omega) = H(\omega) \, X(\omega)]$$

From figure 2.51, it may be observed that $H(\omega) = 0$ for all the frequencies except for the narrow band $-\omega_m$ to ω_m for which it is unity.

Therefore,

$$E_0 = \frac{1}{2\pi} \int_{-w_m}^{w_m} |1 \cdot X(\omega)|^2 \, d\omega$$

or

$$E_0 = \frac{1}{2\pi} \int_{-\omega_m}^{\omega_m} |X(\omega)|^2 \, d\omega \qquad ...(2.48)$$

Further, we may assume that Fourier transform $X(\omega)$ is constant with frequency for a narrow band $-\omega_m$ to ω_m. Therefore, the energy of the signal over this narrow-band $\Delta w = 2w_m$ will be

$$E_0 = \frac{1}{2\pi} \cdot |X(\omega)|^2 \int_{-\omega_m}^{\omega_m} 1 \cdot d\omega = \frac{1}{2\pi} \cdot |X(\omega)|^2 (2\omega_m)$$

Putting $2\omega_m = \Delta\omega$, we get

$$E_0 = \frac{1}{2\pi} \cdot |X(\omega)|^2 (\Delta\omega) = |X(\omega)|^2 (\Delta f) \qquad ...(2.49)$$

From above equation, it is clear that E_0 represents the contribution of energy due to bandwidth $(\Delta\omega)$ of the signal including negative frequencies.

Therefore, energy contribution per unit bandwidth will be

$$\frac{E_0}{\Delta f} = |X(\omega)|^2 \qquad ...(2.49a)$$

Hence, $|X(\omega)|^2$ represents energy per unit bandwidth and is known as **Energy spectral Density** or **Energy Density spectrum.**

It is generally denoted by $\psi(\omega)$.

Hence $\quad \psi(\omega) = |X(\omega)|^2 \qquad ...(2.50)$

Since $\psi(\omega)$ represents the energy per unit bandwidth, hence the total energy of signal may be obtained by integrating $\psi(\omega)$ over the bandwidth of the signal.

Thus,

$$\therefore \qquad E = \frac{1}{2\pi} \int_{-\infty}^{\infty} |X(\omega)|^2 \, d\omega$$

So

$$E = \frac{1}{2\pi} \int_{-\infty}^{\infty} \psi(\omega) \, d\omega$$

However, for real signals,

$$|X(\omega)|^2 = |X(-\omega)|^2$$

This means that spectrum is symmetrical and the contribution by negative and positive frequencies will be same.

Therefore,

$$E = \frac{1}{\pi} \int_{0}^{\infty} \psi(\omega) \, d\omega$$

Now, we can find the relationship between energy-densities of input and output (response) as under:

Since, we know that

$$Y(\omega) = H(\omega) \cdot X(\omega)$$

therefore, $\quad |Y(\omega)|^2 = |H(\omega) \, X(\omega)|^2 = |H(\omega)|^2 \, |X(\omega)|^2 \qquad ...(2.51)$

Now, let $\psi_y(\omega)$ be the energy spectral density of output $y(t)$ and $\psi_x(\omega)$ be the energy spectral density of input $x(t)$, then

$$\psi_y(\omega) = |Y(\omega)|^2 \qquad ...(2.52)$$

and $\quad \psi_x(\omega) = |X(\omega)|^2 \qquad ...(2.53)$

Using equations (2.49), (2.49a) and (2.50), we get

$$\psi_y(\omega) = |H(\omega)|^2 \, \psi_x(\omega)$$

This is the relationship between input and output energy spectral densities.

2.12.3. Properties of Energy Spectral Density Function

Following are the properties of the Energy spectral density function:

Property 1: The total area under the energy spectral density function is equal to the total energy of that signal.

Mathematically, $$E = \int_{-\infty}^{\infty} \psi(f)\, df$$

Property 2: If $x(t)$ is input to a linear time invariant (LTI) system with transfer function $H(\omega)$, then input and output energy spectral density functions are related as

$$\psi_0(\omega) = |H(\omega)|^2\, \psi_i(\omega)$$

where $\psi_0(\omega)$ = output energy spectral density function

$\psi_i(\omega)$ = input energy spectral density function

$|H(\omega)|^2$ = energy gain at frequency ω

Proof: Let us consider that input $x(t)$ is applied to a linear time invariant (LTI) system with transfer function $H(\omega)$.

Let the ouput be denoted by $y(t)$.

Then in frequency-domain, we may write

$$Y(\omega) = H(\omega)\, X(\omega)$$

Squaring the magnitudes of both sides, we get

$$|Y(\omega)|^2 = |H(\omega)|^2 |X(\omega)|^2$$

But, we know that $|X(\omega)|^2 = \psi_i(\omega)$

and $|Y(\omega)|^2 = \psi_0(\omega)$

Therefore, $\psi_0(\omega) = |H(\omega)|^2 \cdot \psi_i(\omega)$

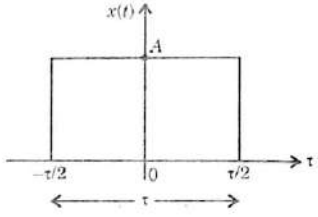


Fig. 2.52. *A gate function*

Property 3: The autocorrelation function $R(\tau)$ and energy spectral density function $\psi(\omega)$ form a Fourier transform pair.

Mathematically.

$$R(\tau) \longleftrightarrow \psi(\omega)$$

Example 2.27. Obtain the energy spectrum density of a Gate function of width τ and amplitude A.

Solution: A gate function of width τ and amplitude A is shown in figure 2.52.

The expression for this gate function is expressed as

$$x(t) = \begin{cases} A & |t| < \tau/2 \\ 0 & \text{otherwise} \end{cases}$$

The Fourier transform is expressed as

$$X(\omega) = \int_{-\infty}^{\infty} x(t)\, e^{-j\omega t}\, dt$$

For a gate function, we have

$$X(\omega) = \int_{-\tau/2}^{\tau/2} A\, e^{-j\omega t}\, dt = \frac{A}{j\omega}\left(e^{\frac{j\omega\tau}{2}} - e^{-\frac{j\omega\tau}{2}}\right) = A\tau \frac{\sin\left(\frac{\omega\tau}{2}\right)}{\left(\frac{\omega\tau}{2}\right)} = A\tau\, Sa\left(\frac{\omega\tau}{2}\right)$$

which is a sampling function.

Its graph is shown in figure 2.53.

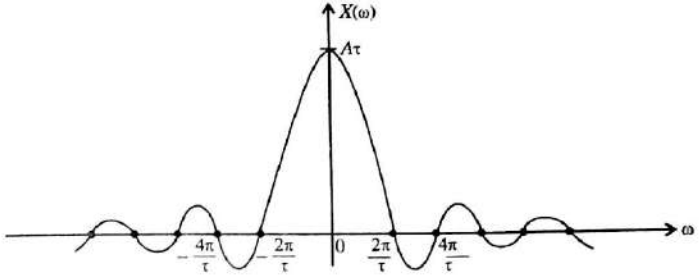


Fig. 2.53. *Sampling function $X(\omega) = A\tau\, S_a(\omega\tau/2)$.*

We know that the energy spectrum density is given as

$$\Psi(\omega) = |X(\omega)|^2 = A^2\tau^2 Sa^2\left(\frac{\omega\tau}{2}\right)$$

The energy spectrum density $\Psi(\omega)$ is shown in figure 2.54. It may be noted that $\Psi(\omega)$ being a square function has no negative value anywhere in the frequency spectrum.

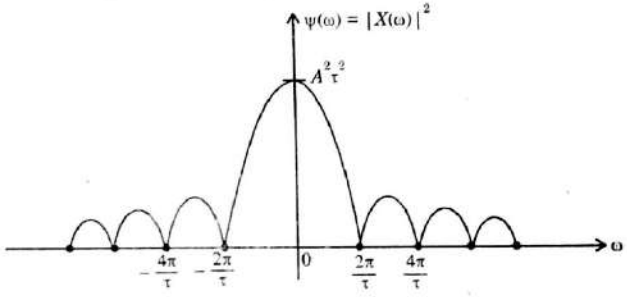


Fig. 2.54. *Energy spectrum density $\psi(j\omega)$.*

2.13. Power Signals

As stated earlier, a power signal is one which has finite average power and infinite energy.

Hence, $x(t)$ is a power signal if

$$0 < P < \infty$$

and $$E = \infty$$

where, P is the average power and E is the energy of the signal.

Almost all the practical periodic signals are power signals since their average power is finite and non-zero.

For continuous-time signal, the average power P of a signal is defined as

$$P = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x^2(t)\, dt \qquad \text{...(2.54)}$$

2.13.1. Parseval's Theorem for Power Signals

Parseval's theorem for power signals states that the power of a signal may be defined in terms of its Fourier series coefficients.

Proof: Let their be a function $x(t)$.

We know that

$$|x(t)|^2 = x(t)\, x^*(t) \qquad ...(2.55)$$

Here, $x^*(t)$ is the complex conjugate of the function $x(t)$.
The power of the signal $x(t)$ over a cycle is expressed as

$$P = \frac{1}{T}\int_{-T/2}^{T/2} |x(t)|^2\, dt = \frac{1}{T}\int_{-T/2}^{T/2} x(t)\, x^*(t)\, dt \qquad ...(2.56)$$

Replacing $x(t)$ by its complex Fourier exponential series, we get

$$P = \frac{1}{T}\int_{-T/2}^{T/2} x^*(t)\left[\sum_{n=-\infty}^{\infty} C_n\, e^{jn\omega_0 t}\right] dt \qquad ...(2.57)$$

Here $\qquad w_0 = \dfrac{2\pi}{T}$

Interchanging the order of integration and summation, we get

$$P = \frac{1}{T}\sum_{n=-\infty}^{\infty} C_n \cdot \int_{-T/2}^{T/2} x^*(t)\, e^{jn\omega_0 t}\, dt \qquad ...(2.58)$$

Since, $\qquad C_n = \dfrac{1}{T}\displaystyle\int_{-T/2}^{T/2} x(t)\, e^{-jn\omega_0 t}\, dt$

therefore $\qquad C_n^* = \dfrac{1}{T}\displaystyle\int_{-T/2}^{T/2} x^*(t)\, e^{jn\omega_0 t}\, dt \qquad ...(2.59)$

Using equations (2.58) and (2.59), we have

$$P = \frac{1}{T}\sum_{n=-\infty}^{\infty} C_n \cdot (T C_n^*) = \sum_{n=-\infty}^{\infty} C_n \cdot C_n^* = \sum_{n=-\infty}^{\infty} |C_n|^2 \qquad ...(2.60)$$

The above equation is known as Parseval's theorem for power signals.

2.13.2. The Power Spectral Density

The expression for power spectral density may be derived by assuming the power signal as a limiting case of an energy signal.

Let us consider a power signal $x(t)$ which is extended to infinity (figure 2.55).

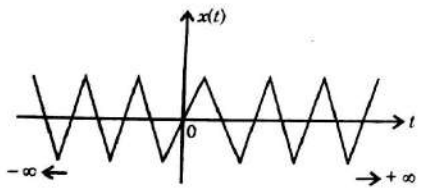


Fig. 2.55.

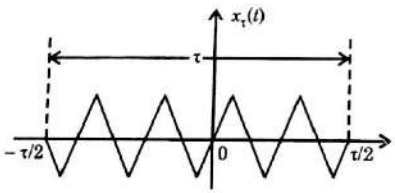


Fig. 2.56.

Let us terminate, this signal such that it is zero outside the interval, $\pm\,\tau/2$. (figure 2.56)
Let this terminated signal be denoted as $x_\tau(t)$.
The signal $x_\tau(t)$ may be expressed as

$$x_\tau(t) = \begin{cases} x(t) & |t| < \tau/2 \\ 0 & \text{elsewhere} \end{cases}$$

Now since the terminated signal $x_\tau(t)$ is of finite duration τ, therefore it is an energy signal. Let the energy of this signal be denoted by E_τ and may be expressed as

$$E_\tau = \int_{-\infty}^{\infty} |x_\tau(t)|^2\, dt = \int_{-\infty}^{\infty} |X_\tau(\omega)|^2\, df \qquad \text{(Parseval's theorem)}$$

Here, $X_\tau(\omega)$ is the Fourier transform of $x_\tau(t)$. It may be observed that $x(t)$ over the interval $\left(-\dfrac{\tau}{2}, \dfrac{\tau}{2}\right)$ will be same as $x_\tau(t)$ over the interval $(-\infty, \infty)$.

Therefore, we have

$$\int_{-\infty}^{\infty} |x_\tau(t)|^2\, dt = \int_{-\tau/2}^{\tau/2} |x(t)|^2\, dt$$

Therefore,

$$\frac{1}{\tau}\int_{-\tau/2}^{\tau/2} |x(t)|^2\, dt = \frac{1}{\tau}\int_{-\infty}^{\infty} |X_\tau(\omega)|^2\, df$$

Taking limit $\tau \to \infty$ of both sides of above equation

$$\lim_{\tau\to\infty} \frac{1}{\tau}\int_{-\tau/2}^{\tau/2} |x(t)|^2\, dt = \lim_{\tau\to\infty} \frac{1}{\tau}\int_{-\infty}^{\infty} |X_\tau(\omega)|^2\, df$$

The left-hand side of above equation represents the average power P of the function $x(t)$. Therefore,

$$P = \int_{-\infty}^{\infty} \lim_{\tau\to\infty} \frac{|X_t(\omega)|^2}{\tau}\, df \qquad ...(2.61)$$

In the limit $\tau \to \infty$, the ratio $\dfrac{|X_\tau(\omega)|^2}{\tau}$ may be approach a finite value.

Let this finite value be $S(\omega)$.

So that $\qquad S(\omega) = \lim\limits_{\tau\to\infty} \dfrac{|X_\tau(\omega)|^2}{\tau} \qquad ...(2.62)$

Using equations (2.61) and (2.62), we have

$$P = \int_{-\infty}^{\infty} S(\omega)\, df = \frac{1}{2\pi}\int_{-\infty}^{\infty} S(\omega)\, d\omega$$

According to above equation, the total power of the signal is obtained by multiplying $S(\omega)$ with bandwidth $\Delta\omega(d\omega)$ and integrating over the entire bandwidth.

Hence, $S(\omega)$ may be treated as average power per unit bandwidth and so is called as **power spectral density** or **power density spectrum.**

Since $\qquad |X_\tau(\omega)|^2 = |X_\tau(-\omega)|^2$

Therefore, the power contribution by positive and negative frequencies is identical.

Thus, average power is expressed as

$$P = \int_{-\infty}^{\infty} S(\omega)\, df = 2\int_{0}^{\infty} S(\omega)\, df = \frac{1}{\pi}\int_{0}^{\infty} S(\omega)\, d\omega$$

Relationship between input and output power spectral densities

This relationship, like energy spectral density, is expressed as

$$S_r(\omega) = |H(\omega)|^2 \, S_x(\omega)$$

where $S_r(\omega)$ = power spectral density of response $r(t)$

$S_x(\omega)$ = power spectral density of input $x(t)$

$H(\omega)$ = transfer function of the system

2.13.3. Properties of Power Spectral Density Function

Following are the properties of the power spectral density function:

Property 1: The area under the power density function is equal to the average power of that signal.

Mathematically $\quad P = \int_{-\infty}^{\infty} S(f)\, df$

Property 2: If $x(t)$ is input to a linear time invariant (LTI) system with transfer function $H(j\omega)$, then input and output power spectral density functions are related as

$$S_0(\omega) = |H(\omega)|^2 \cdot S_i(\omega)$$

where $S_0(\omega)$ = output power spectral density

$S_i(\omega)$ = input power spectral density

and $|H(\omega)|^2$ = power gain of the linear time invariant (LTI) system

Property 3: The autocorrelation function $R(\tau)$ and power spectral density function $S(\omega)$ form a Fourier transform pair.

Mathematically,

$$R(\tau) \longleftrightarrow S(\omega)$$

Table 2.1. Comparison of Energy spectral density function and Power Spectral Density Function

Energy spectral density $\psi(\omega)$	*Power spectral density $S(\omega)$*
1. Energy spectral density gives the distribution of energy of a signal in frequency-domain.	Power spectral density gives the distribution of power of a signal in frequency-domain.
2. Energy spectral density is given as $\psi(\omega) = \lvert X(\omega)\rvert^2$	Power spectral density is given as $S(\omega) = \lim_{\tau\to\infty} \frac{\lvert X_\tau(\omega)\rvert^2}{\tau}$
3. Total energy $E = \frac{1}{2\pi}\int_{-\infty}^{\infty} \psi(\omega)\, d\omega$ or $E = \int_{-\infty}^{\infty} \psi(f)\, df$	Total power $P = \frac{1}{2\pi}\int_{-\infty}^{\infty} S(\omega)\, d\omega$ or $P = \int_{-\infty}^{\infty} S(f)\, df$
4. Autocorrelation function for energy signal and energy spectral density form a Fourier transform pair.	Autocorrelation function for power signal and power spectral density form a Fourier transform pair.

2.14. Cross-Correlations

The cross-correlation between two different waveforms or two signals may be defined as the measure of match or similarity between one signal and time-delayed version of another signal.

This means that the cross-correlation between two signals explains how much one signal is related to the time-delayed version of another signal.

Let us consider two general complex signals $x_1(t)$ and $x_2(t)$ which may or may not be periodic and not restricted to a finite interval. Then the cross-correlation between these two signals $x_1(t)$ and $x_2(t)$ is defined as

$$R_{1,2}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{T/2} x_1(t)\, x_2^*(t-\tau)\, dt \qquad \text{...(2.63)}$$

From the cross-correlation expression given in equation (2.63), few conclusions may be drawn as under:

(*i*) From the expression of cross-correlation, it is clear that the cross-correlation represents the overlapping area between the two signals.

(*ii*) In the expression of cross-correlation, the conjugate symbol '*' may be neglected if the signals are real valued signals.

(*iii*) In the expression of cross-correlation, the parameter 'τ' is the delay parameter. It is also known as searching or scanning parameter.

(*iv*) From the expression of cross-correlation, it may be observed that the time t is a dummy variable and the cross-correlation $R_{1,2}(t)$ is a function of the delay or searching parameter 'τ'.

(*v*) The searching parameter 'τ' is required to determine the maximum possible correlation between two signals. Infact, the two signals having no cross-correlation at $\tau = 0$. (*i.e.* no delay between them) can have significant cross-correlation by adjusting the suitable value of searching parameter 'τ'.

This point may be further illustrated with the help of figure 2.57.

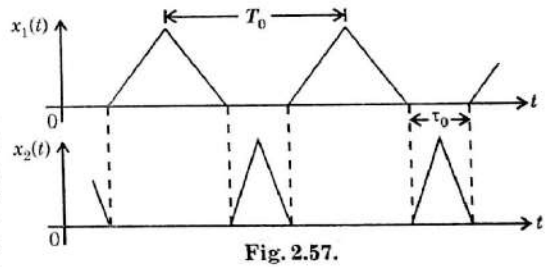


Fig. 2.57.

In figure 2.57. the two signals seem to have considerable similarity, however the cross-correlation expression in equation gives zero value of cross-correlation at $\tau = 0$, since they are non-overlapping. But when one signal is delayed, cross-correlation is not zero. This means that the cross-correlation increases as the dealy or searching parameter τ increases and it will be maximum at a particular value $\tau = \tau_0$. Therefore, we can say that the parameter 'τ' is a searching or scanning parameter that corresponds to the maximum possible correlation between two signals.

(*vi*) The signals, for which cross-correlation $R_{1,2}(\tau)$ is zero for all values of delay-parameter τ are known as uncorrelated or incoherent signals.

2.14.1. Cross-Correlation of Energy and Power Signals

The cross-correlation is defined separately for energy signals and for power or periodic signals.

2.14.2. Cross-Correlation of Energy Signal

Let us consider two general complex signals $x_1(t)$ and $x_2(t)$ of finite energy.

The cross-correlation of these two energy signals is given as

Cross-correlation of energy signals

$$R_{1,2}(\tau) = \int_{-\infty}^{\infty} x_1(t)\, x_2^*(t-\tau)\, dt$$

If $x_1(t)$ and $x_2(t)$ have some similarity, then the cross-correlation function $R_{1,2}(t)$ will have some finite value over the range of τ.

Also, if $$\int_{-\infty}^{\infty} x_1(t)\, x_2^*(t) = 0 \quad ...(2.64)$$

i.e. $$R_{1,2}(0) = 0$$

then the two signals $x_1(t)$ and $x_2(t)$ are known as **orthogonal signals**. This means that the two signals have no cross-correlation between them.

The another cross-correlation function for $x_1(t)$ and $x_2(t)$ is expressed as

$$R_{2,1}(\tau) = \int_{-\infty}^{\infty} x_2(t) \cdot x_1^*(t-\tau)\, dt \quad ...(2.65)$$

2.14.3. Properties of Cross-Correlation Function for Energy Signals

Following are the properties of cross-correlation function for energy signals :

Property 1: The cross-correlation functions exhibit conjugate symmetry.

Mathematically,

$$R_{1,2}(\tau) = R^*_{2,1}(\tau) \quad ...(2.66)$$

At this point, it may be noted that unlike convolution, cross-correlation function is not in general commutative, *i.e.*,

$$R_{1,2}(\tau) \neq R_{2,1}(\tau)$$

Property 2: If $$R_{12}(0) = 0$$

i.e. $$\int_{-\infty}^{\infty} x_1(t) \cdot x_2^*(t)\, dt = 0 \quad ...(2.67)$$

then the two signals $x_1(t)$ and $x_2(t)$ are called **orthogonal** over the entire time interval.

Property 3: The cross-correlation function of the energy signals corresponds to the multiplication of the Fourier transform of one signal and the complex conjugate of the Fourier transform of another signal.

Mathematically,

$$R_{1,2}(\tau) \longleftrightarrow X_1(\omega)\, X_2^*(\omega) \quad ...(2.68)$$

The above expression is also called as **correlation theorem**.

2.14.4. Cross-Correlation of Power or Periodic Signals

The cross-correlation function $R_{1,2}(\tau)$ for two periodic signals $x_1(t)$ and $x_2(t)$ may be defined with the help of average form of correlation.

Mathematically,

Cross-correlation of two period signals $x_1(t)$ and $x_2(t)$ of time-period T_0, is expressed as

$$R_{1,2}(\tau) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x_1(t) \cdot x_2^*(t-\tau)\, dt \quad ...(2.69)$$

The another cross-correlation function of two periodic signals is defined as

$$R_{2,1}(\tau) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x_2(t) \cdot x_1^*(t-\tau)\, dt \quad ...(2.70)$$

2.14.5. Properties of Cross-Correlation Function for Periodic Signals

Following are the properties of cross-correlation function for periodic signals:

Property 1: If cross-correlation is taken at origin (*i.e.*, $\tau = 0$)

or If $$R_{12}(0) = 0$$

i.e., $$\frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x_1(t) \cdot x_2^*(t)\, dt = 0 \quad ...(2.71)$$

then the two signals are called as **orthogonal** over the entire time interval.

Property 2: The cross-correlation function exhibits conjugate symmetry.

Mathematically,

$$R_{1,2}(\tau) = R_{2,1}^*(\tau) \quad ...(2.72)$$

Property 3: Unlike convolution, the cross-correlation function of two signals is not commutative.

Mathematically,

$$R_{1,2}(\tau) \neq R_{2,1}(\tau) \quad ...(2.73)$$

Example 2.28. Prove the following relation:

$$R_{1,2}(\tau) = R_{2,1}^*(\tau)$$

Solution: We know that the cross-correlation of two signals $x_1(t)$ and $x_2(t)$ is expressed as

$$R_{1,2}(\tau) = \int_{-\infty}^{\infty} x_1(t)\, x_2^*(t-\tau)\, dt \quad ...(i)$$

Substituting $t - \tau = y$ in equation (*i*), we get

$$R_{1,2}(t) = \int_{-\infty}^{\infty} x_1(y+\tau)\, x_2^*(y)\, dy \quad ...(ii)$$

Also, we know that $$R_{2,1}(\tau) = \int_{-\infty}^{\infty} x_2(t)\, x_1^*(t-\tau)\, dt \quad ...(iii)$$

Substituting, $t = y$ in equation (*iii*), we get

$$R_{2,1}(\tau) = \int_{-\infty}^{\infty} x_2(y)\, x_1^*(y-\tau)\, dy$$

Therefore, $$R_{2,1}^*(\tau) = \int_{-\infty}^{\infty} x_2^*(y)\, x_1(y+\tau)\, dy \quad ...(iv)$$

Comparing equation (*iv*) with equation (*ii*), we get

$$R_{1,2}(\tau) = R_{2,1}^*(\tau)$$

which is the given relationship. **Ans.**

2.15. Autocorrelation Function

Autocorrelation function gives the measure of similarity, match or coherence between a signal and its delayed replica. This means that autocorrelation function is a special form of a cross-correlation function.

Let us consider a signal $x(t)$.

Then the autocorrelation function of this signal with its delayed version will be

$$R(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)\, x^*(t-\tau)\, dt \quad ...(2.74)$$

The autocorrelation function is defined separately for energy signals and for power or period signals.

2.15.1. Autocorrelation Function of Energy Signals

Let us consider an energy signal $x(t)$. Then the autocorrelation function of this signal may be obtained by integrating the product of signal $x(t)$ and delayed version of its complex conjugate.

Mathematically,

Autocorrelation of energy signals

$$R(\tau) = \int_{-\infty}^{\infty} x(t)\, x^*(t-\tau)\, dt \qquad ...(2.75)$$

where parameter 'τ' is defined as the time-delay parameter. In above equation, the complex valued signal $x(t)$ is delayed in positive-direction. Now, if the signal $x(t)$ is shifted by the same period 'τ' in negative direction, then the autocorrelation function in equation (2.74), becomes

$$R(\tau) = \int_{-\infty}^{\infty} x(t+\tau) \cdot x^*(t)\, dt \qquad ...(2.76)$$

2.15.2. Properties of Autocorrelation Function for Energy Signals

Following are the properties of autocorrelation function for energy signals:

Property 1: The autocorrelation function exhibits conjugate symmetry.

Mathematically,

$$R(\tau) = R^*(-\tau) \qquad ...(2.77)$$

This means that the real part of $R(\tau)$ is an even function of τ whereas the imaginary part of $R(\tau)$ is an odd function of τ.

Proof: We know that the autocorrelation function of energy signals is given as

$$R(\tau) = \int_{-\infty}^{\infty} x(t)\, x^*(t-\tau)\, dt \qquad ...(2.78)$$

Taking complex conjugate, we have

$$R^*(\tau) = \int_{-\infty}^{\infty} x^*(t)\, x(t-\tau)\, dt \qquad ...(2.79)$$

Therefore,

$$R^*(-\tau) = \int_{-\infty}^{\infty} x^*(t)\, x(t+\tau)\, dt$$

$$R^*(-\tau) = R(t) \qquad \text{[using equation (2.79)]}$$

or

$$R(t) = R^*(-\tau)$$

Property 2: The value of autocorrelation function at $\tau = 0$, (*i.e.* at origin) is equal to the energy of the signal

Mathematically,

$$R(0) = \int_{-\infty}^{\infty} |x(t)|^2\, dt = E \qquad ...(2.80)$$

Proof: We know that the autocorrelation function of energy signals is given as

$$R(\tau) = \int_{-\infty}^{\infty} x(t)\, x^*(t-\tau)\, dt$$

Substituting $\tau = 0$ provides

$$R(0) = \int_{-\infty}^{\infty} x(t)\, x^*(t)\, dt = \int_{-\infty}^{\infty} |x(t)|^2\, dt = E \qquad ...(2.81)$$

Property 3: If τ is increased in either direction, the autocorrelation reduces. As τ reduces, autocorrelation increases and it is maximum at $\tau = 0$ *i.e.*, at origin.

Mathematically,

$$|R(\tau)| \le R(0) \quad \text{for all } \tau \qquad ...(2.82)$$

Property 4: The autocorrelation function and energy spectrum density function of energy signal forms a Fourier transform pair.

Mathematically,

$$R(\tau) \longleftrightarrow \Psi(\omega) \qquad ...(2.83)$$

2.15.3. Autocorrelation Function of Power (*i.e.*, Periodic) Signals

Let the signal $x(t)$ be periodic with period T_0. The autocorrelation function of periodic signals over one period is given as

$$R(\tau) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t)\, x^*(t-\tau)\, dt$$

If τ is in negative direction, then

$$R(\tau) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} x(t+\tau)\, x^*(t)\, dt$$

For any period T, $R(\tau)$ is given as

$$R(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)\, x^*(t-\tau)\, dt \qquad ...(2.84)$$

2.15.4. Properties of Autocorrelation Function for Periodic Signals

Following are the properties of autocorrelation function for periodic signals :

Property 1: If $R(\tau)$ and $R^*(-\tau)$ are the complex conjugates of each other, then

$$R(\tau) = R^*(-\tau) \qquad ...(2.85)$$

This property is known as conjugate property.

Proof: We know that autocorrelation function for periodic signals is given as

$$R(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t)\, x^*(t-\tau)\, dt$$

Therefore,

$$R^*(\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x^*(t)\, x(t-\tau)\, dt$$

Thus,

$$R^*(-\tau) = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x^*(t)\, x(t+\tau)\, dt = R(\tau)$$

or

$$R(\tau) = R^*(-\tau)$$

Property 2: The autocorrelation function at origin is equal to the average power of the signa

Mathematically,

$$R(0) = \frac{1}{T_0} \int_{-T_0/2}^{T_0/2} |x(t)|^2\, dt = P \qquad ...(2.$$

Proof: The autocorrelation function for periodic signals is given as

$$R(\tau) = \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} x(t)\, x^*(t-\tau)\, dt$$

Substituting $\tau = 0$, we get

$$R(0) = \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} x(t)\cdot x^*(t)\, dt = \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} |x(t)|^2\, dt$$

$$R(0) = \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} |x(t)|^2\, dt = P$$

Property 3: The autocorrelation function has maximum value at the origin.

Mathematically,

$$R(\tau) \le R(0) \qquad ...(2.87)$$

This means that the value of autocorrelation function reduces as τ increases from zero.

Property 4: The autocorrelation function and power spectrum density form a Fourier transform pair.

Mathematically

$$R(\tau) \longleftrightarrow S(\omega) \qquad ...(2.88)$$

Property 5: The autocorrelation function is periodic with the same period as periodic signal itself.

Mathematically,

$$R(\tau) = R(\tau \pm n\, T_0),\ n = 1, 2 ... \qquad ...(2.89)$$

Example 2.29. The autocorrelation function of an aperiodic power signal is given as

$$R(\tau) = e^{(-\tau^2/2\sigma^2)}$$

Determine the power spectral density (psd) and the normalized average power content of the signal. *(U.P.S.C., I.E.S. Examination, 1998)*

Solution: We know that the autocorrelation function and the power spectral density form a Fourier transform pair, *i.e.*,

$$R(\tau) \longleftrightarrow S(\omega)$$

This means that

$$S(\omega) = \int_{-\infty}^{\infty} R(\tau)\, e^{-j\omega\tau} d\tau \qquad ...(i)$$

and

$$R(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S(\omega)\, e^{j\omega\tau}\, d\omega \qquad ...(ii)$$

Using equation (*i*), we get

$$S(\omega) = \int_{-\infty}^{\infty} R(\tau)\, e^{-j\omega\tau}\, d\tau = \int_{-\infty}^{\infty} e^{-\tau^2/2\sigma^2}\, e^{-j\omega\tau}\, d\tau$$

$$S(\omega) = \sqrt{2\pi\sigma^2}\cdot e^{-(\omega\sigma)^2/2} \qquad \textbf{Ans.}$$

The normalized average power is given as

$$P = R(0) = e^0 = 1 \qquad \textbf{Ans.}$$

Example 2.30. Find the autocorrelation function of the sine wave signal expressed as below:

$$x(t) = A \sin(\omega_1 t + \phi)$$

where $\omega_1 = \dfrac{2\pi}{T}$.

Solution: The autocorrelation function for periodic signal is given as

$$R(\tau) = \frac{1}{T_0}\int_{-T_0/2}^{T_0/2} x(t)\, x^*(t-\tau)\, dt$$

Since this signal is a real valued signal and the period T_0 is not specified, let us use the following definition for autocorrelation,

$$R(\tau) = \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2} x^*(t)\, x(t+\tau)\, dt$$

or

$$R(\tau) = \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2} A\sin(\omega_1 t+\phi)\, A\sin(\omega_1 t+\omega_1\tau+\phi)\, dt$$

or

$$R(\tau) = \lim_{T\to\infty}\frac{1}{T}\frac{A^2}{2}\int_{-T/2}^{T/2} [\cos\omega_1\tau - \cos(2\omega_1 t + 2\phi + \omega_1\tau)]\, dt$$

or

$$R(\tau) = \lim_{T\to\infty}\frac{1}{T}\frac{A^2}{2}\cos\omega_1\tau\int_{-T/2}^{T/2} dt - \lim_{T\to\infty}\frac{1}{T}\int_{-T/2}^{T/2}\cos(2\omega_1 t + 2\phi + \omega_1\tau)\, dt$$

Taking limits, we get

$$R(\tau) = \frac{A^2}{2}\cos\omega_1\tau \qquad \textbf{Ans.}$$

2.16. Application of Fourier Transform to the Analysis of LTI Networks

As a matter of fact, for a continuous-time LTI network, if input signal is $x(t)$ and unit-impulse response of LTI system is $h(t)$, then the output or response $y(t)$ of the system is expressed as a convolution integral, *i.e.*,

$$y(t) = \int_{-\infty}^{\infty} x(\tau)\, h(t-\tau)\, d\tau = x(t) \otimes h(t) \qquad ...(2.90)$$

Hence according to above equation, for a continuous-time LTI system, the output of any general input may be found by convolving the input signal $x(t)$ with the systems unit impulse response $h(t)$.

In other words, if the response or output of a continuous-time LTI system is known for an impulse signal then the response to any other signal $x(t)$ may be obtained by above equation.

It may be noted at this stage, that the equation is helpful in time-domain only. In other words, when input and output are in time-domain then only convolution equation is used.

Also, the use of convolution in finding the response is quite complex.

To ease the analysis, Fourier transform is used. With the help of Fourier transform, we get simpler equations in frequency-domain to find the response or output of an LTI system as under:

Applying Time-convolution theorem to the equation $y(t) = x(t) \otimes h(t)$,

we obtain $\quad Y(\omega) = X(\omega)\, H(\omega)$

where $\quad y(t) \longleftrightarrow Y(\omega)$

$$x(t) \longleftrightarrow X(\omega)$$

and $$h(t) \longleftrightarrow H(\omega)$$

Therefore, according to above equation, the output of LTI system in frequency-domain is equal to the product of the input signal in frequency-domain and unit impulse response in frequency-domain.

From last equation, we have

$$H(\omega) = \frac{Y(\omega)}{X(\omega)}$$

This equation indicates that the transform $H(\omega)$ of the impulse response $h(t)$ is equal to the ratio of transforms of output and input. Hence, this $H(\omega)$ is called transfer function of the LTI system. Equally, $H(\omega)$ is also called the **frequency-response** of continuous-time linear-time invariant (LTI) system.

2.16.1. Definition of Frequency-Response or Transfer-Functions

Frequency-response $H(\omega)$ is given as

$$H(\omega) = \frac{Y(\omega)}{X(\omega)} \qquad ...(2.91)$$

Now if $X(\omega) = 1$, then according to above equation, $H(\omega)$ is equal to $Y(\omega)$, *i.e.*,

$$H(\omega) = Y(\omega) \quad \text{if} \quad X(\omega) = 1$$

But since $X(\omega) = 1$ corresponds to an unit-impulse signal in the time-domain, therefore $H(\omega)$ is response (in frequency domain) of the LTI system when its input is an unit impulse signal.

Hence, frequency response or transfer-function may be defined as under:

"The frequncy response or transfer function of an LTI system is transform of its response when the input is an unit impulse function".

$H(\omega)$, in general, is complex and may be expressed as

$$H(\omega) = |H(\omega)|\, e^{j\theta(j\omega)} \qquad ...(2.92)$$

Here $|H(\omega)|$ is the amplitude response and $\theta(j\omega)$ is the phase response.

$H(\omega)$ exhibits the conjugate symmetry for a linear system which implies that the amplitude response is symmetrical, and phase response is antisymmetrical about the vertical axis *i.e.*,

$$|H(\omega)| = |H(-\omega)|$$

and $$\theta(\omega) = -\theta(-\omega)$$

As a matter of fact, it is often preferable to work on the logarithmic scale. Now, taking logarithm on both sides of equation (2.92), we have

$$\ln H(\omega) = \ln|H(\omega)| + j\,\theta(\omega) = \alpha(\omega) + j\,\theta(\omega)$$

Here, the function $\alpha(\omega)$ is known as the gain of the system whereas $\theta(\omega)$ is known as the phase shift introduced by the system. Also, the gain $\alpha(\omega)$ is measured in neper or decibels (dBs) depending upon the base of the logarithm *i.e.*,

$$\alpha(\omega) = \ln|H(\omega)| \text{ nepers}$$

$$= 20 \log_{10}|H(\omega)| \text{ decibel or dB}$$

Further, $\theta(\omega)$ is measured in radians.

Note: *It may be noted that the gain has positive value if the response is greater than excitation and has zero value if the response is same as the excitation. On the other hand, if the gain $\alpha(\omega)$ is negative in dB, then the response would be less than the excitation and hence the system would produce a loss. Further, this negative value of the gain $\alpha(\omega)$ for which the system produces a loss, is known as attenuation.*

The neper and decibal (dB) are related to each other as under:

1 neper = 8.69 dB

Example 2.31. Figure 2.58 shows a low-pass RC network. For this network or system find the output or response if the input signal is $x(t) = e^{-t/RC}$ as shown in figure 2.58(*a*).

(A.M.I.E., Examination, 1999)

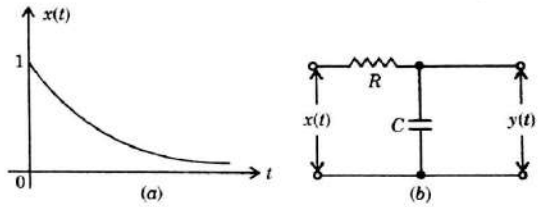


Fig. 2.58.

Solution: Given that the input signal

$$x(t) = e^{-t/RC}$$

Let $h(t)$ be the unit-impulse response of the given system.

Then output or response is given by

$$y(t) = x(t) \otimes h(t)$$

Applying time convolution theorem, we get

$$Y(\omega) = X(\omega)\, H(\omega) \qquad ...(i)$$

Now $$X(\omega) = F[x(t)] = \int_{-\infty}^{\infty} x(t)\, e^{-j\omega t}\, dt$$

According to figure 2.58(*a*), the input signal $x(t)$ exists only for positive values of t. Hence

$$X(\omega) = \int_0^{\infty} e^{-t/RC} \cdot e^{-j\omega t}\, dt = \int_0^{\infty} e^{-(1/RC + j\omega)t}\, dt$$

or $$X(\omega) = \left[\frac{e^{-(1/RC + j\omega)t}}{-\left(\frac{1}{RC} + j\omega\right)}\right]_0^{\infty} = -\frac{1}{\frac{1}{RC} + j\omega}\left[e^{-\infty} - e^{0}\right]$$

or $$X(\omega) = -\frac{1}{j\omega + \frac{1}{RC}}[0 - 1] = \frac{1}{j\omega + \frac{1}{RC}} \qquad ...(ii)$$

Also, from figure 2.58*b*, the transfer function $H(\omega)$ us given as

$$H(\omega) = \frac{\frac{1}{j\omega C}}{R + \frac{1}{j\omega C}} = \frac{1}{j\omega RC + 1} = \frac{1}{RC \cdot \left(j\omega + \frac{1}{RC}\right)} \qquad ...(iii)$$

Using equations (*i*), (*ii*) and (*iii*), we get

$$Y(\omega) = X(\omega)\, H(\omega) = \frac{1}{\left(j\omega + \frac{1}{RC}\right)} \cdot \frac{1}{RC\left(j\omega + \frac{1}{RC}\right)}$$

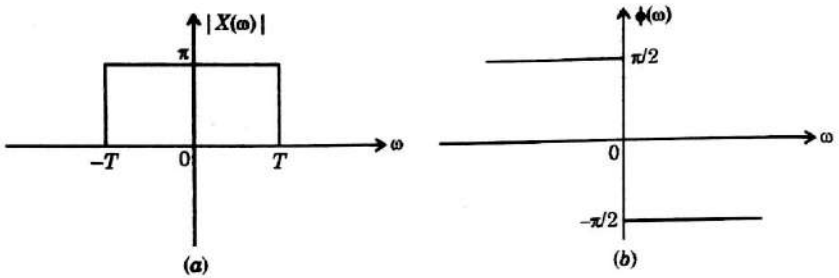


Fig. 2.71. *(a) Magnitude spectrum, (b) Phase spectrum.*

For the limits $-T \le \omega \le 0$, $X(\omega) = \pi\, e^{j\pi/2}$
For the limits $0 \le \omega \le T$, $X(\omega) = \pi\, e^{-j\pi/2}$
Now, $x(t)$ can be determined as under:

$$x(t) = F^{-1}[X(\omega)] = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(\omega)\, e^{j\omega t} d\omega$$

@sTc(small team creation)

ir $$x(t) = \frac{1}{2\pi}\left[\int_{-T}^{0} \pi e^{j\pi} e^{j\omega t} d\omega + \int_{0}^{T} \pi e^{-j\omega t} e^{j\omega t} d\omega\right]$$

or $$x(t) = \frac{1}{t}\left[\frac{e^{j\pi/2} - e^{-j\pi/2}}{2j}\right] - \frac{1}{t}\left[\frac{e^{j(\pi/2)-T_t} - e^{-j(\pi/2)-T_t}}{2j}\right]$$

or $$x(t) = \frac{1}{t}\left[\sin\left(\frac{\pi}{2}\right) - \sin\left(\frac{\pi}{2} - T_t\right)\right]$$

or $$x(t) = \frac{1}{t}[1 - \cos T_t] = \frac{2\sin^2\left(\frac{T_t}{2}\right)}{t} = \frac{T^2 t}{2}\operatorname{sinc}^2\left(\frac{T_t}{2}\right) \quad \textbf{Ans.}$$

Example 2.45. Find the Fourier transform of the trapezoid pulse shown in figure 2.72.
Solution: The function $x(t)$ is given as

$$x(t) = \begin{cases} \dfrac{At}{t_P} + \dfrac{At_P}{t_P - t_Q}, \text{ for } -t_P < -t_Q \\ A, \quad \text{for } -t_Q\ t < t_Q \\ \dfrac{At_P}{t_P - t_Q} - \dfrac{At}{t_P - t_P}, \text{for } t_Q < t < t_P \end{cases} \quad ...(i)$$

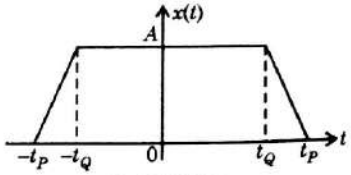


Fig. 2.72.

The Fourier transform of function $x(t)$ is determined by

$$X(\omega) = F[x(t)] = \int_{-\infty}^{\infty} x(t) e^{-j\omega t} dt$$

or $$= \int_{-t_P}^{-t_Q}\left[\frac{At}{t_P - t_Q} + \frac{At_P}{t_P - t_P}\right] e^{-j\omega t} dt + \int_{-t_Q}^{t_Q} [A] e^{-j\omega t} dt + \int_{t_Q}^{t_P}\left[\frac{At_P}{t_P - t_P} - \frac{At}{t_P - t_Q}\right] e^{-j\omega t} dt$$

or $$X(\omega) = \frac{A}{t_P - t_Q}\left[\int_{-t_P}^{-t_P} te^{-j\omega t} dt - \int_{t_Q}^{t_P} te^{-j\omega t} dt\right] + \frac{At_P}{t_P - t_Q}\left[\int_{-t_P}^{-t_Q} e^{-j\omega t} dt + \int_{t_Q}^{t_P} e^{-j\omega t} dt\right] + A\int_{-t_P}^{t_Q} e^{-j\omega t} dt$$

or $$X(\omega) = \frac{2A}{\omega^2(t_P - t_P)}[\cos \omega t_Q - \cos \omega t_P]$$

This is the required Fourier transform of the trapezoidal pulse shown in figure 2.72.

Example 2.46. Determine the Fourier transform of the signal shown in figure 2.73.
Solution: The given signal can be represented analytically as under:

$$x(t) = \begin{cases} 2, \text{for } -1 < t < 1 \\ 1, \text{for } -2 < t < 1 \\ 1, \text{for } 1 < t < 2 \end{cases} \quad ...(i)$$

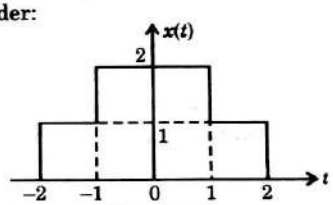


Fig. 2.73.

Therefore, $$X(\omega) = \int_{-\infty}^{\infty} x(t) e^{-j\omega t} dt$$

or $$X(\omega) = \int_{-1}^{1} (2) e^{-j\omega t} dt + \int_{-2}^{-1} (1) e^{-j\omega t} dt + \int_{0}^{2} (1) e^{-j\omega t} dt$$

or $$X(\omega) = \frac{1}{-j\omega}\{[2e^{-j\omega t}]_{-1}^{1} + [e^{-j\omega t}]_{-2}^{-1} + [e^{-j\omega t}]_{1}^{2}\}$$

or $$X(\omega) = \frac{1}{-j\omega}\{2(e^{-j\omega} - e^{j\omega}) + (e^{j\omega} e^{2j\omega}) + (e^{-j2\omega} - e^{-j\omega})\}$$

or $$X(\omega) = \frac{1}{\omega}\left\{4\left[\frac{e^{j\omega} - e^{j\omega}}{2j}\right] + 2\left[\frac{e^{j2\omega} - e^{-j2\omega}}{2j}\right] - 2\left[\frac{e^{j\omega} - e^{-j\omega}}{2j}\right]\right\}$$

or $$X(\omega) = \frac{4}{\omega}\left\{\sin \omega + \frac{\sin 2\omega}{2} - \frac{\sin \omega}{2}\right\} = \frac{2\sin \omega}{\omega} + \frac{2\sin 2\omega}{\omega}$$

or $$X(\omega) = 2 \operatorname{sinc} \omega + 4 \operatorname{sinc} 2\omega \quad \textbf{Ans.}$$

Example 2.47. Determine the Fourier transform of the signal $x(t) = t \cos At$.
Solution: We know that the Fourier transform of the signal $x(t)$ is given as

$$X(\omega) = F[x(t)] = \int_{-\infty}^{\infty} x(t) e^{-j\omega t} dt = \int_{0}^{\infty} t \cos At\, dt$$

or $$X(\omega) = \int_{0}^{\infty} t\left[\frac{e^{jAt} + e^{-jAt}}{2}\right] e^{-j\omega t} dt$$

or $$X(\omega) = \frac{1}{2}\int_{0}^{\infty} [te^{-(A-\omega)t} + te^{-j(A+\omega t)}] dt$$

or $$X(\omega) = \frac{1}{2}\left[t\left\{\frac{e^{-j(-A+\omega)t}}{-j(-A+\omega)}\right\} - \frac{e^{-j(-A+\omega)t}}{[-j(-A+\omega)^2}\right]_0^\infty + \frac{1}{2}\left[t\left\{\frac{e^{-j(-A+\omega)t}}{-j(-A+\omega)}\right\} - \frac{e^{-j(-A+\omega)t}}{[-j(-A+\omega)]^2}\right]_0^\infty$$

Solving, we get

$$X(\omega) = \frac{-(\omega^2 + A^2)}{(\omega^2 + A^2)^2} \quad \textbf{Ans.}$$

Example 2:48. Determine the Fourier transform of the sinusoidal pulse shown in figure 2.74.

Solution: The Fourier transform of signal $x(t)$ is defined as

$$X(\omega) = F[x(t)] = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}dt \qquad ...(i)$$

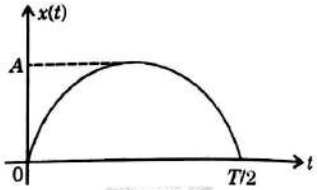


Fig. 2.74.

The signal $x(t)$ shown in figure 2.74, can be analytically given as under:

$$x(t) = \begin{cases} A\sin\omega_0 t \text{ for } 0 < t \le \dfrac{T}{2} \\ 0 \quad \text{otherwise} \end{cases} \qquad ...(ii)$$

Substituting equation (*ii*) in equation (*i*), we get

$$X(\omega) = \int_0^{T/2} A\sin\omega_0 t\, e^{-j\omega t}dt = A\int_0^{T/2}\left[\frac{e^{j\omega_0 t} - e^{-j\omega_0 t}}{2j}\right]e^{-j\omega t}dt$$

or $$X(\omega) = \frac{A}{2j}\int_0^{T/2} e^{j(\omega_0-\omega)t}dt - \frac{A}{2j}\int_0^{T/2} e^{-j(\omega_0+\omega)t}dt = \frac{A}{2j}\left[\frac{e^{j(\omega_0-\omega)t}}{j(\omega_0-\omega)} - \frac{e^{-j(\omega_0+\omega)t}}{-j(\omega_0+\omega)}\right]_0^{T/2}$$

On solving above equation, we obtain

$$X(\omega) = \frac{A\omega_0}{\omega_0^2 - \omega^2}[e^{j\omega T/4} + e^{-j\omega T/4}]\,e^{-j\omega T/4}$$

$$X(\omega) = \left(\frac{2A\omega_0}{\omega_0^2 - \omega^2}\cos\frac{\omega T}{4}\right)e^{-j\omega T/4} \quad \textbf{Ans.}$$

Example 2.49. The Fourier transform of the signal $x(t) = e^{-3t^2}$ is of the following form, where A and B are constants.

(*a*) $Ae^{-B|f|}$ (*b*) Ae^{-Bf}

(*c*) $A + B|f|^2$ (*d*) Ae^{-Bf} *(GATE Examination, 2002)*

Solution: The correct option is (*d*). **Ans.**

SUMMARY

1. Signals may be classified broadly into two types: Continuous-time signals and discrete-time signals.
2. When we analyze continuous-time signal with the help of Fourier series and Fourier transform, then Fourier series and Fourier transform are called continuous-time Fourier series and Continuous-time Fourier transform respectively or simply Fourier series and Fourier transform. Thus, in this chapter, we shall study continuous-time Fourier series (CTFS) and continuous-time Fourier transformer (CTFT).

3. When we analyze discrete-time signal with the help of Fourier series and Fourier transform, then Fourier series and Fourier transform are called discrete-time Fourier series and discrete-time Fourier transform respectively
4. Signals contain information about a variety of things and activities in physical world. An electrical signal may be represented in two equivalent forms, *i.e.*, as a voltage signal or a current signal. This means that an electrical signal may be represented either in the form of a voltage source or in the form of a current source.
5. An electrical signal, either a voltage signal or a current signal, may further be represented in two forms:

   These two types of representation are as under:

   (*i*) Time domain representation

   (*ii*) Frequency domain representation
6. The continuous-time Fourier series exists only when the function $x(t)$ satisfies the following three conditions called **Dirichlet's conditions**.

   (*i*) $x(t)$ is well-defined and a single-valued function.

   (*ii*) $x(t)$ must possess only a finite number of discontinuities in the period T.

   (*iii*) $x(t)$ must have a finite number of positive and negative maxima in the period T.
7. A periodic function $x(t)$ may be expressed in the form of trigonometric Fourier series comprising the following sine and cosine terms:

$$x(t) = a_0 + a_1\cos\omega_0 t + a_2\cos_2\omega_0 t + ... + a_n\cos n\omega_0 t + ...$$
$$+ b_1\sin\omega_0 t + b_2\sin_2\omega_0 t + ... + b_n\sin n\omega_0 t + ...$$

$$x(t) = a_0 + \sum_{n=1}^{\infty}(a_n\cos\omega_0 t + b_n\sin n\omega_0 t)$$

8. Negative frequency signals are not physical signals. They are only a mathematical tool or concept required to provide a real signal by a combination of complex exponentials of positive and negative frequencies.
9. Following are the important properties of continuous-time Fourier Series (CTFS):

1. Linearity	5. Time-Differentiation
2. Time-shifting	6. Convolution
3. Frequency-shifting	7. Multiplication
4. Time scaling	8. Symmetry property.

10. Fourier transform may be expressed as

$$X(\omega) = F[x(t)] = \int_{-\infty}^{\infty} x(t)e^{-j\omega t}dt$$

In above equation, $X(\omega)$ is called the Fourier transform of $x(t)$. In other words, $X(\omega)$ is the frequency domain representation of time domain function $x(t)$. This means that we are converting a time domain signal into its frequency domain representation with the help of Fourier transform.

11. Inverse Fourier transform

$$F^{-1}[X(\omega)] = x(t)$$

$$x(t) = F^{-1}[X(\omega)] = \frac{1}{2\pi}\int_{-\infty}^{\infty} X(\omega)e^{j\omega t}d\omega$$

12. Symbolically, this may be expressed as

$$x(t) \longleftrightarrow X(\omega)$$

Here, double ended arrow means that Fourier transform of $x(t)$ is $X(\omega)$ and inverse Fourier transform of $X(\omega)$ is $x(t)$.

In general, Fourier transform $X(\omega)$ is a complex function of ω and may be expressed as,

$$X(\omega) = |X(\omega)|e^{j\theta(\omega)}$$

where $|X(\omega)|$ is called the amplitude spectrum and $X(\omega)$ is called the phase spectrum.

13. For a function $x(t)$ to be Fourier transformable, it is sufficient that $x(t)$ satisfies Dirichlet's conditions given below:
    (*i*) The function $x(t)$ is a single valued function with a finite number of maxima and minima and a finite, number of discontinuities in any finite time interval.
    (*ii*) The function $x(t)$ is absolutely integrable. Mathematically,

$$\int_{-\infty}^{\infty} |x(t)|\,dt < \infty$$

14. The functions $\frac{\sin x}{x}$ is the "sine over argument" function and is denoted by $\sin c(x)$. This function plays an important role in signal processing. It is also known as the filtering or interpolating function. Mathematically,

$$\sin c(x) = \frac{\sin x}{x} = \frac{\sin x}{x}$$

SHORT QUESTIONS WITH ANSWERS

Q.1. What is Fourier series as a complex Exponential?

Ans. Given a signal $x(t)$ over the interval $(t_0, t_0 + T)$ the complex Exponential Furier series is

$$x(t) = \sum_{n=-\infty}^{\infty} X_n e^{jn\omega_0 t}, t_0 \le t \le t_0 + T_0$$

where $$X_n = \frac{1}{T_0}\int_{t_0}^{t_0+T_0} x(t)e^{-jn\omega_0 t}.dt$$

Q.2. What are Fourier coefficients in time averages?

Ans. $$X_n = |X_n|e^{jX_n}$$

where $$X_n = \langle x(t)\cos n\omega_0 t\rangle - j\langle x(t)\sin n\omega_0 t\rangle.$$

and $$\langle x(t)\rangle = \begin{cases} \lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} x(t).dt\ [x(t) - \text{non-periodic}] \\ \frac{1}{T_0}\int_{T_0}^{-T} x(t)\,dt\ [x(t) - \text{Periodic}] \end{cases}$$

Q.3. What is Fourier integral?

Ans. $$X(\omega) = \int_{-\infty}^{\infty} x(t).e^{-j\omega t}\,dt$$

Q.4. State the superposition Theorem of Fourier transform.

Ans. $$F[a_1x_1(t) + a_2x_2(t)] = a_1X_1(f) + a_2X_2(f)$$

where $$F(x_1(t)] = X_1(f)$$

and $$F[x_2(t)] = X_2(f).$$

Q.5. State the Time-delay Theorem of Fourier transform.

Ans. $$F[x(t - t_0) = X(f).e^{-2\pi f t_0}$$

where $$F[x(t)] = X(f)$$

Q.6. Write down the scale-change Theorem of Fourier transform.

Ans. $$F[x(at)] = \frac{1}{|a|}.X\left(\frac{f}{a}\right).$$

Q.7. Write down the duality Theorem of Fourier transform.

Ans. $$X(t) \leftrightarrow x(-f)$$

where $$F(x(t)) = X(f)$$

Q.8. Write the Frequency Translation Theorem of Fourier transform.

Ans. $$x(t)e^{j2\pi f_0 t} \leftrightarrow X(f - f_0)$$

Q.9. Write the modulation Theorem of Fourier transform.

Ans. $$x(t)\cos(2\pi f_0 t) \leftrightarrow \frac{1}{2}X(f - f_0) + \frac{1}{2}X(f + f_0)$$

Q.10. State the differentiation Theorem of Fourier transform.

Ans. $$\frac{d^n x(t)}{dt^n} \leftrightarrow (j^{2\pi}f)^n.X(f).$$

Q.11. State the Integration Theorem of Fourier transform.

Ans. $$\int_{-\infty}^{t} x(\lambda).d\lambda \leftrightarrow (j^{2\pi}f)^{-1}.X(f) + \frac{1}{2}X(0).\delta(f)$$

Q.12. State the convolution Theorem of Fourier transform. *(Madras University, 1994)*

Ans. $$\int_{-\infty}^{\infty} x_1(\lambda).x_2(t-\lambda)d\lambda = \int_{-\infty}^{\infty} x_1(t-\lambda).x_2(\lambda).\,d\lambda \leftrightarrow X_1(f).X_2(f)$$

Q.13. State the multiplication Theorem of Fourier transform.

Ans. $$x_1(t).x_2(k) \leftrightarrow X_1(f) \otimes X_2(f) = \int_{-\infty}^{\infty} X_1(\lambda)X_2(t-\lambda)\,d\lambda$$

Q.14. What is direct-delta function?

Ans. A function $x(t)$ whose spectrum (or Fourier transform) is $\delta(f)$ is called direct delta function.

i.e., $$X(f) = A.\delta(f)$$

where $[x(t) = A, -\infty \le t \le \infty]$

Q.15. State the properties of direct delta function.

Ans. (*a*) $\int_{-\infty}^{\infty} \delta(f).df = 1;$

(*b*) Sifting property: $\int_{-\infty}^{\infty} \delta(f - f_0)g(f)df = g(f_0);$

(*c*) $2\pi\delta(t) = \int_{-\infty}^{\infty} e^{jut}.du$

REVIEW QUESTIONS

1. Write the Dirichlet's conditions for Fourier series.
2. Write the trigonometric form of the Fourier series representation of a periodic signal.
3. Explain even and odd functions.
4. Explain the concept of negative frequency.
5. What is the difference between Fourier transform and Laplace transform. *(WBTU, Kolkata-2001)*
6. Explain the linearity and time-scaling properties of Fourier transform.
7. Explain the time-shifting and frequency shifting properties of Fourier transform.
8. What is convolution. Explain time-convolution and frequency convolution theorems. *(Anna University, Chennai-2000)*
9. Define energy and power of a signal
10. State and prove Parseval's theorem for energy signals.
11. Derive a relationship for energy-spectral density of a signal.
12. Write the Dirichlet's conditions for Fourier series.
13. Write the trigonometric form of the Fourier series representation of a periodic signal.
14. Explain even and odd functions.
15. Explain the concept of negative frequency.
16. What is the difference between Fourier transform and Laplace transform.
17. Explain the linearity and times-scaling properties of Fourier transform.
18. Explain the time-shifting and frequency shifting properties of Fourier transform. *(A.M.I.E. Examination, 1997)*
19. What is convolution. Explain time-convolution and frequency theorems.
20. Define energy and power of a signal.
21. State and prove Parseval's theorem density of a signal.
22. Derive a relationship for energy-spectral density of a signals.
24. State and proved Rayleigh's energy theory. *(A.M.I.E. Examination, 1999)*
25. Explain energy spectral density and state all of its properties.
26. Explain power spectral density and state all of its properties.
27. Define autocorrelation function of energy signals and prove all its properties.
28. Define autocorrelation function signals and state and prove all its properties.

NUMERICAL PROBLEMS

1. Find the trigonometric Fourier series for the waveform shown in figure 2.75.

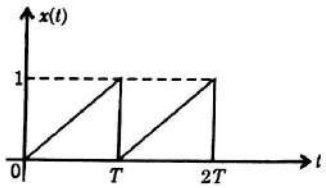


Fig. 2.75.

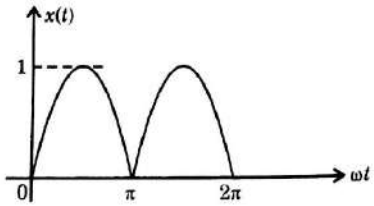


Fig. 2.76.

$$\left[\text{Ans. } x(t) = \frac{1}{2} - \frac{1}{\pi}\sum_{n=1}^{\infty}\frac{\sin(n\omega t)}{n}\right]$$

2. Find the trigonometric Fourier series for the full wave rectified sine wave shown in figure 2.76.

$$\left[\text{Ans. } x(t) = \frac{2}{\pi} - \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{\cos 2nt}{4n^2-1}\right]$$

3. Determine the Fourier transform of a gate pulse, of unit height, unit width and centred at $t = 0$.
4. Obtain the Fourier transform of the function *(A.M.I.E. Examination, 1998)*

$$x(t) = 5[u(t+3) + u(t+2) - u(t-2) - u(t-3)$$

$$\left[\text{Ans. } \frac{10}{\omega}(\sin 3\omega + \sin 2\omega)\right]$$

5. Find the Fourier transform of the following:
   (i) $x(2t)$, (ii) $x(t-2)\,e^{jt}$
6. Find the Fourier transform of the signal: *(RGP University, MP-2002)*

$$x(t) = e^{-b|t|}\cos\omega_0 t$$

7. Find the power spectral density of $x(t) = A\sin(2\pi f_0 t)$ and hence find total power.
8. Find the output responses of the .low pass RC network due to an input $x(t) = te^{-t/RC}$ by convolution. *(GATE Examination, 1999)*

$$\left[\text{Ans. } y(t) = \frac{1}{RC}\frac{t^2 e^{-t/RC}}{2}u(t)\right]$$

9. Find the output signal $y(t)$ when $B = A$ for a continuous-time system whose impulse response $h(t) = e - Bt\,u(t), A > 0$.

$$\left[\text{Ans. } y(t) = \frac{1}{(B-A)}\left[e^{-Btu(t)} \text{ for } A \neq B;\ y(t) = te^{-At}u(t) \text{ for } A = B\right]\right]$$

□□□

CHAPTER 3

Amplitude (Linear) Modulation

Inside this Chapter

3.1. Introduction 115
3.2. Multiplexing 115
3.3. Amplitude Modulation 116
3.4. Modulation Index 120
3.5. Single Tone Amplitude Modulation (AM) 121
3.6. Power Content in AM Wave 123
3.7. Transmission Efficiency of Amplitude Modulated Signal 124
3.8. Power of a Single Tone Amplitude Modulated (AM) Signal 125
3.9. Current Calculation for Single Tone AM 126
3.10. Power Content in Multiple-tone Amplitude Modulation (AM) 127
3.11. Generation of Amplitude Modulation (AM) 130
3.12. Collector Modulation Method 135
3.13. Demodulation of AM Wave 136
3.14. Double Sideband Suppressed Carrier (DSB-SC) System 139
3.15. Generation of DSB-SC Signal 141
3.16. Demodulation of DSB-SC 145
3.17. Quadrature-Amplitude Modulation (QAM) 148
3.18. Effect of Phase and Frequency Errors in Synchronous Detection 149
3.19. Carrier Acqwsition in DSB-SC System or Synchronization Techniques in DSB-SC system 150
3.20. Single sideband suppressed-Carrier (SSB-SC) Modulation 152
3.21. Hilbert Transform 155
3.22. SSB-SC for a General Modulating Signal 157
3.23. Generation of SSB-SC Signal 158
3.24. Demodulation of SSB-SC Signals 160
3.25. Vestigial Sideband (VSB) Modulation Systems 164
3.26. Generation of VSB Signals 165
3.27. Demodulation of VSB Signals 169
3.28. Comparison of Various AM Systems 170
3.29. Frequency Translation and Mixing 170

3.1. Introduction

As discussed earlier in Chapter 1, Modulation is a fundamental requirement of a communication system. Modulation may be defined as a process by which some characteristic of a signal known as carrier is varied according to the instantaneous value of another signal known as modulating signal. The signals containing intelligence or information to be transmitted are called modulating signals. These modulating signals containing information are also called baseband signals. Also the carrier frequency is greater than the modulating frequencies and the signal which results from the process of modulation is known as modulated signal.

Modulation may be classified as continuous wave modulation and pulse modulation. If the carrier waveform is continuous in nature then the modulation process is called as continuous-wave (CW) modulation. The examples of this type of modulation are amplitude modulation and angle modulation. On the other hand, if the carrier waveform is a pulse-type waveform, then the modulation process is called as pulse modulation. The examples of this type of modulation are Pulse-Amplitude Modulation (PAM), Pulse-Width Modulation (PWM), Pulse Code Modulation (PCM) etc.

Amplitude Modulation and Angle Modulation are the two families of continuous-wave (CW) modulation systems. In amplitude modulation, the amplitude of a sinusoidal carrier wave is varied in accordance with the baseband (modulating) signal. On the other hand, in angle modulation, the angle of the sinusoidal carrier wave is varied in accordance with baseband signal.

Amplitude modulation is discussed in this chapter, whereas the angle modulation will be discussed in next chapter.

3.2. Multiplexing

Multiplexing is a technique in which several message signals are combined into a composite signal for transmission over a common channel. In order to transmit a number of these signals over the same channel, the signals must be kept apart so that they do not interfere with each other, and hence they can be separated easily at the receiver end.

Basically, multiplexing is of two types as under:

(*i*) frequency division multiplexing (FDM)

(*ii*) time-division multiplexing (TDM)

(*i*) Frequency Division Multiplexing (FDM) ***(Important)***

The FDM scheme is illustrated in figure 3.1 with the simultaneous transmission of three message or baseband signals. The spectra of the message signals and the sum of the modulated carriers are indicated in the figure. DSB modulation is used in illustrating the spectra of figure 3.1. Any type of modulation can be used in FDM as long as the carrier spacing is sufficient to avoid spectral overlapping. However, the most widely used method of modulation is SSB modulation. At the receiving end of the channel the three modulated signals are separated by bandpass filters (BPFs) and then demodulated.

FDM is used in telephone system, telemetry, commercial broadcast, television, and communication networks. Commercial AM broadcast stations use carrier frequency spaced 10 kHz apart in the frequency range from 540 to 1640 kHz. This separation is not sufficient to avoid spectral overlap for AM with a reasonably high fidelity (50 Hz to 15 kHz) audio signal. Therefore, AM stations on adjacent carrier frequencies are placed geographically far apart to minimize interference. Commercial FM (Frequency-Modulation) broadcast uses carrier frequencies spaced 200 kHz apart. In a long-distance telephone system, up to 600 or more voice signals (200 Hz to 3.2 kHz) are transmitted over a coaxial cable or microwave links by using SSB modulation with carrier frequencies spaced 4 kHz apart. In practice, the composite signal formed by spacing several signals in frequency may, in turn, be modulated by using another carrier frequency. In this case, the first carrier frequencies are often called ***sub-carriers.***

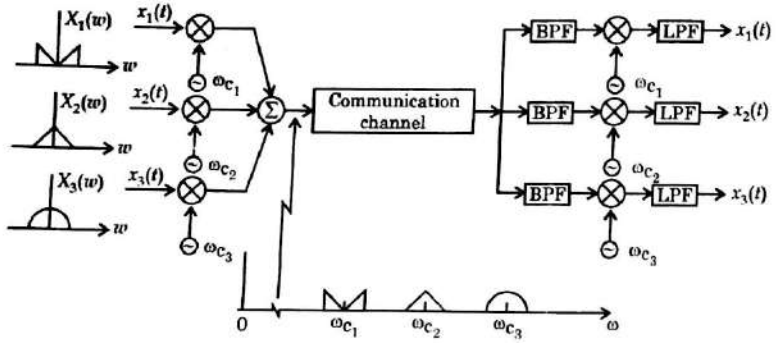


Fig. 3.1. *Frequency-division multiplexing.*

(ii) Time Division Multiplexing (TDM)

In case of Time Division Multiplexing (TDM), the complete channel bandwidth is allotted to one user for a fixed time slot. As an example, if there are ten users, then every user can be given the time slot of one second. Thus, complete channel can be used by each user for one second time in every ten seconds. This techniques is suitable for digital signals. Because digital signals are transmitted intermittently and the time spacing between two successive digital codewords can be utilized by other signals. There is possibility of crosstalk in FDM whereas Intersymbol Interference is possible in TDM. These problems can be overcome by some special cares.

3.3. Amplitude Modulation

Amplitude Modulation may be defined as a system in which the maximum amplitude of the carrier wave is made proportional to the instantaneous value (amplitude) of the modulating or baseband signal.

Let us consider a sinusoidal carrier wave $c(t)$ given as

$$c(t) = A \cos \omega_c t \qquad ...(3.1)$$

Here A is the maximum amplitude of the carrier wave and ω_c is the carrier frequency. For simplicity here we have assumed that the phase of the carrier wave is zero in equation (3.1).

Let $x(t)$ denotes the modulating or baseband signal, then according to amplitude modulation, the maximum amplitude A of the carrier will have to be made proportional to the instantaneous amplitude of modulating signal $x(t)$.

The standard equation for amplitude modulated (*AM*) wave may be expressed as

$$s(t) = x(t) \cos \omega_c t + A \cos \omega_c t \qquad ...(3.2)$$

or

$$s(t) = [A + x(t)] \cos \omega_c t \qquad ...(3.3)$$

Figure 3.2 shows the modulating signal or baseband signal, carrier signal and modulated signal.

Few Points about Amplitude Modulation (AM):

(*i*) It may be observed that equation (3.2) or equation (3.3) describes the time-domain behaviour of amplitude-modulated signal.

(*ii*) It may be noted that carrier signal [*i.e.* $c(t) = A \cos \omega_c t$] is a fixed frequency signal having frequency ω_c. The modulating or baseband signal $x(t)$ contains the information or intelligence to be transmitted. In the process of amplitude modulation, this information is superimposed upon the carrier signal in the form of amplitude variations of the carrier signal. This means that the information to be transmitted, is now, contained in the amplitude variations of the carrier signal. In other words, in amplitude modulation, the information is transmitted in the form of amplitude variations of the carrier signal.

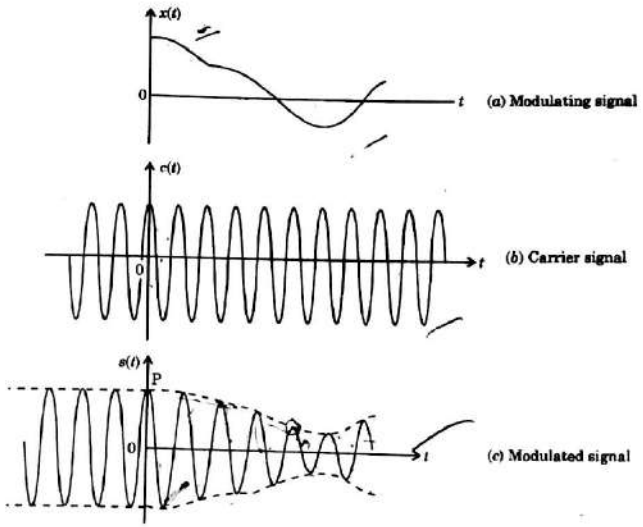


Fig. 3.2. *Illustration of amplitude modulation.*

(*iii*) The resulting signal from the process of amplitude-modulation is called amplitude modulated signal or simply *AM* wave.

(*iv*) In the process of amplitude modulation, the frequency and phase of the carrier remain constant whereas the maximum amplitude varies according to the instantaneous value of the information signal.

(*v*) Equation (3.3) represents an amplitude modulated wave. This wave has a constant frequency ω_c and amplitude $A + x(t)$. This implies that the amplitude of the wave is changing around A in accordance with the value of the modulating signal $x(t)$. The frequency of the *AM* signal remains unchanged and is equal to ω_c.

(*vi*) Figure 3.2 (*c*) illustrates the process of amplitude modulation. It may be observed that upto point P modulating signal is not applied so there is no modulation and maximum amplitude of the carrier remains constant at A. Now, at point P, the modulating signal is applied. Therefore, after point P, the amplitude modulation occurs. This means that the maximum amplitude A of the carrier now varies in accordance with the instantaneous value of the modulating signal $x(t)$.

(*vii*) The *AM* wave has a time-varying amplitude called as the envelope of the *AM* wave. Figure 3.2(*c*) shows that the envelope of *AM* wave consists of the modulating or baseband signal. This means that the unique property of *AM* is that the envelope of the modulated carrier has the same shape as the message signal or baseband signal.

We know that the expression for *AM* wave is

$$s(t) = [A + x(t)] \cos \omega_c t = E(t) \cos \omega_c t \qquad ...(3.4)$$

where

$$E(t) = A + x(t)$$

$E(t)$ is called the envelope of *AM* wave. This envelope consists of the baseband signal $x(t)$. Hence, the modulating or baseband signal may be recovered from an *AM* wave by detecting the envelope.

3.3.1. Spectrum of *AM* Wave or Frequency-Domain Representation

If $x(t)$ is modulating signal and carrier signal is given by the expression

$$c(t) = A \cos \omega_c t \qquad ...(3.5)$$

Then the equation for *AM* wave will be

$$s(t) = x(t) \cos \omega_c t + A \cos \omega_c t \qquad ...(3.6)$$

This equation describes the *AM* wave in time-domain. However, if we want to know the frequency description or frequencies present in *AM* wave, we will have to find its spectrum or frequency-domain representation. For this purpose, first we have to take the Fourier transform of *AM* wave.

Let $S(\omega)$ denote the Fourier transform of $s(t)$, $C(\omega)$ denotes the Fourier transform of $c(t)$ and $X(\omega)$ denotes the Fourier transform of $x(t)$.

Let the modulating signal $x(t)$ and its Fourier transform $X(\omega)$ be as shown in figure 3.3(*a*).

Let the modulating signal or message signal $x(t)$ be band limited to the interval $-\omega_m \le \omega \le \omega_m$ as shown in figure 3.3(*a*). This means that the modulating signal does not have any frequency component outside the interval $(-\omega_m, \omega_m)$.

It may be noted at this point that in figure 3.3(*a*) the modulating signal frequency ranges extend from $-\omega_m$ to ω_m *i.e.* it includes negative frequencies also from $-\omega_m$ to 0. Practically there is no meaning of negative frequency. In fact, the negative frequency is used for mathematical convenience only.

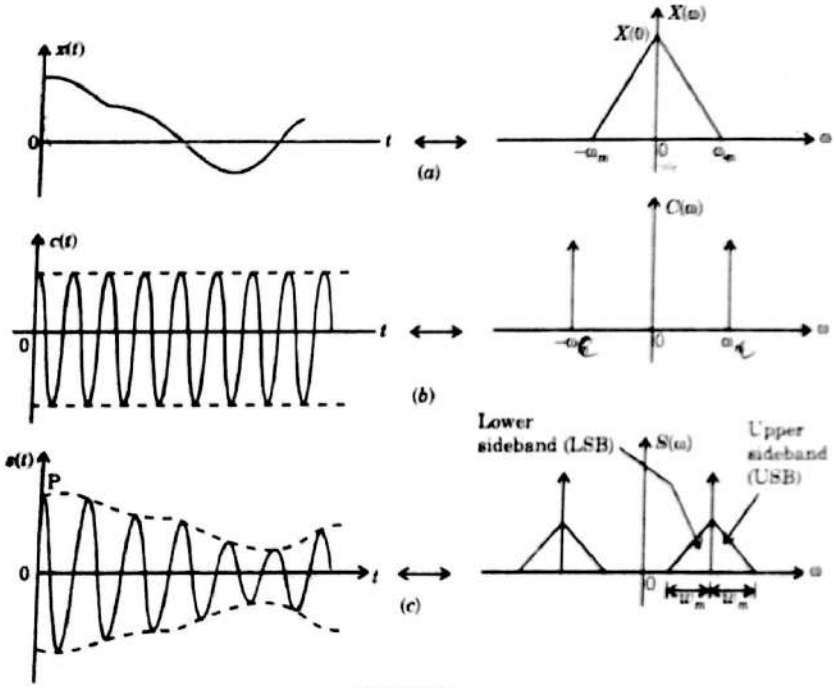


Fig. 3.3.

Hence, we can say that the modulating signal contains frequencies from 0 to ω_m or simply the bandwidth of modulating signal is ω_m. We know that the Fourier transform of a cosine signal $\cos \omega_c t$ consists of two impulse at ω_c and $-\omega_c$ as

$$\cos \omega_c t \leftrightarrow \pi[\delta(\omega + \omega_c) + \delta(\omega - \omega_c)] \qquad ...(3.7)$$

Since the carrier signal is $c(t) = A \cos \omega_c t$, therefore

$$A \cos \omega_c t \leftrightarrow \pi A[\delta(\omega + \omega_c) + \delta(\omega - \omega_c)] \qquad ...(3.8)$$

Figure 3.3(*b*) shows the carrier signal $A \cos \omega_c t$ and its Fourier transform.

Now, the *AM* wave is given as

$$s(t) = x(t) \cos \omega_c t + A \cos \omega_c t \qquad ...(3.9)$$

To find the spectrum of *AM* wave, we take its Fourier transform.

The Fourier transform of $s(t)$ may be found by considering the two factors $x(t) \cos \omega_c t$ and $A \cos \omega_c t$ separately as follow:

To find the Fourier transform of $x(t) \cos \omega_c t$, we first note the frequency-shifting theorem of Fourier transform as

If $\quad x(t) \leftrightarrow X(\omega)$

then $\quad e^{j\omega_c t} x(t) \leftrightarrow X(\omega - \omega_c) \qquad ...(3.10)$

This property states that if a signal $x(t)$ is multiplied by $e^{j\omega_c t}$ in time-domain, then its spectrum $X(\omega)$ in frequency-domain is shifted by an amount ω_c.

Similarly, $e^{-j\omega_c t} x(t) \leftrightarrow X(\omega + \omega_c) \qquad ...(3.11)$

But, since $e^{j\omega_c t}$ is not a real function and cannot be generated practically, therefore frequency shifting in practice is achieved by multiplying $x(t)$ by a sinusoid such as $\cos \omega_c t$.

Therefore,

$$x(t) \cos \omega_c t = x(t)\left[\frac{1}{2}(e^{j\omega_c t} + e^{-j\omega_c t})\right] \qquad ...(3.12)$$

$$x(t) \cos \omega_c t = \frac{1}{2} x(t) e^{j\omega_c t} + \frac{1}{2} x(t) e^{-j\omega_c t} \qquad ...(3.13)$$

Hence, using equations (3.10) and (3.11), we get

$$x(t) \cos \omega_c t \leftrightarrow \frac{1}{2}[X(\omega - \omega_c) + X(\omega + \omega_c)]. \qquad ...(3.14)$$

This means that the multiplication of a signal $x(t)$ by a sinusoid of frequency ω_c shifts the spectrum $X(\omega)$ by $\pm \omega_c$.

The Fourier transform of the second factor $A \cos \omega_c t$ will be as in equation (3.8)

$$A \cos \omega_c t \leftrightarrow \pi A[\delta(\omega + \omega_c) + \delta(\omega - \omega_c)] \qquad ...(3.15)$$

This means that Fourier transform of $A \cos \omega_c t$ consists of two impulses at $\pm \omega_c$.

Therefore, the Fourier transform of *AM* wave given by equation (3.6) will be the sum of equations (3.14) and (3.15).

$$\text{Thus,} \quad S(\omega) = \frac{1}{2}[X(\omega - \omega_c) + X(\omega + \omega_c)] + \pi A[\delta(\omega + \omega_c) + \delta(\omega - \omega_c)] \qquad ...(3.16)$$

Above equation for Amplitude Modulated Wave contains two factors. The first factor given as $1/2\ [X(\omega - \omega_c) + X(\omega + \omega_c)]$ represents the spectrum of original or baseband signal shifted in the positive as well as in the negative direction by the factor ω_c. On the other hand, the second factor given as $\pi A[\delta(\omega + \omega_c) + \delta(\omega - \omega_c)]$ represents the presence of carrier signal *i.e.*, two impulses each having strength equal to πA.

Thus the spectrum of modulated signal contains shifted spectrum of modulating signal and the spectrum of carrier signal as shown in figure 3.3(*c*).

From figure 3.3(*c*) following points may be observed:

(*i*) For positive frequencies, a portion of the spectrum of AM wave is lying above the carrier frequency ω_c. This band of frequency which is lying above the carrier frequency ω_c is known as the **upper sideband (USB)** whereas the symmetrical portion below carrier frequency ω_c is known as the **lower sideband (LSB)**. For negative frequencies, the upper sideband (USB) is represented by the portion of the spectrum below $-\omega_c$ and the lower sideband by the symmetrical portion above $-\omega_c$. Two sidebands have been shown in figure 3.3(*c*).

(*ii*) We generally keep $\omega_c > \omega_m$ which ensures that the two sidebands do not overlap each other.

(*iii*) At this point, it may be noted that the negative frequencies appeared in spectrum analysis is due to the exponential representation of Fourier transform. These negative frequencies are used for mathematical convenience. For a real function, these negative frequencies have no

meaning. For general purpose, it is sufficient to consider only the positive frequency region while treating the negative frequency region as a replica of the positive frequencies. This means that to calculate the BW of a signal, we consider only positive side.

(*iv*) From figure, it is obvious that for positive side, the highest frequency component present in the spectrum of AM wave is $\omega_c + \omega_m$ and the lowest frequency component is $\omega_c - \omega_m$.
Infact, the difference between these two extreme frequencies is equal to the bandwidth of the AM wave.

Therefore, Bandwidth $B = (\omega_c + \omega_m) - (\omega_c - \omega_m)$

$$B = 2\omega_m$$

Thus, it is clear that the bandwidth of the amplitude modulated wave is twice the highest frequency present in the baseband or modulating signal.

3.4. Modulation Index

In AM system the modulation index is defined as the measure of extent of amplitude variation about an unmodulated maximum carrier. It is represented by m_a.

Mathematically,

Modulation Index $$m_a = \frac{|x(t)|_{max}}{\text{Maximum carrier amplitude}}$$

or Modulation Index $$m_a = \frac{|x(t)|_{max}}{A} \quad ...(3.17)$$

where $|x(t)|_{max}$ represents the maximum amplitude of modulating signal and A represents the maximum amplitude of carrier signal.

The modulation index is also known as depth of modulation, degree of modulation or modulation factor.

Also, the absolute value of m_a multiplied by 100 is known as **percentage modulation.**

3.4.1. Overmodulation

We know that modulation index is given as

$$m_a = \frac{|x(t)|_{max}}{A} \quad ...(3.18)$$

where 'A' is the maximum amplitude of the carrier signal.

The baseband or modulating signal will be preserved in the envelope of the *AM* signal only if we have

$$|x(t)|_{max} \le A$$

i.e. modulation index m_a is less than or equal to unity.

In other words, the modulating signal is preserved in the envelope of *AM* signal only if the per centage modulation is less than or equal to 100 per cent.

On the other hand, if $m_a > 1$ or the per centage modulation is greater than 100, the baseband signal is not preserved in the envelope. It means that in this case, the baseband signal recovered from the envelope will be distorted. This type of distortion is called envelope distortion and the AM signal with $m_a > 1$ or $m_a > 100\%$ is called overmodulated signal.

Let us consider a baseband signal $x(t)$ and a carrier signal $A \cos \omega_c t$ as shown in figure 3.4(*a*) and 3.4(*b*). On Amplitude modulation, following two cases arise depending upon the amplitude of the baseband signal relative to the maximum amplitude of the carrier signal.

(*i*) Amplitude Modulation with $m_a < 1$

In this case, the waveform of the AM signal is as shown in figure 3.4(*c*). Here the maximum amplitude of baseband signal is less than maximum carrier amplitude A, *i.e.*,

$$|x(t)|_{max} < A$$

From figure, it may be observed that the envelope is not reaching the zero-amplitude axis of the *AM* waveform and so the baseband signal may be fully recovered from the envelope of AM wave.

(*ii*) Amplitude Modulation with $m_a > 1$

In this case, the waveform of the *AM* signal is as shown in figure 3.4(*d*). Here the amplitude of a baseband signal exceeds maximum carrier amplitude *i.e.*

$$|x(t)|_{max} < A$$

In this case, the modulation index m_a is more than 1 or 100%.

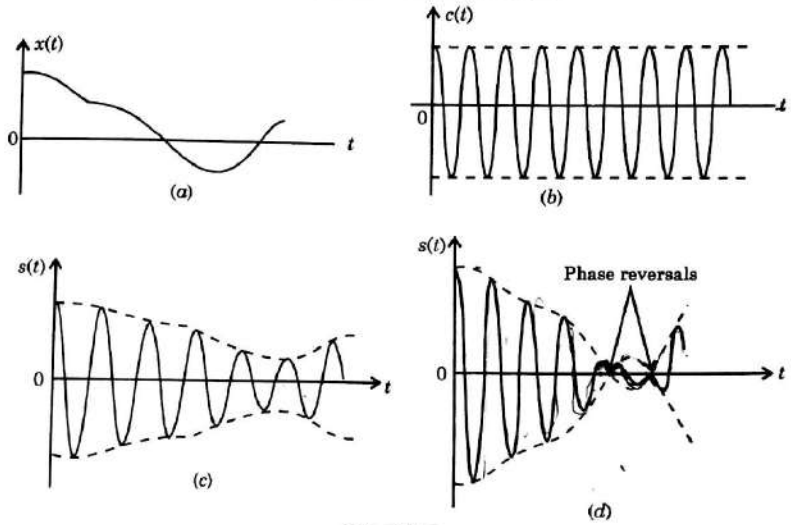


Fig. 3.4.

3.5. Single Tone Amplitude Modulation (*AM*) *(Important)*

Till now, we discussed amplitude modulation in which we assumed that baseband or modulating signal is a random signal which contains a large number of frequency components. This means that a carrier signal (fixed frequency signal) is modulated by a large number of frequency components.

In this section, we shall discuss amplitude modulation in which the modulating or baseband signal consists of only one (single) frequency *i.e.* modulation is done by a single frequency or tone. This type of amplitude modulation is known as **single tone amplitude modulation.**

Let us consider a single tone modulating signal as

$$x(t) = V_m \cos \omega_m t^* \quad ...(3.19)$$

which contains a single frequency ω_m.

Let the carrier signal be

$$c(t) = A \cos \omega_c t \quad ...(3.20)$$

We know that the general expression for *AM* signal is

$$s(t) = [A + x(t)] \cos \omega_c t$$

or $$s(t) = A \cos \omega_c t + x(t) \cos \omega_c t \quad ...(3.21)$$

Putting the value of $x(t)$, we get

$$s(t) = A \cos \omega_c t + V_m \cos \omega_m t \cos \omega_c t$$

or $$s(t) = A \cos \omega_c t + V_m \cos \omega_c t \cos \omega_m t$$

* Signal $x(t)$ may be a voltage signal or a current signal. Here, we have assumed that $x(t)$ is a voltage signal with maximum amplitude equal to V_m.

or $$s(t) = A\cos\omega_c t\left[1+\frac{V_m}{A}\cos\omega_m t\right] \quad ...(3.22)$$

But we know that for *AM*, the modulation index m_a is given as

$$m_a = \frac{|x(t)|_{max}}{A}$$

where $|x(t)|_{max}$ denotes the maximum amplitude of modulating signal and A is the maximum amplitude of the carrier signal.

In this case, we have

$$|x(t)|_{max} = V_m$$

Therefore, $$m_a = \frac{V_m}{A}$$

Putting this value of m_a in equation (3.22), we get

$$s(t) = A\cos\omega_c t\,[1 + m_a\cos\omega_m t] \quad ...(3.23)$$

This is the desired expression for single-tone modulating signal.

The expression in equation (3.23) may further be simplified to observe the frequency components present in *AM* signal.

$$s(t) = A\cos\omega_c t\,[1 + m_a\cos\omega_m t]$$

or $$s(t) = A\cos\omega_c t + A\cdot m_a\cos\omega_c t\cos\omega_m t$$

or $$s(t) = A\cos\omega_c t + \frac{A\cdot m_a}{2}[2\cos\omega_c t\cos\omega_m t]$$

$$= A\cos\omega_c t + \frac{A\cdot m_a}{2}[\cos(\omega_c t+\omega_m)t + \cos(\omega_c-\omega_m)t]$$

or $$s(t) = A\cos\omega_c t + \frac{A\cdot m_a}{2}\cos(\omega_c+\omega_m)t + \frac{A\cdot m_a}{2}\cos(\omega_c-\omega_m)t \quad ...(3.24)$$

Equation (3.24) reveals that the *AM* signal has three frequency components as follow:

(*i*) carrier frequency ω_c having amplitude A

(*ii*) upper sideband $(\omega_c + \omega_m)$ having amplitude $\frac{m_a\cdot A}{2}$

(*iii*) lower sideband $(\omega_c - \omega_m)$ having amplitude $\frac{m_a\cdot A}{2}$

With the help of these frequency components, we can plot the frequency-spectrum of single-tone amplitude modulated (AM) wave. Figure 3.4*a* shows the one-sided frequency spectrum of single-tone AM wave.

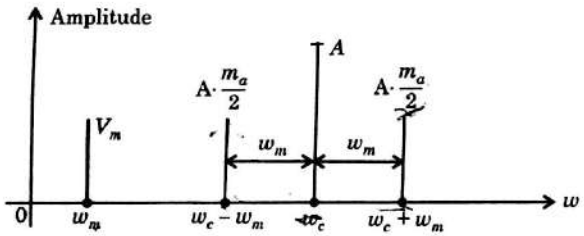


Fig. 3.4*a*. *Single-sided frequency spectrum of single-tone AM wave.*

Example 3.1. The tuned-circuit of the oscillator in an AM transmitter uses a 50 μH coil and a 1 nF capacitor. Now, if the oscillator output is modulated by audio frequencies upto 8 kHz, then find the frequency range occupied by the sidebands.

Solution: The oscillator in AM transmitter is used to generate high carrier frequency. Hence, the resonance frequency of the oscillator will be the carrier frequency.

Therefore,

Carrier frequency, $$f_c = \frac{1}{2\pi\sqrt{LC}}$$

Here given that $L = 50\ \mu H$

$L = 50\times10^{-6}$ H

and $C = 1\ nF = 1\times10^{-9}$ F

Thus, $$f_c = \frac{1}{2\pi\sqrt{50\times10^{-6}\times1\times10^{-9}}} = \frac{1}{2\pi\sqrt{5\times10^{-14}}} = \frac{1}{2\pi\times10^{-7}\times\sqrt{5}}$$

$$f_c = 7.12\times10^5\ \text{Hz} = 712\ \text{kHz}$$

Now, it is given that the highest modulating frequency is 8 kHz.

Therefore, the frequency range occupied by the sidebands will range from 8 kHz above to 8 kHz below the carrier frequency, extending from 7 to 4 kHz to 720 kHz. **Ans.**

3.6. Power Content in AM Wave

It may be observed from the expression of *AM* wave that the carrier component of the amplitude modulated wave has the same amplitude as unmodulated carrier. In addition to carrier component, the modulated wave consists of two sideband components. It means that the modulated wave contains more power than the unmodulated carrier. However, since the amplitudes of two sidebands depend upon the modulation index, it may be anticipated that the total power of the amplitude modulated wave would depend upon the modulation index also. In this section, we shall find the power contents of the carrier and the sidebands.

We know that the general expression of *AM* wave is given as

$$s(t) = A\cos\omega_c t + x(t)\cos\omega_c t \quad ...(3.25)$$

The total power P of the *AM* wave is the sum of the carrier power P_c and sideband power P_s.

Carrier Power

The carrier power P_t is equal to the mean-square (*ms*) value of the carrier term $A\cos\omega_c t$ *i.e.*

$$P_c = \text{mean square value of } A\cos\omega_c t$$

$$P_c = [A\cos\omega_c t]^2 = \frac{1}{2\pi}\int_0^{2\pi} A^2\cos^2\omega_c t.\,dt^* = \frac{A^2}{2} \quad ...(3.26)$$

Sideband Power

The sideband power P_s is equal to the mean square value of the sideband term $x(t)\cos\omega_c t$, *i.e.*

$$P_s = \text{mean square value of } x(t)\cos\omega_c t$$

$$P_s = [x(t)\cos\omega_c t]^2 = \frac{1}{2\pi}\int_0^{2\pi} x^2(t)\cos^2\omega_c t\,dt$$

$$P_s = \frac{1}{2\pi}\int_0^{2\pi}\frac{1}{2}\left[2\cos^2\omega_c t\right]x^2(t)\,dt$$

or $$P_s = \frac{1}{2\pi}\int_0^{2\pi}\frac{1}{2}x^2(t)\,dt + \frac{1}{2\pi}\int_0^{2\pi}x^2(t)\cos2\omega_c t\,dt \quad ...(3.27)$$

* Since period of the signal $A\cos\omega_c t$ is 2π.

In AM generation, a bandpass filter (BPF) or a tuned circuit tuned to carrier frequency ω_c is used to filter out the second integral term.

Therefore,

$$P_s = \frac{1}{2\pi}\int_0^{2\pi}\left[\frac{1}{2}x^2(t)\right]dt$$

or $\quad P_s$ = mean square (*ms*) value of $\frac{1}{2}x^2(t)$

or $$P_s = \frac{1}{2}\overline{x^2(t)} \qquad ...(3.28)$$

However, the total sideband power P_s is due to the equal contributions of the upper and lower sidebands. Hence, the power carried by the upper and the lower sidebands will be

$$P_{s\,(LSB)} = P_{s(USB)} = \frac{P_s}{2} = \frac{1}{4}\overline{x^2(t)} \qquad ...(3.29)$$

Therefore, the total power P_t of the *AM* signal is the sum of the carrier power P_c and sideband power P_s.

$$P_t = P_c + P_s = \frac{1}{2}A^2 + \frac{1}{2}\overline{x^2(t)}$$

$$P_t = \frac{1}{2}[A^2 + \overline{x^2(t)}] \qquad ...(3.30)$$

3.7. Transmission Efficiency of Amplitude Modulated Signal

We know that the total modulated power of an *AM* signal is expressed as

$$P_t = P_c + P_s = \frac{1}{2}[A^2 + \overline{x^2(t)}] \qquad ...(3.31)$$

Out of this total power P_t, the useful message or baseband power is the power carried by the sidebands, *i.e.* P_s. The large carrier power P_c is a waste from the transmission point of view because it does not carry any information or message. This large carrier power P_c is transmitted alongwith the sideband power only for the convenient and cheap detection. Hence, P_s is the only useful message power present in the AM wave.

In AM wave, the amount of useful message power P_s may be expressed by a term known as **transmission efficiency η.**

Hence transmission efficiency of AM wave may be defined as the per centage of total power contributed by the sidebands.

Mathematically,

Transmission efficiency, $$\eta = \frac{P_s}{P_t}\times 100 \qquad ...(3.32)$$

or $$\eta = \frac{\frac{1}{2}\overline{x^2(t)}}{\frac{1}{2}A^2 + \frac{1}{2}\overline{x^2(t)}}\times 100$$

or $$\eta = \frac{100\,\overline{x^2(t)}}{A^2 + \overline{x^2(t)}} \qquad ...(3.33)$$

The maximum transmission efficiency of the *AM* is only 33.33%. This implies that only one-third of the total power is carried by the sidebands and the rest two-third is wasted.

3.8. Power of a Single-Tone Amplitude-Modulated (AM) Signal

In article 3.6, we have found the power content of the AM signal when modulating-signal is any random signal and may consist of several frequency components. Likewise we can find power content of single-tone Amplitude Modulated (AM) signal.

Let us consider that a carrier signal $A\cos\omega_c t$ is amplitude-modulated by a single-tone modulating signal $x(t) = V_m \cos\omega_m t$.

Then the unmodulated or carrier power

P_c = mean square (*ms* value)

$$P_c = \overline{(A\cos\omega_c t)^2} = \frac{A^2}{2}$$

The sideband power $$P_s = \frac{1}{2}\overline{x^2(t)} = \frac{1}{2}\overline{(V_m\cos\omega_m t)^2}$$

$$P_s = \frac{1}{2}\frac{V_m^2}{2} = \frac{1}{4}V_m^2 \qquad ...(3.34)$$

We know that the total modulated power P_t is the sum of P_c and P_s.

Therefore

$$P_t = P_c + P_s = \frac{A^2}{2} + \frac{1}{4}\cdot V_m^2$$

or $$P_t = \frac{A^2}{2}\left[1 + \frac{1}{2}\left(\frac{V_m}{A}\right)^2\right]$$

But $$\frac{V_m}{A} = \frac{\text{Maximum baseband amplitude}}{\text{Maximum carrier amplitude}} = m_a$$

= modulation index for *AM*

Hence

$$P_t = \frac{A^2}{2}\left[1 + \frac{1}{2}\cdot m_a^2\right]$$

But $$\frac{A^2}{2} = P_c = \text{carrier power}$$

Therefore, $$P_t = P_c\left(1 + \frac{m_a^2}{2}\right) \qquad ...(3.35)$$

Example 3.2. A 400 watts carrier is modulated to a depth of 75 per cent. Find the total power in the amplitude-modulated wave. Assume the modulating signal to be a sinusoidal one.

Solution: We know that for a sinusoidal modulating signal, the total power is expressed as

$$P_t = P_c\left(1 + \frac{m_a^2}{2}\right)$$

where P_t = total power or modulated power

P_c = carrier power or unmodulated power

m_a = modulation index

Given that, P_c = 400 watts

m_a = 75 per cent = 0.75

Therefore, $$P_t = P_c\left(1+\frac{m_a^2}{2}\right) = 400\left(1+\frac{0.75^2}{2}\right)$$

$$P_t = 512.5 \text{ watts} \qquad \textbf{Ans.}$$

Example 3.3. An AM broadcast radio transfer radiates 10 K watts of power if modulation percentage is 60. Calculate how much of this is the carrier power.

Solution: We know that the total power is expressed as

$$P_t = P_c\left(1+\frac{m_a^2}{2}\right) \qquad ...(i)$$

where P_t = total power or modulated power
P_c = carrier power or unmodulated power
m_a = modulation index

Given that,

P_t = 10 K watts
m_a = 60 per cent = 0.6

From equation (*i*), we get

$$P_c = \frac{P_t}{1+\frac{m_a^2}{2}} = \frac{10}{1+\frac{0.6^2}{2}} = \frac{10}{1.18}$$

$$P_c = 8.47 \text{ kW} \qquad \textbf{Ans.}$$

3.9. Current Calculation for Single-Tone AM

In AM, it is generally more convenient to measure the *AM* transmitter current than the power. In this case, the modulation index may be calculated from the values of unmodulated and modulated currents in the AM transmitter.

Let I_c be the r.m.s. value of the carrier or unmodulated current and I_t be the r.m.s. value of the total or modulated current of an *AM* transmitter. Let R be the antenna resistance through which these currents flow.

Now, we know that for a single-tone modulation the power relation is expressed as

$$P_t = P_c\left(1+\frac{m_a^2}{2}\right) \qquad ...(3.36)$$

where P_t = total or modulated power
P_c = carrier or unmodulated power
m_a = modulation index

From equation (3.36), we may write

$$\frac{P_t}{P_c} = 1+\frac{m_a^2}{2}$$

or $$\frac{I_t^2 \cdot R}{I_c^2 \cdot R} = 1+\frac{m_a^2}{2}$$

or $$\frac{I_t}{I_c} = \sqrt{1+\frac{m_a^2}{2}}$$

or $$I_t = I_c\sqrt{1+\frac{m_a^2}{2}} \qquad ...(3.37)$$

Example 3.4. The antenna current of an AM transmitter is 8 A, if only the carrier is sent, but it increases to 8.93 A, if the carrier is modulated by a single sinusoidal wave. Determine the percentage modulation. Also find the antenna current if the per cent of modulation changes to 0.8. *(U.P. Tech. Sem., Exam., 2004-05) (06 marks)*

Solution: (*i*) The current relation for a single-tone amplitude modulation is expressed as

$$I_t = I_c\sqrt{1+\frac{m_a^2}{2}} \qquad ...(i)$$

where I_t = total or modulated current
I_c = carrier or unmodulated current
m_a = modulation index

Using equation (*i*), we get

$$\frac{I_t}{I_c} = \sqrt{1+\frac{m_a^2}{2}}$$

or $$\left(\frac{I_t}{I_c}\right)^2 = 1+\frac{m_a^2}{2}$$

or $$\frac{m_a^2}{2} = \left(\frac{I_t}{I_c}\right)^2 - 1$$

or $$m_a^2 = 2\left[\left(\frac{I_t}{I_c}\right)^2 - 1\right]$$

or $$m_a = \sqrt{2\left[\left(\frac{I_t}{I_c}\right)^2 - 1\right]}$$

Putting all the given values, we have

$$m_a = \sqrt{2\left[\left(\frac{8.93}{8}\right)^2 - 1\right]} = \sqrt{2\left[(1.116)^2 - 1\right]}$$

$$m_a = \sqrt{2(1.246-1)} = \sqrt{0.492} = 0.701 = 70.1\%$$

(*ii*) Since $$I_t = I_c\sqrt{1+\frac{m_a^2}{2}}$$

Here, $I_c = 8$ A and $m_a = 0.8$

Therefore, $$I_t = 8\times\sqrt{1+\frac{0.8^2}{2}} = 8\sqrt{1+\frac{0.64}{2}}$$

or $$I_t = 8\sqrt{1.32} = 8\times 1.149 = 9.19 \text{ A} \qquad \textbf{Ans.}$$

3.10. Power Content in Multiple-Tone Amplitude Modulation (AM)

A multiple-tone amplitude modulation is that type of modulation in which the modulating signal consists of more than one frequency components.

Let us consider that a carrier signal $A\cos\omega_c t$ is modulated by a baseband or modulating signal $x(t)$ which is expressed as

$$x(t) = V_1\cos\omega_1 t + V_2\cos\omega_2 t + V_3\cos\omega_3 t \qquad ...(3.38)$$

We know that the general expression for AM wave is

$$s(t) = A\cos\omega_c t + x(t)\cos\omega_c t \quad ...(3.39)$$

Putting the value of $x(t)$, we get

$$s(t) = A\cos\omega_c t + [V_1\cos\omega_1 t + V_2\cos\omega_2 t + V_3\cos\omega_3 t]\cos\omega_c t$$

or

$$s(t) = A\left[1+\frac{V_1}{A}\cos\omega_1 t + \frac{V_2}{A}\cos\omega_2 t + \frac{V_3}{A}\cos\omega_3 t\right]\cos\omega_c t \quad ...(3.40)$$

But we know that

$$\frac{V}{A} = \frac{\text{Maximum amplitude of modulating signal}}{\text{Maximum amplitude of carrier signal}} = \text{modulation index } m_a$$

Therefore,

$$s(t) = A[1 + m_1\cos\omega_1 t + m_2\cos\omega_2 t + m_3\cos\omega_3 t]\cos\omega_c t \quad ...(3.41)$$

where

$m_1 = \frac{V_1}{A}$, $m_2 = \frac{V_2}{A}$ and $m_3 = \frac{V_3}{A}$, are the modulation indexes of the corresponding frequency components.

The expression for AM wave in equation (3.41) may further be expanded as

$$s(t) = A\cos\omega_c t + m_1 A\cos\omega_c t\cos\omega_1 t + m_2 A\cos\omega_c t\cos\omega_2 t + m_3 A\cos\omega_c t\cos\omega_3 t$$

Now we know that the total power in *AM* is given as

$$P_t = \text{carrier power} + \text{sideband power}$$

$$P_t = P_c + P_s \quad ...(3.42)$$

The carrier power P_c is given as

$$P_c = (A\cos\omega_c t)^2 = \frac{A^2}{2}$$

Similarly, the sideband power is given as

$$P_s = \frac{1}{2}\overline{x^2(t)}$$

$$P_s = \frac{1}{2}[\overline{(V_1\cos\omega_1 t)^2} + \overline{(V_2\cos\omega_2 t)^2} + \overline{(V_3\cos\omega_3 t)^2}] \quad ...(3.43)$$

But we know that

$$m_1 = \frac{V_1}{A} \text{ so that } V_2 = m_1 A$$

$$m_2 = \frac{V_2}{A} \text{ so that } V_2 = m_2 A$$

and

$$m_3 = \frac{V_3}{A} \text{ so that } V_3 = m_3 A$$

Putting the values of V_1, V_2 and V_3 in equation (3.43), we get

$$P_s = \frac{1}{2}[\overline{(m_1 A\cos\omega_1 t)^2} + \overline{(m_1 A\cos\omega_1 t)^2} + \overline{(m_3 A\cos\omega_3 t)^2}]$$

$$P_s = \frac{1}{2}\left[\frac{m_1 A^2}{2} + \frac{m_2 A^2}{2} + \frac{m_3 A^2}{2}\right]$$

or

$$P_s = \frac{1}{4}A^2[m_1^2 + m_2^2 + m_3^2] \quad ...(3.44)$$

Now putting the value of P_c and P_s in equation (3.41), we get

$$P_t = P_c + P_s = \frac{A^2}{2} + \frac{1}{4}A^2[m_1^2 + m_2^2 + m_3^2]$$

or

$$P_t = \frac{A^2}{2}\left[1+\frac{1}{2}(m_1^2 + m_2^2 + m_3^2)\right] = P_c\left[1+\frac{m_1^2 + m_2^2 + m_3^2}{2}\right]$$

or

$$P_t = P_c\left[1+\frac{m_1^2}{2}+\frac{m_2^2}{2}+\frac{m_3^2}{2}\right]$$

This expression may be extended upto to n-modulating terms *i.e.*

$$P_t = P_c\left[1+\frac{m_1^2}{2}+\frac{m_2^2}{2}+\frac{m_3^2}{2}+.....+\frac{m_n^2}{2}\right] \quad ...(3.45)$$

3.10.1. Total or Net Modulation Index for Multiple-Tone Modulation

Let us consider that m_t is the total or net modulation indexes for a multiple-tone modulation. We know that for a multiple-tone modulation, the total power is expressed as

$$P_t = P_c\left[1+\frac{m_1^2}{2}+\frac{m_2^2}{2}+\frac{m_3^2}{2}+.....+\frac{m_n^2}{2}\right] \quad ...(3.46)$$

where m_1, m_2 .... m_n are the modulation indexes for different modulating signals.

The power for *AM* wave is also expressed as

$$P_t = P_c\left(1+\frac{m_t^2}{2}\right) \quad ...(3.47)$$

Comparing equations (3.46) and (3.47), we get

$$m_t^2 = m_1^2 + m_2^2 + m_3^2 + ... + m_n^2$$

or

$$m_t = \sqrt{m_1^2 + m_2^2 + m_3^2 + ... + m_n^2} \quad ...(3.48)$$

This is the desired expression for the total or net modulation index.

Example 3.5. An AM transmitter radiates 9 K watts of power when the carrier is unmodulated and 10.125 K watts when the carrier is sinusoidally modulated. Find the modulation index, per centage of modulation. Now, if another sine wave, corresponding to 40 per cent modulation is transmitted simultaneously, then calculate the total radiated power. ***(Very Important)***

Solution: (*i*) We know that for a single-tone sinusoidal amplitude-modulation the total power is expressed as

$$P_t = P_c\left(1+\frac{m_a^2}{2}\right) \quad ...(i)$$

where P_t = modulated or total power

P_c = unmodulated or carrier power

m_a = modulation index

Given that, $P_t = 10.125$ kW

and $P_c = 9$ kW

Using equation (*i*), we get

$$1+\frac{m_a^2}{2} = \frac{P_t}{P_c}$$

or $$\frac{m_a^2}{2} = \frac{P_t}{P_c} - 1$$

or $$\frac{m_a^2}{2} = \frac{10.125}{9} - 1 = 1.125 - 1 = 0.125$$

or $$m_a^2 = 0.125 \times 2 = 0.250 = 0.50$$

(*ii*) We know that in case of modulation by two sinusoidal waves, the total modulation index m_t is expressed as

$$m_t = \sqrt{m_1^2 + m_2^2}$$

Let $m_1 = m_a = 0.5$

Given that $m_2 = 0.4$

Therefore, $$m_t = \sqrt{(0.5)^2 + (0.4)^2} = \sqrt{0.25 + 0.16} = \sqrt{0.41} = 0.64$$

The total radiated power in this case will be

$$P_t = P_c\left(1 + \frac{m_t^2}{2}\right) = 9 \times \left(1 + \frac{0.64^2}{2}\right)$$

$$P_t = 9\,(1 + 0.205) = 10.84 \text{ kW}$$ Ans.

3.11. Generation of Amplitude Modulation (AM)

The device which is used to generate an amplitude modulated (AM) wave is known as Amplitude Modulator.

The methods of AM Generation may be broadly classified as follow:

(*i*) Low-level AM Modulation

(*ii*) High-level AM Modulation

3.11.1. Low Level Amplitude Modulation

Figure 3.5(*a*) shows the block diagram of a low level AM modulation system. In a low-level amplitude modulation system, the modulation is done at low power level. At low power levels, a very small power is associated with the carrier signal and the modulating signal. Because of this, the output power of modulation is low. Therefore, the power amplifiers are required to boost the amplitude-modulated signals upto the desired output level. From block diagram in figure 3.5(*a*), it is clear that modulation is done at low power level. After this, the amplitude-modulated signal (*i.e.* a signal containing a carrier and two sidebands) is applied to a wideband power amplifier. A wideband power amplifier is used just to preserve the sidebands of the modulated signal. Amplitude modulated systems, employing modulation at low power levels are also called **low-level amplitude modulation transmitters.**

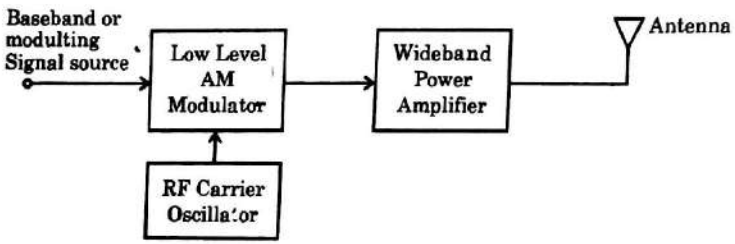


Fig. 3.5. *(a) Block Diagram for Low level AM Modulation.*

Square-low diode modulation and switching modulation are examples of low-level modulation.

3.11.2. High Level Amplitude Modulation

Figure 3.5(*b*) shows the block diagram of a high level AM modulation system. In a high-level amplitude modulation system, the modulation is done at high power level. Therefore, to produce amplitude-modulation at these high power levels, the baseband signal and the carrier signal must be at high power levels. In block diagram of figure 3.5(*b*) the modulating signal and carrier signal are first power amplified and then applied to AM high-level modulator. For modulating signal, the wide-band power amplifier is required just to preserve all the frequency components present in modulating signal. On the other hand, for carrier signal, the narrow band power amplifier is required because it is a fixed-frequency signal. The collector modulation method is the example of high-level modulation.

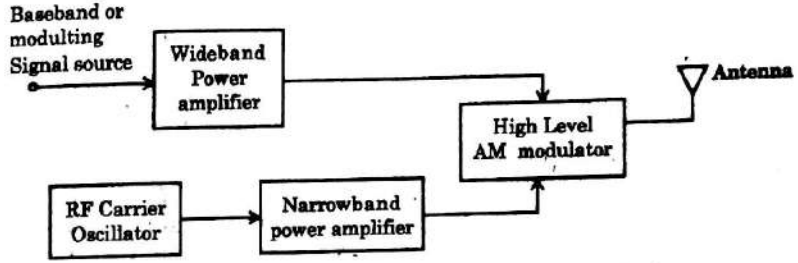


Fig. 3.5. *(b) Block diagram for high level AM Modulation.*

Before we discuss low level and high level modulation methods in detail, we shall establish the fact that a non-linear resistance of a non-linear device can be made to produce Amplitude Modulation when two different frequencies are passed together through it.

3.11.3. Non-linear resistance or Non-linear Circuits

We know that the relationship between voltage and current in a linear resistance is expressed as

$$i = bv \qquad ...(3.49)$$

where

v = voltage across the linear resistance

i = current through linear resistance

and b = any constant of proportionality

If equation (3.49) is applied to a resistor, then constant b is clearly its conductance.

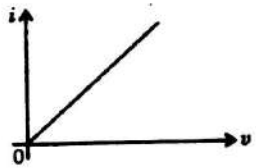


Fig. 3.6. *Linear resistance characteristics*

Also, if equation (3.49) is applied to the linear portion of the transistor characteristic then i is the collector current and v is the voltage applied to the base.

As more general, equation (3.49), may be written as

$$i = a + bv \qquad ...(3.50)$$

where a is the d.c. component of the current.

Now, let us consider a non-linear resistance. For a non-linear resistance the current-voltage characteristics will be non-linear as shown in figure 3.7.

The non-linear relationship between voltage and current may be expressed as

$$i = a + bv + cv^2 + dv^3 + ... \qquad ...(3.51)$$

This means that due to non-linearity in the v-i characteristics of a non-linear resistance, the current becomes proportional not only to voltage but also to the square, cube and higher powers of the voltage.

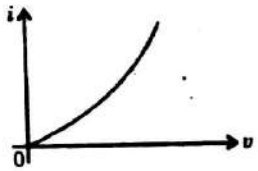


Fig. 3.7. *Non-linear resistance characteristics*

For simplicity, neglecting the higher terms in equation (3.51), we have

$$i = a + bv + cv^2 \qquad ...(3.52)$$

Devices or circuits having non-linear v-i characteristics can be treated as a non-linear resistance. For example, devices like diodes, transistors, *FET* etc. exhibit non-linear characteristics and hence work as a non-linear resistance.

Equation (3.52), may be used in relating the output current to the input voltage of a non-linear resistance.

We can apply this equation to the gate-voltage-drain-current characteristics of a *FET*, then

$$i = a + bv + cv^2$$

If two voltages are applied at gate simultaneously, then

$$i = a + b(v_1 + v_2) + c(v_1 + v_2)^2$$

$$i = a + b(v_1 + v_2) + c(v_1^2 + v_2^2 + 2\,v_1 v_2) \qquad ...(3.53)$$

Let us consider that the two input voltages are sinusoidal having different frequencies, then,

$$v_1 = V_1 \sin \omega t$$

and $$v_2 = V_2 \sin pt$$

Here ω and p are the two different frequencies.

Putting the values of v_1 and v_2 in equation (3.53), we get

$$i = a + b(V_1 \sin \omega t + V_2 \sin pt) + c(V_1^2 \sin^2 \omega t + C_2^2 \sin^2 pt + 2V_1 V_2 \sin \omega t \sin pt)$$

$$i = a + bV_1 \sin \omega t + bV_2 \sin pt + cV_1^2 \sin^2 \omega t + cV_2^2 \sin^2 pt + 2cV_1 V_2 \sin \omega t \sin pt$$

or $$i = a + bV_1 \sin \omega t + bV_2 \sin pt + \frac{1}{2} cV_1^2 (2\sin^2 \omega t) + \frac{1}{2} cV_2^2 (2\sin^2 pt) + cV_1 V_2 (2\sin \omega t \sin pt)$$

$$i = a + bV_1 \sin \omega t + bV_2 \sin pt + \frac{1}{2} cV_1^2 (1 - \cos 2\omega t) + \frac{1}{2} cV_2^2 (1 - \cos 2pt)$$

$$+ cV_1 V_2 [\cos(\omega - p)t - \cos(\omega + p)t \qquad ...(3.49)$$

or $$i = a + bV_1 \sin \omega t + bV_2 \sin pt + \frac{1}{2} cV_1^2 - \frac{1}{2} cV_1^2 \cos 2\omega t + \frac{1}{2} cV_2^2$$

$$- \frac{1}{2} cV_2^2 \cos 2pt + cV_1 V_2 \cos(\omega - p)t - cV_1 V_2 \cos(\omega + p)t$$

or $$i = \underbrace{\left(a + \frac{1}{2} cV_1^2 + \frac{1}{2} cV_2^2\right)}_{(1)} + \underbrace{bV_1 \sin \omega t}_{(2)} + \underbrace{bV_2 \sin pt}_{(3)}$$

$$- \underbrace{\left(\frac{1}{2} cV_1^2 \cos 2\omega t + \frac{1}{2} cV_2^2 \cos 2pt\right)}_{(4)} + \underbrace{cV_1 V_2 \cos(\omega - p)t}_{(5)} - \underbrace{cV_1 V_2 \cos(w + p)t}_{(6)} \qquad ...(3.54)$$

In equation (3.54), if ω is assumed as the carrier frequency and p is assumed as modulating frequency then we can identify the different terms as follow:

term (1) is the d.c. component
term (2) is the carrier signal
term (3) is the modulating signal
term (4) contains harmonics of the carrier signal and the modulating signal
term (5) is the lower sideband
term (6) is the upper sideband

Since the process of amplitude modulation consists of carrier signal, lower sideband and upper sideband, therefore we can conclude from above that when two voltages of different frequencies are passed through a non-linear resistance, the Amplitude Modulation takes place.

In a practical modulation circuit, the frequencies other than the carrier and two sideband frequencies are rejected with the help of a tuned circuit.

3.11.4. Square Law Diode Modulation

Square law diode modulation circuit make use of nonlinear current-voltage characteristics of diode.* This method is suited at low voltage levels because of the fact that current-voltage chara-cteristic of a diode is highly nonlinear particularly in the low voltage region as shown in figure 3.8.

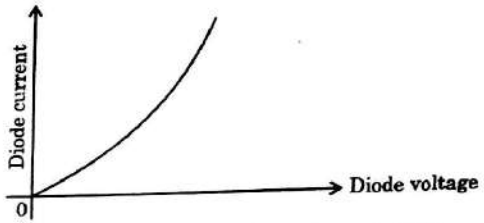


Fig. 3.8. *Current-voltage characteristic of a diode.*

Figure 3.9 shows the circuit of square law diode modulation.

It may be observed from the figure 3.9, that carrier and modulating signals are applied across the diode. A d.c. battery V_{cc} is connected across the diode to get a fixed operating point on the v–i characteristics of diode. The working of this circuit may be explained by considering the fact when two different frequencies are passed through a non-linear device, the process of amplitude modulation takes place. Hence, when carrier and modulating frequencies are applied at the input of diode, then different frequency terms appear at the output of diode. These different frequency terms are applied across a tuned circuit which is tuned to the carrier frequency and has a narrow bandwidth just to pass two sidebands alongwith the carrier and reject other frequencies. Hence at the output of tuned circuit, carrier and two sidebands are obtained *i.e.*, Amplitude Modulated (AM) wave is produced.

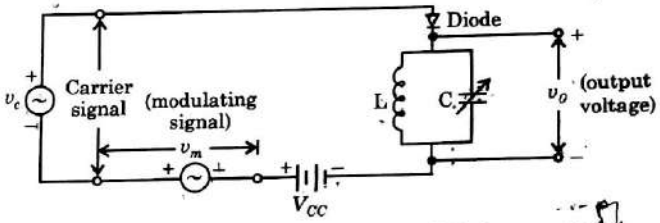


Fig. 3.9. *Square-low diode modulation.*

Mathematical Analysis

Let us consider that carrier voltage is expressed as

$$v_c = V_c \cos \omega_c t \qquad ...(3.55)$$

where ω_c is the carrier frequency.

Let the modulating voltage be expressed as

$$v_m = V_m \cos \omega_m t \qquad ...(3.56)$$

where w_m is the modulating frequency.

The total a.c. voltage across the diode is given as

$$v_s = v_c + v_m \qquad ...(3.57)$$

$$v_s = V_c \cos \omega_c t + V_m \cos \omega_m t \qquad ...(3.58)$$

The non-linear relationship between voltage and current for a diode is expressed as

$$i = a + bv_s + cv_s^2 \qquad ...(3.59)$$

where a, b, and c are constants

i = current through the diode

v_s = voltage across the diode

* Explain the working of a modulator for generating AM wave.
(U.P. Tech, Sem. Exam., 2003-04) (05 marks)

Putting the value of v_s from equation (3.58) in equation (3.59), we get

$$i = a + bv_s + cv_s^2 = a + b\,(V_c \cos \omega_c t + V_m \cos \omega_m t) + c(V_c \cos \omega_c t + V_m \cos \omega_m t)^2$$

or $$i = a + bV_c \cos \omega_c t + bV_m \cos \omega_m t + c(V_c^2 \cos^2 \omega_c t + V_m^2 \cos^2 \omega_m t + 2V_cV_m \cos \omega_c t \cos \omega_m t)$$

or $$i = a + bV_c \cos \omega_c t + bV_m \cos \omega_m t + cV_c^2 \cos^2 \omega_c t + cV_m^2 \cos^2 \omega_m t + 2c.V_c.V_m \cos \omega_c t \cos \omega_m t$$

or $$i = a + bV_c \cos \omega_c t + bV_m \cos \omega_m t + \frac{1}{2} cV_c^2 (2 \cos^2 \omega_c t)$$

$$+\frac{1}{2} cV_m^2 (2 \cos^2 \omega_m t) + cV_c V_m (2 \cos \omega_c t \cos \omega_m t)$$

or $$i = a + bV_c \cos \omega_c t + bV_m \cos \omega_m t + \frac{1}{2} cV_c^2 (1 + \cos^2 \omega_c t) + \frac{1}{2}$$

$$cV_m^2 (1 + \cos^2 \omega_m t) + cV_cV_m [\cos (\omega_c + \omega_m) t + \cos (\omega_c - \omega_m)t]$$

or $$i = a + bV_c \cos \omega_c t + bV_m \cos \omega_m t + \frac{1}{2} cV_c^2 + \frac{1}{2} cV_c^2 \cos^2 \omega_c t +$$

$$\frac{1}{2} cV_m^2 + \frac{1}{2} cV_m^2 \cos 2\,\omega_m t + cV_cV_m \cos (\omega_c+\omega_m)t + cV_cV_m \cos (\omega_c - \omega_m)t$$

or $$i = \underbrace{\left(a+\frac{1}{2}cV_c^2+\frac{1}{2}cV_m^2\right)}_{(1)} + \underbrace{bV_c \cos \omega_c t}_{(2)} + \underbrace{bV_m \cos \omega_m t}_{(3)} + \underbrace{\left(\frac{1}{2}cV_c^2 \cos 2\omega_c t + \frac{1}{2}cV_m^2 \cos 2\omega_m t\right)}_{(4)}$$

$$+ \underbrace{cV_cV_m \cos (\omega_c+\omega_m)t}_{(5)} + \underbrace{cV_c V_m \cos (\omega_c - \omega_m)t}_{(6)} \qquad ...(3.60)$$

Equation (3.60), consists of six terms in all as follow:

term (1) is the d.c. term
term (2) is the carrier signal
term (3) is the modulating signal
term (4) consists of harmonics of carrier and modulating signals
term (5) represents the upper sideband
term (6) represents the lower sideband

As discussed earlier, in the diode modulation circuit, the load impedance is a tuned circuit which is tuned to the carrier frequency ω_c. Therefore, this tuned circuit responds to a narrowband of frequencies centred about the carrier frequency ω_c. Thus the frequency components which are actually developed in the output are terms of frequency ω_c, $(\omega_c + \omega_m)$ and $(\omega_c - \omega_m)$. The rest of the frequency components are rejected by the tuned circuit.

Therefore, the required expression of output current will be

$$i_0 = bV_c \cos \omega_c t + c\,V_c\,V_m \cos (\omega_c + \omega_m)t + cV_cV_m \cos (\omega_c - \omega_m)t$$

or $$i_0 = bV_c \cos \omega_c t + cV_c.\,V_m[\cos (\omega_c + \omega_m)t + \cos (\omega_c - \omega_m)t]$$

or $$i_0 = bV_c \cos \omega_c t + 2\,cV_c\,V_m \cos \omega_c t \cos \omega_m t$$

or $$i_0 = bV_c \left(1+\frac{2cV_m}{b} \cos \omega_m t\right) \cos \omega_c t$$

$$i_0 = bV_c\,(1 + m_a \cos \omega_m t) \cos \omega_c t \qquad ...(3.61)$$

where $m_a = \dfrac{2cV_m}{b}$ is the modulation index. Equation (3.61) is the required expression for AM current.

3.12. Collector Modulation Method

Collector modulation method is a very popular method for AM generation. Figure 3.10 illustrates the circuit diagram of a collector modulation method. Here, the transistor T_1 makes a radio frequency (*RF*) class-*C* amplifier. At the base of T_1, the carrier signal is applied. V_{cc} makes the collector supply used for biasing purpose. Also, the transistor T_2 makes a class-*B* amplifier which is used to amplify the audio or modulating signal. The baseband or modulating signal appears across the modulation transformer after amplification. This amplified baseband or modulating signal appears in series with the collector supply V_{cc}. The function of the capacitor *C* is to offer low impedance path for the high-frequency carrier signal and hence the carrier signal is prevented from flowing through the modulation transformer.

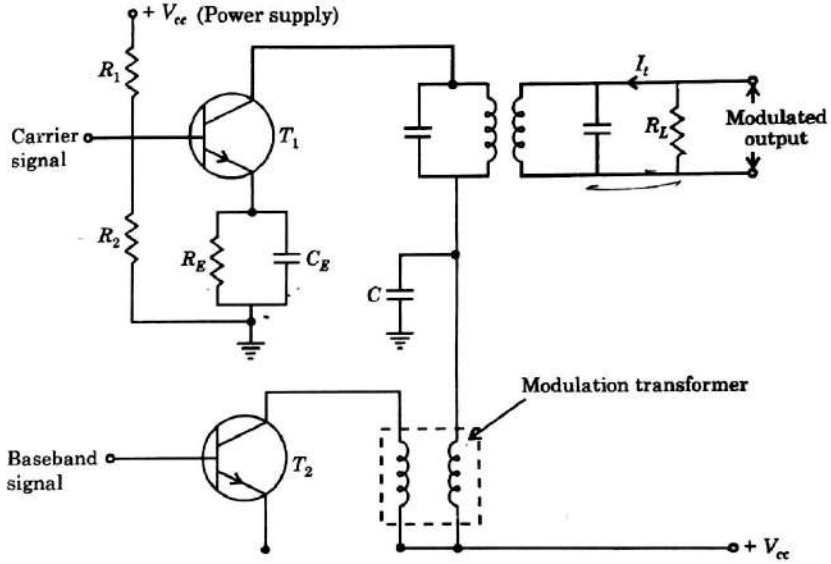


Fig. 3.10. *Illustration of collector modulation method.*

Operating Principle

It is a known fact that in a class-*C* amplifier, the magnitude of the output voltage is a definite fraction of or at the most equal to the supply voltage V_{cc}. In addition to this, a linear relationship exists between the output tank current I_t and the variable supply Voltage V_c (*i.e.*, here we have assumed that the supply voltage V_{cc} is a varying quantity and its varying value is denoted by V_c). This means that in class-C amplifier, the output voltage will be an exact replica of the input voltage waveform and the magnitude of the output voltage will be approximately equal to the carrier supply voltage V_{cc}.

Now if *R* is the resistance of the output tank circuit at resonance, then the magnitude of the output voltage is given as

$$RI_t \cong V_{cc}$$

Therefore, the unmodulated carrier is amplified by class-*C* modulated amplifier using transistor T_1 and its magnitude will remain constant at V_{cc} since there appears no voltage across the modulating transformer in the absence of baseband or modulating voltage.

But now if a baseband or modulating voltage $v_m = V_m \cos \omega_m t$ appears across the modulating transformer, this signal will be added to the carrier supply voltage V_{cc}. This results in a quite slow

variation in carrier supply voltage V_{cc}. This type of slow variation in carrier supply voltage changes the magnitude of the carrier signal voltage at the output of the modulated class-C amplifier as shown in figure 3.11.

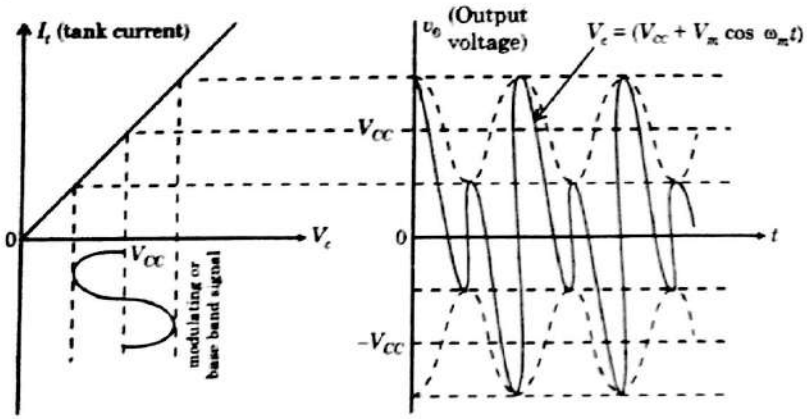


Fig. 3.11.

It may be observed that the envelope of the output voltage is identical with the baseband or modulating voltage and hence an AM signal is generated.

Mathematical Analysis

The slowly changing carrier supply voltage V_c may be expressed as

$$V_c = V_{cc} + v_m$$

or

$$V_c = V_{cc} + V_m \cos \omega_m t \quad ...(3.62)$$

But we know that for AM, the modulation index m_a is given by

$$m_a = \frac{\text{Maximum modulating voltage}}{\text{Maximum carrier voltage}}$$

In this case,

$$m_a = \frac{V_m}{V_{cc}} \quad ...(3.63)$$

Therefore,

$$V_c = V_{cc} + m_a . V_{cc} . \cos \omega_m t = V_{cc} (1 + m_a \cos \omega_m t) \quad ...(3.64)$$

Let the carrier voltage be

$$v_c = V_{cc} \cos \omega_c t$$

Then the modulated output voltage will be

$$v_0 = V_c \cos \omega_c t \quad ...(3.65)$$

Putting the value of V_c in above equation from equation (3.64), we get

$$v_0 = V_{cc} (1 + m_a \cos \omega_m t) \cos \omega_c t$$

which is the required expression for AM wave.

3.13. Demodulation of AM Waves

The process of extracting a modulating or baseband signal from the modulated signal is called **demodulation or detection**. In other words, demodulation or detection is the process by which the message is recovered from the modulated signal at receiver. The devices used for demodulation or detection are called demodulators or detectors. For amplitude modulation, detectors or demodulators are categorized as:

(*i*) Square-Law detectors

(*ii*) Envelope detectors

AM signal with large carrier are detected by using the envelope detector. The envelope detector uses the circuit which extracts the envelope of the AM wave. Infact, the envelope of the AM wave is the baseband or modulating signal. But a low-level amplitude modulated signal can only be detected by using square-law detectors in which a device operating in the non-linear region is used to detect the modulating signal.

3.13.1. Square-Law Detector

(U.P. Tech. Sem., Exam., 2003-04) (05 marks)

The square law detector circuit is used for detecting modulated signal of small magnitude (*i.e.* below 1 volt) so that the operating region may be restricted to the non-linear portion of the V-I characteristics of the device. Figure 3.12 shows the circuit of a square-law detector.* It may be observed that the circuit is very similar to the square-law modulator. The only difference lies in the filter circuit. In a square law modulator, the filter used is a bandpass filter whereas in a square law detector, a low-pass filter is used.

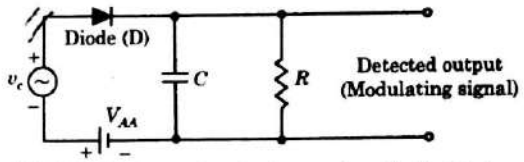


Fig. 3.12. *Basic circuit of square law diode detector.*

In the circuit, the dc supply voltage V_{AA} is used to get the fixed operating point in the non-linear portion of the diode V-I characteristic. Since, the operation is limited to the non-linear region of the diode characteristics, the lower half-portion of the modulated waveform is compressed. This produces envelope applied distortion. Due to this, the average value of the diode-current is no longer constant, rather it varies with time as shown in figure 3.13.

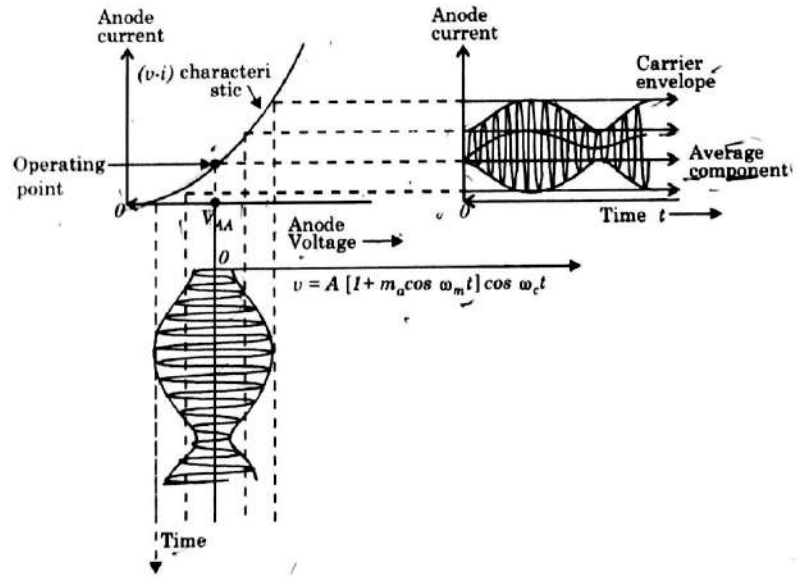


Fig. 3.13.

This distorted output diode current is expressed by the non-linear v-i relationship (*i.e.* square law) as

* Explain the working of square law demodulator for detection of AM wave.
(U.P. Tech, Sem. Exam., 2003-04) (05 marks)

$$i = av + bv^2 \qquad ...(3.66)$$

Here, v is the input modulated voltage.

We know that AM wave is expressed as

$$v = A(1 + m_a \cos \omega_m t) \cos \omega_c t \qquad ...(3.67)$$

Substituting, the value of v in equation (3.66), we get

$$i = a[A(1 + m_a \cos \omega_m t) \cos \omega_c t] + b[A(1 + m_a \cos \omega_m t) \cos \omega_c t]^2 \qquad ...(3.68)$$

Now, if above expression is expanded, then we may observe the presence of terms of frequencies like $2\omega_c$, $2(\omega_c \pm \omega_m)$, ω_m and $2\omega_m$ besides the input frequency terms.

Hence, this diode current i containing all these frequency terms is passed through a low-pass filter which allows to pass the frequencies below or upto modulating frequency ω_m and rejects the other higher frequency components. Therefore, the modulating or baseband signal with frequency ω_m is recovered form the input modulated signal.

3.13.2. Linear Diode or Envelope Detector

It is a known fact that a diode operating in a linear region of its *V-I* characteristics can extract the envelope of an AM wave. This type of detector is known as **envelope detector** or **a linear detector.** Envelope detector is most popular in commercial receiver circuits since it is very simple and is not expensive, also at the sametime, it gives satisfactory performance for the reception of broadcasting programmes. Figure 3.14 shows the circuit diagram of a linear diode detector or envelope detector. In the input portion of the circuit, the tuned transformer provides perfect tuning at the desired carrier frequency. *R-C* network is the time-constant network. If the magnitude of the modulated signal at the input of the detector is 1 volt or more, the operation takes place in the linear portion of the V-I characteristics of diode. Figure 3.15 shows the idealized linear characteristics of the diode along with the input voltage and output current waveforms.

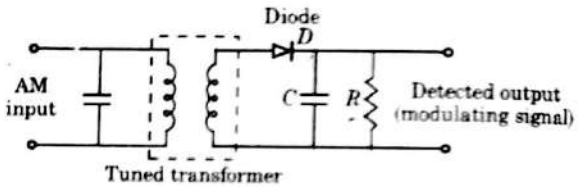


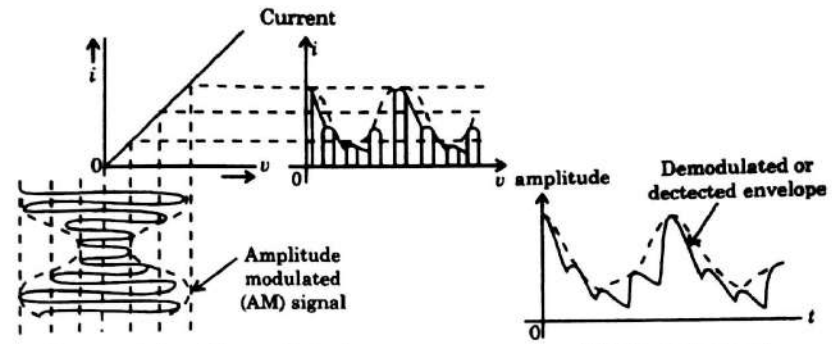
Fig. 3.14. *A linear diode detector.*

Operating Principle

First of all, let us assume that the capacitor is absent in the circuit. In this case, the detector circuit will work as a half-wave rectifier. Therefore, the output waveform would be a half-rectified modulated signal as shown in figure 3.15. Now let us consider that the capacitor is introduced in the circuit. For the positive half-cycle, the diode conducts and the capacitor is charged to the peak value of the carrier voltage. However, for a negative half-cycle, the diode is reverse-biased and does not conduct. This means that the input carrier voltage is disconnected from the *R-C* circuit. Therefore, the capacitor starts discharging through the resistance R with a time-constant $\tau = RC$. If the time-constant $\tau = RC$ is suitably chosen, the voltage across the capacitor C will not fall appreciably during the small period of negative half-cycle, and by that time the next positive cycle appears. This positive cycle again charges the capacitor C to the peak value of the carrier voltage and thus this process repeats again and again.

Hence, the output voltage across the capacitor C is a spiky modulating or baseband signal. This means that the voltage across the capacitor C is same as envelope of the modulated carrier signal, however, spikes are introduced because of charging and discharging of the capacitor C.

Figure 3.15(*b*) illustrates the resulting detected modulating or baseband signal. We can reduce discharges negligibly amount by keeping the time constant RC large so that the capacitor C discharges negligible small. However, the large value of $\tau = RC$ produces another problem known as diagonal clipping. Thus, we cannot increase the time-constant beyond a certain limit. Therefore, the time-constant is an important consideration for envelope detector.



(a) Characteristics of linear diode detector *(b) Detected output*

Fig. 3.15.

3.14. Double Sideband Suppressed Carrier (DSB-SC) System

The equation of AM wave in its simplest form *i.e.*, single-tone modulation, is expressed as

$$s(t) = A \cos w_c t + A.\frac{m_a}{2} \cos (w_c + w_m)t + A.\frac{m_a}{2} \cos (w_c - w_m)t \qquad ...(3.68)$$

From this equation, it is obvious that the carrier component in AM wave remains constant in amplitude and frequency. This means that the carrier of amplitude modulated wave does not convey any information. In power calculation of AM signal, it has been observed that for single-tone sinsusoidal modulation, the ratio of the total power to the carrier power is $\left(1 + \frac{m_a^2}{2}\right)$, m_a being the modulation index. Thus for 100% modulation about 67% of the total power is required for transmitting the carrier which does not contain any information. Hence, if the carrier is suppressed, only the sidebands remain and in this way a saving of two-third power may be achieved at 100% modulation, This type of suppression of carrier does not affect the baseband signal in any way. The resulting signal obtained by suppressing the carrier from the modulated wave is called **Double sideband suppressed carrier (DSB-SC) system.**

Thus, in a double-sideband suppressed carrier modulation there is no carrier signal only sidebands are present.

We know that the frequency-shifting property of Fourier transform is given as

If $\qquad x(t) \longleftrightarrow X(w)$

then $\qquad e^{jw_c t}\, x(t) \longleftrightarrow X(w - w_c) \qquad ...(3.69)$

This property states the if a signal $x(t)$ is multiplied by $e^{jw_c t}$ in time-domain then its spectrum $X(w)$ in frequency-domain is shifted by an amount w_c.

Similarly,

$$e^{-jw_c t}\, x(t) \longleftrightarrow X(w + w_c) \qquad ...(3.70)$$

But, since $e^{jw_c t}$ is not a real function and cannot be generated practically, therefore, frequency shifting in practice is achieved by multiplying $x(t)$ by a sinusoid such as $\cos w_c t$.

Hence, $\quad x(t) \cos w_c t = x(t) \cdot \frac{1}{2}(e^{jw_c t} + e^{-jw_c t})$

$$x(t)\cos w_c t = \frac{1}{2}x(t)\,e^{jw_c t} + \frac{1}{2}x(t)e^{-jw_c t} \qquad ...(3.71)$$

Using frequency-shifting property in equation (3.71), we get

$$x(t)\cos w_c t \longleftrightarrow \frac{1}{2}\left[X(w - w_c) + X(w + w_c)\right] \qquad ...(3.72)$$

This means that the multiplication of a signal $x(t)$ by a sinusoid of frequency w_c shifts the spectrum $X(w)$ by $\pm w_c$.

Now if $x(t)$ is taken as modulating or baseband signal and $\cos w_c t$ is taken as carrier signal, then $x(t)\cos w_c t$ represents the modulated signal. Further, the Fourier transform of this modulated signal is given by the equation (3.72). This equation shows that the spectrum of modulated signal contains only shifted spectrum of signal and there is no carrier components. However, we know that the modulated signal, which contains no carrier but two sidebands is called **Double-Sideband Suppressed Carrier (DSB-SC) modulation.**

This means that the term $x(t)\cos w_c t$ represents a DSB-SC signal.

Therefore, a DSB-SC signal is obtained by simply multiplying modulating signal $x(t)$ with carrier signal $\cos w_c t$. This is achieved by a product modulator. The block diagram of a product modulator is shown in figure 3.16.

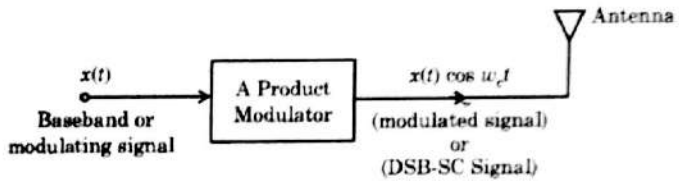


Fig. 3.16.

Figure 3.17 shows the modulating signal and its Fourier transform, carrier signal and its Fourier transform and DSB-SC signal and its frequency spectrum.

Few Points

(*i*) It is obvious from the figure 3.17, that the DSB-SC signal exhibits phase-reversal at zero crossings, *i.e.*, whenever the baseband signal $x(t)$ crosses zero. Because of this, the envelope of a DSB-SC modulated signal is different from the message signal. This is unlike the case of an AM wave.

(*ii*) From figure 3.17, it is also clear that the impulses at $\pm w_c$ are missing which means that the carrier term is suppressed in the spectrum and only two sideband terms, USB and LSB are left. Therefore, it is called double sideband suppressed carrier (DSB-SC) system.

(*iii*) In figure 3.17, considering only positive side the upperside band frequency is $w_c + w_m$ whereas the lowerside band frequency is $w_c - w_m$. The difference of these two is equal to the **transmission bandwidth of a DSB-SC signal**, *i.e.*,

Bandwidth $B = (w_c + w_m) - (w_c - w_m) = 2\,w_m$

It is obvious that the bandwidth of DSB-SC modulation is same as that of general AM wave.

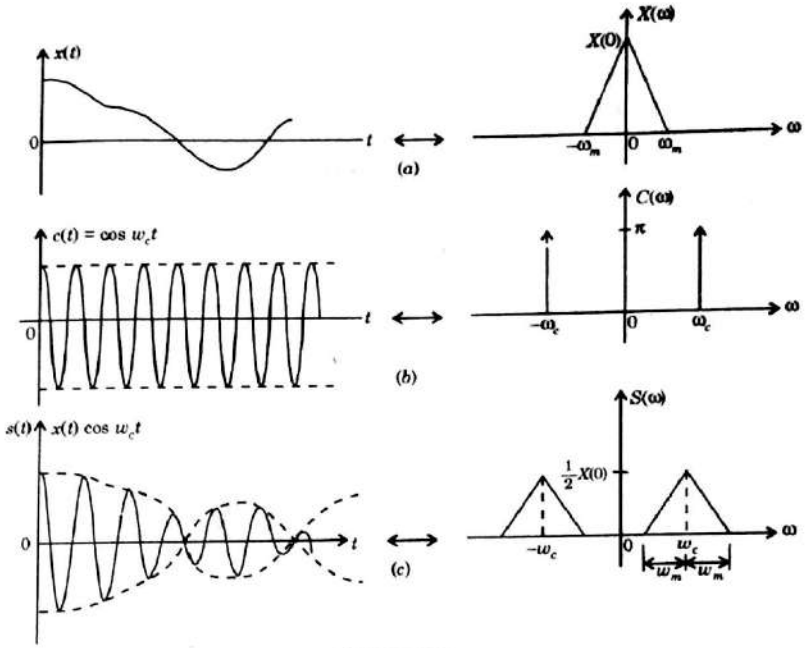


Fig. 3.17.

3.15. Generation of DSB-SC Signal

The expression for DSB-SC signal is given as

$$s(t) = x(t)\cos w_c t$$

where $x(t)$ = Baseband signal

$\cos w_c t$ = Carrier signal

From this expression, we see that a DSB-SC signal is basically the product of the modulating or baseband signal and the carrier signal. Unfortunately, a single electronic device cannot generate a DSB-SC signal. As discussed earlier, a circuit to achieve the generation of a DSB-SC signal is called a product modulator. In this section, we shall discuss two types of product modulator, namely the **Balanced Modulator** and the **Ring Modulator.**

3.15.1. The Balanced Modulator

We know that a non-linear resistance or a non-linear device may be used to produce Amplitude Modulation, *i.e.*, one carrier and two sidebands. However, a DSB-SC signal contains only two sidebands. Thus, if two non-linear devices such as diodes, transistors etc. are connected in a balanced mode so as to suppress the carriers of each other, then only sidebands are left, *i.e.*, a DSB-SC signal is generated.

Therefore, a balanced modulator may be defined as a circuit in which two non-linear devices are connected in a balanced mode to produce a DSB-SC signal.

In the following section, we shall discuss a balanced modulator circuit using diodes.

Figure 3.18 shows a balanced modulator circuit using two diodes. A modulating signal $x(t)$ is applied to the two diodes through a centre-tapped transformer with the carrier signal $\cos w_c t$.

A non-linear v-i relationship is given as

$$i = av + bv^2 \quad ...(3.73)$$

Here, we have neglected the higher power terms.

In above expression, v is the input voltage applied across a non-linear device and i is the current through the non-linear device.

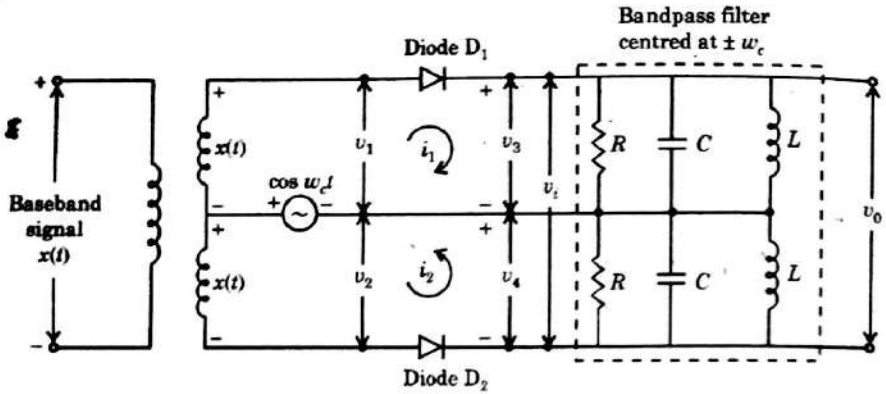


Fig. 3.18. *Balanced Modulator using Diodes.*

From figure 3.18, we write the two input voltages v_1 and v_2 across the two diodes as

$$v_1 = \cos w_c t + x(t) \quad ...(3.74)$$

$$v_2 = \cos w_c t - x(t) \quad ...(3.75)$$

For diode D_1, the nonlinear v-i relationship becomes

$$i_1 = av_1 + bv_1^2 \quad ...(3.76)$$

Similarly, for diode D_2, the nonlinear v-i relationship becomes

$$i_2 = av_2 + bv_2^2 \quad ...(3.77)$$

In the expression of current i_1, substituting the value of v_1, we get

$$i_1 = a\,[\cos w_c t + x(t)] + b\,[\cos w_c t + x(t)]^2$$

or $$i_1 = a \cos w_c t + ax(t) + b\,[\cos^2 w_c t + x^2(t) + 2x(t) \cos w_c t]$$

or $$i_1 = a \cos w_c t + ax(t) + b \cos^2 w_c t + bx^2(t) + 2bx(t) \cos w_c t \quad ...(3.78)$$

Similarly, in the expression of current i_2, substituting the value of v_2, we get

$$i_2 = a\,[\cos w_c t - x(t)] + b[\cos w_c t - x(t)]^2$$

or $$i_2 = a \cos w_c t - ax(t) + b\cos^2 w_c t + bx^2(t) - 2bx(t) \cos w_c t \quad ...(3.79)$$

Due to currents i_1 and i_2, the net voltage v_i at the input of bandpass filter is expressed as

$$v_i = v_3 - v_4 \quad ...(3.80)$$

But from figure, we have

$$v_3 = i_1 R$$

and $$v_4 = i_2 R$$

Therefore, $$v_i = i_1 R - i_2 R$$

or $$v_i = R(i_1 - i_2)$$

In above equations, substituting the values of i_1 and i_2 from equations (3.78) and (3.79), we get

$$v_i = R[2ax(t) + 4bx(t) \cos w_c t]$$

or $$v_i = 2R[ax(t) + 2bx(t) \cos w_c t]$$

This voltage v_i is the input to the bandpass filter centered around $\pm w_c$.

A bandpass filter is that type of filter which allows to pass a band of frequencies. Here, since the bandpass filter is centred around $\pm w_c$, it will pass a narrowband of frequencies centred at $\pm w_c$ with a small bandwidth of $2\,w_m$ to preserve the sidebands.

Therefore, the output of BPF centred around $\pm w_c$ is given by

$$v_0 = 4a\,Rx(t) \cos w_c t = K\,x(t) \cos w_c t,$$

which is the expression for a DSB-SC signal

3.15.2. Ring Modulator

Ring modulator is another product modulator, which is used to generate DSB-SC signal.* Figure 3.19 shows the circuit diagram of a ring modulator. In a ring modulator circuit, four diodes are connected in the form of a ring in which all four diodes point in the same manner. All the four diodes in ring are controlled by a square wave carrier signal $c(t)$ of frequency f_c applied through a centre-tapped transformer.

In case, when diodes are ideal and transformer are perfectly balanced, the two outer diodes are switched on if the carrier signal is positive whereas the two inner diodes are switched off and thus presenting very high impedance as shown in figure 3.20(a). Under this condition, the modulator multiplies the modulating signal $x(t)$ by +1.

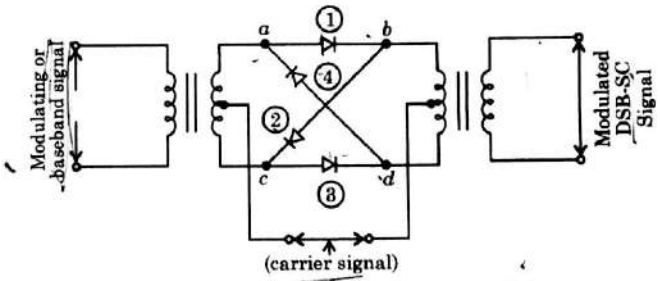


Fig. 3.19. *Circuit diagram of a ring modulator.*

Now, in case when carrier signal is negative, the situation becomes reversed as shown in figure 3.20(b). In this case, the modulator multiplies the modulating signal by −1.

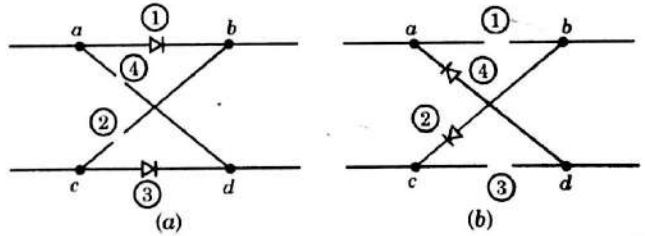


Fig. 3.20. *(a) Illustration of condition when the diodes (1) and (3) are switched ON and diodes (2) and (4) are switched OFF.*
(b) Illustration of condition when the diodes (2) and (4) are switched ON and diodes (1) and (3) switched. OFF.

Hence, the ring modulator is a product modulator for a square wave carrier and modulating signal.

Figure 3.21 illustrates the modulating signal (assuming a sinusoidal signal), square wave carrier signal and the modulated (DSB-SC) signal.

* What is a DSB-SC modulator? Explain the working principle of a ring modulator.
(U.P. Tech Sem. Exam., 2003-04) (05 marks)

The square wave carrier may be represented in Fourier series as under:

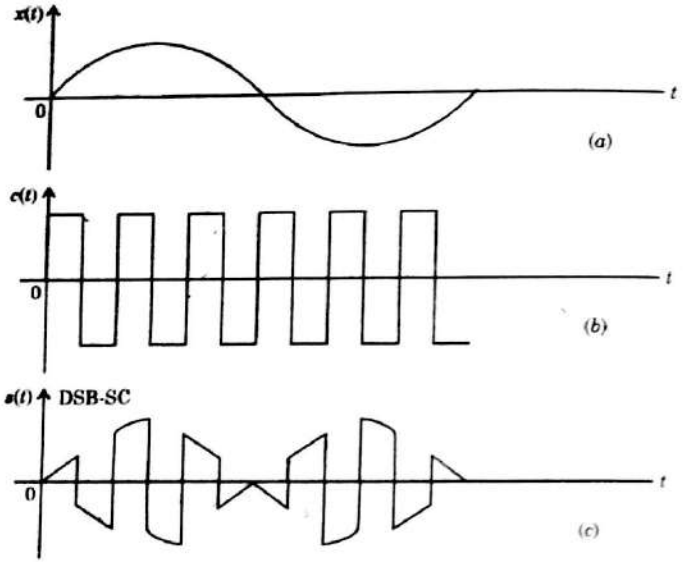


Fig. 3.21. *(a) Sinusoidal modulating wave (b) Square wave carrier (c) Modulated wave (DSB-SC) output of the Ring Modulator.*

$$c(t) = \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{(2n-1)}\left\{\cos\left[2\pi f_c t\,(2n-1)\right]\right\}$$

We have $\quad s(t) = x(t)\,c(t)$

Substituting the value of $c(t)$, we have

or $$s(t) = x(t)\,\frac{4}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{(2n-1)}\left\{\cos\left[2\pi f_c t\,(2n-1)\right]\right\}$$

or $$s(t) = \frac{4}{\pi}\sum_{n=1}^{\infty}\frac{(-1)^{n-1}}{(2n-1)}\left\{\cos\left[2\pi f_c t(2n-1)\right]\right\}x(t)$$

Thus, with the help of above mathematical analysis, it may be verified that the output from ring modulator does not have any component at carrier frequency. Hence, the modulated output does not have any component at carrier frequency. Hence, the modulated output contains only product terms.

A ring modulator is also known as a **double-balanced modulator** since it is balanced with respect to the baseband signal as well as the square wave carrier signal.

The frequency spectrum of the ring modulator output contains sidebands around each of the odd harmonics of the square wave carrier signal as illustrated in figure 3.22. We have assumed that the modulating or baseband signal is bandlimited to $-f_m \le f \le f_m$. The desired sideband around the carrier frequency f_c may be selected by using a bandpass filter (BPF) having center frequency w_c and bandwidth $2f_m$.

From figure 3.22, it may be observed that to avoid overlapping of sidebands we must have $f_c > f_m$.

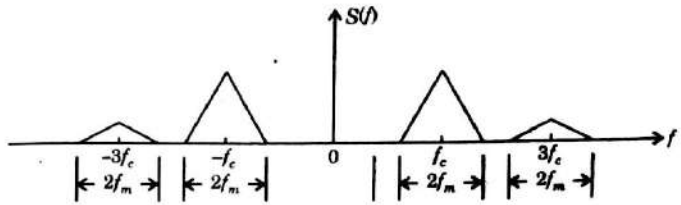


Fig. 3.22. *Spectrum of the DSB-SC output of the ring modulator.*

3.16. Demodulation of DSB-SC Signals

The DSB-SC signal may be demodulated by following two methods:

(*i*) Synchronous detection method

(*ii*) Using envelope detector after carrier reinsertion.

3.16.1. Synchronous Detection Method

We know that the DSB-SC system is used at the transmitter end to shift the modulating signal (having maximum frequency w_m) to a higher carrier frequency $\pm w_c$. Now, this modulated (DSB-SC) signal is transmitted from the transmitter and it reaches the receiver through a transmission medium. At the receiver end, the original modulating signal $x(t)$ is recovered from the modulated (DSB-SC) signal. This can be achieved by simply retranslating the baseband or modulating signal from a higher spectrum, centered at $\pm w_c$, to the original spectrum. This process of retranslation is called **demodulation** or **detection**. Hence, the original or baseband signal is recovered from the modulated signal by the detection process.

A method of DSB-SC detection is known as **synchronous detection.** Figure 3.23 shows the block diagram of synchronous detection method.

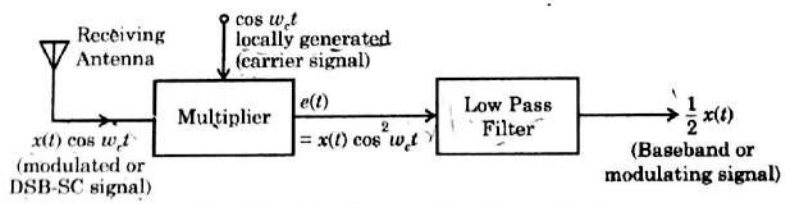


Fig. 3.23. *Synchronous detection method.*

Working Principle

In synchronous detection method, the received modulated or DSB-SC signal is first multiplied with a locally generated carrier signal $\cos \omega_c t$ and then passed through a low-pass filer (LPF). At the output of a low-pass filter (LPF), the original modulating signal is recovered.

Mathematically,

$$e(t) = \underbrace{x(t)\cos w_c t}_{\text{DSB-SC signal}} \cdot \underbrace{\cos w_c t}_{\text{Locally generated carrier signal}} \qquad ...(3.81)$$

or $$e(t) = x(t)\cos^2 w_c t = \frac{1}{2}x(t)\left[2\cos^2 w_c t\right]$$

or $$e(t) = \frac{1}{2}x(t)\left[1+\cos 2w_c t\right] = \frac{1}{2}x(t) + \frac{1}{2}x(t)\cos 2w_c t \qquad ...(3.82)$$

Now, it may be observed that when multiplied signal $e(t)$ is passed through a low-pass filter (LPF), then the term $\frac{1}{2}x(t)\cos 2w_c t$, centred at $\pm 2w_c$ is suppressed by low-pass filter and thus at the output of low-pass filter, the original modulating signal $\frac{1}{2}x(t)$ is obtained.

We may get the frequency-spectrum of multiplied signal $e(t)$ using Fourier Transform as

$$x(t)\cos^2 w_c t \longleftrightarrow \frac{1}{2}X(w) + \frac{1}{4}\left[X(w+2w_c) + X(w-2w_c)\right] \quad ...(3.83)$$

Figure 3.24 shows this frequency spectrum

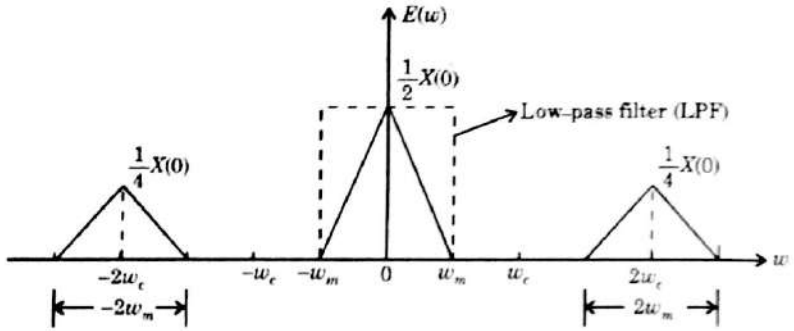


Fig. 3.24. *Frequency spectrum of $x(t)\cos^2 w_c t$.*

Note: *The spectrum reveals the fact that the original baseband or modulating signal (0 to w_m) is present alongwith another frequency spectrum centered around $\pm w_c$ when this signal i.e., $x(t)\cos 2w_c t$ is passed through a low pass filter (LPF) with cut-off frequency w_m, the original baseband signal or modulating signal appears at the output of low-pass filter. It may be noted $w_c >> w_m$ and $2w_c$ is still greater than w_m and thus is easily filtered out.*

Hence, the original modulating or message signal $x(t)$ is recovered from the DSB-SC signal.

Synchronous or Coherent Detection

We have observed in last article that the detection process for DSB-SC requires a local oscillator at the receiver end. The frequency and phase of the locally generated carrier signal and the carrier signal at the transmitter carrier must be identical. This means that the local oscillator signal must be exactly coherent or synchronized with the carrier signal at the transmitter, both in frequency and phase, otherwise the detected signal would get distorted. Therefore, this method of recovery is called **synchronous detection** or **coherent detection**. Thus, the demerits of the synchronous detection is that if requires an additional system at the receiver to ensure that the locally generated carrier is synchronized with the transmitter carrier making the receiver complex and costly.

3.16.2. Envelope Detection after Suitable Carrier Re-insertion

The other possible method of demodulating DSB-SC signal is by inserting a carrier generated at the receiver end with the help of a local oscillator. However, the phase and the frequency of the re-inserted carrier must be properly synchronized with those at the transmitter end in order to avoid distortion. We know that if we insert a sufficient carrier of same frequency and phase to DSB-SC signal, it converts DSB-SC signal into a conventional AM wave. Now, this AM wave is demodulated by an envelope detector. However, phase and frequency errors will result in similar type of distortion as obtained in coherent detection.

Let us consider that the received DSB-SC signal is expressed by

$$s(t) = c(t)\cdot x(t)$$

or $$s(t) = A\cos(2\pi f_c t)\,x(t) \quad ...(3.84)$$

Let us assume $A = 1$ in equation (3.84). Then equation (3.84) can be written as

$$s(t) = \cos(2\pi f_c t)\,x(t) \quad ...(3.85)$$

The inserted carrier at the receiver will be

$$c'(t) = A\cos(2\pi f_c t + \phi)$$

where ϕ = amount of phase discrepancy.

Then the resulting signal will be

$$r(t) = s(t) + \text{re-inserted carrier at the receiver}$$

or $$r(t) = s(t) + c'(t) \quad ...(3.86)$$

Substituting equations (3.84) and (3.85) in equation (3.86), we get

$$r(t) = s(t) + c'(t) = \cos(2\pi f_c t)x(t) + A\cos(2\pi f_c t + \phi)$$

or $$r(t) = x(t)\cos(2\pi f_c t) + A\cos(2\pi f_c t)\cos\phi - A\sin(2\pi f_c t)\sin\phi$$

or $$r(t) = [x(t) + A\cos\phi]\cos(2\pi f_c t) - (A\sin\phi)\sin(2\pi f_c t)$$

or $$r(t) = e(t)\cos[(2\pi f_c t) + \theta(t)] \quad ...(3.87)$$

where $$e(t) = \sqrt{[A + x(t)]^2 - 2A\,x(t)[1-\cos\phi]}$$

and $$\theta(t) = \tan^{-1}\left[\frac{A\sin\phi}{x(t) + A\cos\phi}\right]$$

Now, from the expression

$$r(t) = e(t)\{\cos[(2\pi f_c t) + \theta(t)]\}$$

it may be observed that $e(t)$ is the envelope of the resulting signal $r(t)$.

Also, if we take $\phi = 0$, then envelope will be given by

$$e(t) = A + x(t)$$

Hence, modulating signal $x(t)$ can be recovered from $r(t)$ using an envelope detector since the $r(t)$ is basically a conventional AM wave given by

$$r(t) = [A + x(t)]\cos 2\pi f_c t \quad ...(3.88)$$

This is however possible only when $[A + x(t)] > 0$ for all values of t.

It is possible only when the modulation index m is less than unity.

If $\phi \neq 0$, then the phase error exists between the two carriers. It is given as

$$e(t) = A\left[1 + \frac{2x(t)}{A}\cos\phi + \left\{\frac{x(t)}{A}\right\}^2\right]^{\frac{1}{2}} \quad ...(3.89)$$

If $A >> |x(t)|$, then, we have

$$e(t) \cong A + x(t)\cos\phi \quad ...(3.90)$$

The desired signal output will thus be $x(t)\cos\phi$. If $\phi = 0$ and there is a difference in frequency Δf between the two oscillators, then the envelope of the resulting signal $r(t)$ will be given by

$$e(t) = A + x(t)\cos[2\pi\Delta f t] \quad \text{for } A >> |x(t)| \quad ...(3.91)$$

The envelope can be detected by an envelope detector but here distortion will be identical to coherent detection.

3.17. Quadrature-Amplitude Modulation (QAM)

Definition

This modulation scheme is also called Quadrature Carrier Multiplexing. In fact, this modulation scheme enables two DSB-SC modulated signals to occupy the same transmission bandwidth and therefore it allows for the separation of the two message signals at the receiver output. It is therefore known as a **bandwidth-conservation scheme.**

Figure 3.25 shows a block diagram of the Quadrature Amplitude Modulation (QAM) system. The QAM transmitter has been shown in figure 3.26(a) and QAM receiver has been shown in figure 3.25(b). The QAM transmitter consists of two separate balanced modulators (BM) which are supplied with two carrier waves of the same frequency but differing in phase by 90°. The output of the two balanced modulators are added in the adder and transmitted. The transmitted signal is thus given by

$$s(t) = x_1(t)\, A \cos(2\pi f_c t) + x_2(t)\, A \sin(2\pi f_c t)$$

or

$$s(t) = A\, x_1(t) \cos(2\pi f_c t) + A\, x_2(t) \sin(2\pi f_c t) \qquad \text{...(3.92)}$$

where, $x_1(t)$ and $x_2(t)$ are the two different message signals applied to the product modulators. Both $x_1(t)$ and $x_2(t)$ are band-limited in the interval $-f_m \le f \le f_m$, then $s(t)$ will occupy a bandwidth of $2f_m$. This bandwidth $2f_m$ is centred at the carrier frequency f_c, where f_m is the bandwidth of message signal $x_1(t)$ or $x_2(t)$.

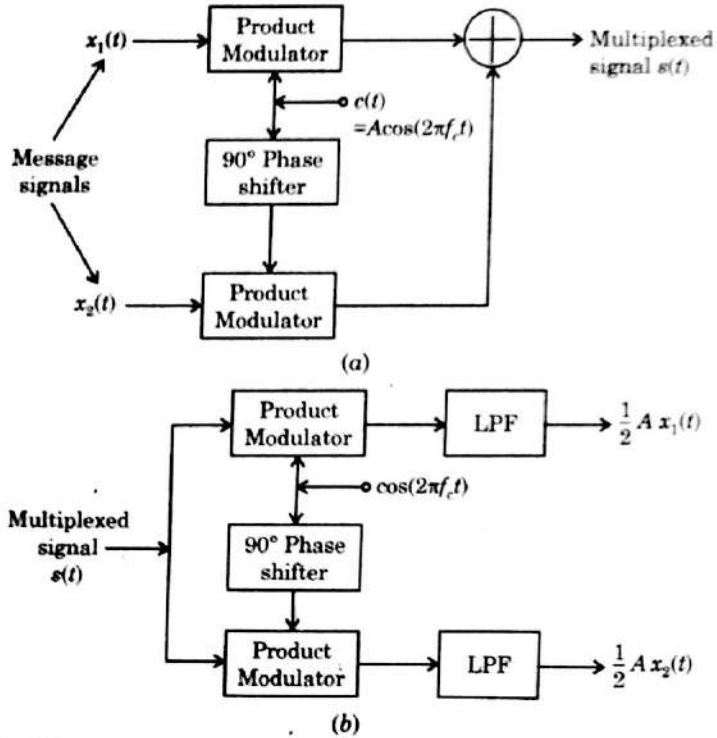


Fig. 3.25. *Illustration of Quadrature Amplitude Modulation (QAM) system (a) QAM transmitter (b) QAM receiver.*

Hence, the multiplexed signal consists of the in-phase component $A\, x_1(t)$ and the quadrature phase component $-A\, x_2(t)$.

The multiplexed signal $s(t)$ from QAM transmitter is applied simultaneously to two separate coherent detectors that are supplied with two local carriers of the same frequency, but differing in phase by 90°. The output of one detector is $1/2\, A\, x_1(t)$, whereas the output of the second detector is $1/2\, A\, x_2(t)$. For satisfactory operation of the coherent detector, it is essential to maintain coherent phase and frequency relationship between the oscillators used in the QAM transmitter and receiver parts of the system. The Quadrature Amplitude Modulation (QAM) finds application in colour television (CTV).

3.18. Effect of Phase and Frequency Errors in Synchronous Detection

As discussed earlier, the frequency and phase of the local oscillator signal in coherent detection method at the receiver end, must be identical to the transmitted carrier signal. Any kind of discrepancy in frequency or phase produces a distortion in the detected output at the receiver end.

In this subsection, let us examine the nature of distortion produced by a frequency or phase discrepancy.

For this, let us assume that the modulated signal reaching the receiver is denoted as $x(t) \cos \omega_c t$. Considering, a locally generated signal with frequency and phase error equal to $\Delta\omega$ and ϕ respectively, the product of the two signals in the synchronous detector provides

$$e_d(t) = x(t) \cos \omega_c t.\, \cos[(\omega_c + \Delta\omega)\, t + \phi]$$

or

$$e_d(t) = \frac{1}{2} x(t) \{\cos[(\Delta\omega)\, t + \phi] + \cos[(2\omega_c t + \Delta\omega) t + \phi]\} \qquad \text{...(3.93)}$$

Now, when this signal is allowed to pass through a low-pass filter (LPF) having a cut-off frequency ω_m, the terms centered around $\pm 2\omega_c$ are filtered out and the filter output is given as

$$e_0(t) = \frac{1}{2} x(t) \cos[(\Delta\omega) t + \phi] \qquad \text{...(3.94)}$$

The baseband signal $x(t)$ is multiplied by a slow-time varying function $\cos[(\Delta\omega)\, t + \phi]$ which distorts the message signal $x(t)$.

Now, let us consider the following special cases:

(*i*) When the frequency error $\Delta\omega$ and phase error ϕ are both zero, then last equation (3.94) gives

$$e_0(t) = \frac{1}{2} x(t)$$

This means that there is no distortion in the detected output signal.

(*ii*) When there is only the phase error *i.e.*,

$$\Delta\omega = 0 \quad \text{but} \quad \phi \ne 0$$

In this case, equation (3.94) gives

$$e_0(t) = \frac{1}{2} x(t) \cos\phi$$

This shows that the output signal is multiplied by $\cos\phi$. When ϕ is time-independent, there is no distortion, rather, there is only attenuation. The output will be maximum when $\phi = 0$ and minimum when $\phi = 90°$. But, in general ϕ randomly varies with respect to the time due to random variation of propagation media (*i.e.*, iono sphere). This causes undesirable distortion in the detected output.

It may be noted that the detected output is zero when $\phi = 90°$. This is known as a **quadrature null effect** since the signal is zero when the local carrier is in-phase quadrature with the transmitted carrier signal.

(iii) When there is only the frequency error, *i.e.*,

$$\Delta\omega \neq 0 \quad \text{and} \quad \phi = 0$$

In this case, equation (3.94) gives

$$e_0(t) = \frac{1}{2} x(t) \cos(\Delta\omega) t$$

Here, the multiplying factor cos $(\Delta\omega)$ t is time-dependent and produces distortion in the detected output signal. The error $\Delta\omega$ is usually small and thus a message signal $x(t)$ is multiplied by a slow varying sinusoidal signal. This is a more serious distortion. Therefore, a frequency error must be avoided

(iv) When both errors are non-zero, *i.e.*,

$$\Delta\omega \neq 0 \quad \text{and} \quad \phi \neq 0$$

In this case equation (3.94) itself provides the detected output signal. Also, in this case, the constant phase error provides attenuation and the frequency error produces distortion in the detected output signal.

Hence, we get an attenuated and distorted output signal at the receiver end.

3.19. Carrier Acquisition in DSB-SC System or Synchronization Techniques in DSB-SC System

As discussed in the last subsection, the phase and the frequency of the locally generated carrier signal in synchronous detector is very critical. Precision phase and frequency control of the local carrier requires an expensive and a complex circuitry at the receiver end. Some important synchronization techniques are given as under:

(i) Pilot Carrier

A small amount of carrier signal known as pilot carrier is transmitted alongwith the modulated signal from the transmitter. This small amount of carrier signal is called pilot carrier. This pilot carrier, separated at the receiver by an appropriate filter, is amplified, and is used to phase lock the locally generated carrier signal at the receiver. The phase locking provides synchronization. This system, where a weak carrier is transmitted alongwith the DSB-SC signal is also referred to as partially suppressed carrier system as the carrier is not totally suppressed. The process in which a large carrier is transmitted alongwith DSB-SC signal is known as amplitude modulation. This has been already discussed. The large carrier simplifies the reception system. The DSB-SC with partially suppressed carrier is equivalent to an over modulated AM signal.

(ii) Costa's Receiver

This system used for synchronous detection of DSB-SC signal has been shown in figure 3.26.

This system has two synchronous detectors–one detector is fed with a locally generated carrier signal which is in phase with the transmitted carrier signal. This detector circuit is called inphase coherent detector or I-channel. The other synchronous detector employs a local carrier which is in phase quadrature with the transmitted carrier signal and is called Quadrature phase coherent detector or Q-channel. On combining, the two detectors constitute a negative feedback system which synchronizes the local carrier signal with the transmitted carrier signal.

Figure 3.26 shows a costa's receiver.

Operating Principle

To start with, let us assume that the local carrier signal is synchronized with the transmitted carrier signal and $\phi = 0$. As shown in figure 3.26, the output of the I-channel is the desired modulating signal (since cos $\phi = 1$), but the output of the Q-channel is zero (since sin $\phi = 0$) because of the quadrature null effect. Now, assuming that the local oscillator frequency drifts slightly *i.e.*, ϕ is a very small non-zero quantity, I-channel output will be almost unchanged, but Q-channel output

now is not a zero, rather some signal would appear at its output and is proportional to sin ϕ. Thus, the output of the Q-channel,

(i) is proportional to ϕ (since sin $\phi = \phi$ for small ϕ)

(ii) would have a polarity same as the I-channel for one direction of phase shift in local oscillator, whereas, the polarity would be opposite to I-channel for the other direction of phase shift.

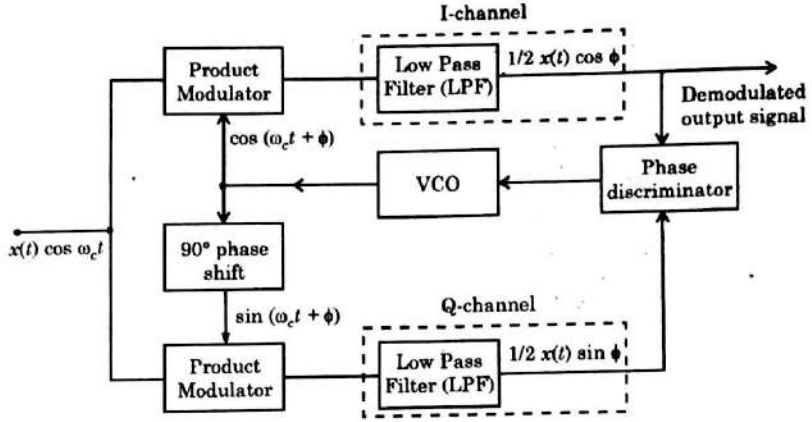


Fig. 3.26. *A Costa's receiver.*

The phase discriminator provides a d.c. control signal which may be used to correct local oscillator phase error. The local oscillator is a voltage controlled oscillator (VCO). Its frequency may be adjusted by an error control d.c. signal.

Limitation

The costa's receiver cases phase control when there is no modulation *i.e.*, $x(t) = 0$. The phase control reestablishes itself on the reappearance of modulation. However, the reestablishment is so fast that distortion is not perceptible in case of voice communication.

(iii) Squaring Loop

In this method, the received signal is squared by a squaring circuit as shown in figure 3.27. The output of the squarer will be given as

$$[A.\ x(t) \cos \omega_c t]^2 = A^2 x^2(t) \cos^2 \omega_c t$$

For simplicity let us assume that $x(t)$ is a single tone sinusoid denoted as cos $\omega_m t$ *i.e.*,

$$x(t) = \cos \omega_m t$$

then the output of the squarer becomes

$$[A \cos \omega_c t.\ \cos \omega_m t]^2 = A^2 \cos^2 \omega_m t \cos^2 \omega_c t$$

$$= \frac{A^2}{4} (1 + \cos 2\omega_m)(1 + \cos 2\omega_c t)$$

$$= \frac{A^2}{4} [1 + \cos 2\omega_m t + \cos 2\omega_c t + \cos 2\omega_c t \cos 2\omega_m t] \qquad \text{...(3.95)}$$

The term cos $2\omega_c t$ can be obtained by using a narrowband filter centred at $\pm 2\omega_c$. This frequency $\pm 2\omega_c$ is kept constant by tracking through a phase locked loop (PLL). The PLL uses negative feedback technique to provide a constant frequency signal, cos $2\omega_c t$.

Any drift in frequency is corrected by an error signals $e(t)$, generated at the output of the lowpass filter of PLL as depicted in figure 3.27. PLL will be discussed later on.

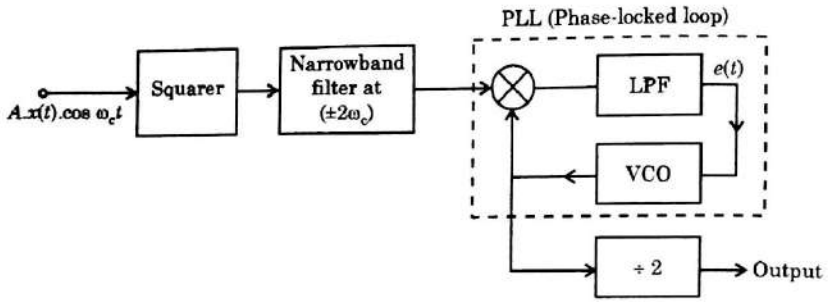


Fig. 3.27. *A squaring circuit for synchronization.*

The VCO output is frequency divided by 2, to yield a synchronized local carrier of frequency ω_c. This local carrier signal is used in synchronous detector. The frequency division can be accomplished by using a bistable multivibrator.

3.20. Single Sideband Suppressed-Carrier (SSB-SC) Modulation *(Important)*

As discussed earlier that amplitude modulation and double-sideband suppressed-carrier (DSB-SC) modulation are wasteful of bandwidth since they both need a transmission bandwidth equal to twice the message signal bandwidth. In either case one half of the transmission bandwidth is occupied by the upper sideband of the modulated signal whereas the other half is occupied by the lower sideband. However, the lower and upper sidebands are uniquely related to each other by virtue of their symmetry about the carrier frequency, *i.e.*, if amplitude and phase spectra of either sideband is given, we can uniquely determine the other. This means that as far as the transmission of information is concerned, only one sideband is necessary. Thus, if the carrier and one of the two sidebands are suppressed at the transmitter, no information is lost. Modulation of this type which provides a single sideband with suppressed carrier is known as single sideband suppressed carrier (SSB-SC) system. Thus, SSB-SC system reduces the transmission bandwidth by half. This means that in a given frequency band we can accommodate twice the number of channels by using a single sideband in place of both the sidebands.

Figure 3.28 further illustrated the concept of single sideband modulation with the help of different frequency spectrums.

Figure 3.28(*a*) shows the frequency spectrum of modulating or baseband signal. It contains ω_m as the maximum frequency component.

Figure 3.28(*b*) shows the frequency spectrum of a DSB-SC modulation which contains no carrier but two sidebands, *i.e.*, lower sideband and upper sideband.

Figure 3.28(*c*) shows the frequency spectrum of single sideband suppressed carrier modulation consisting of upper sideband only, *i.e.*, lower sideband and carrier signal are suppressed.

Figure 3.28(*d*) shows the frequency spectrum of single sideband suppressed carrier modulation consisting of lower sideband only, *i.e.*, upper sideband and carrier signal are suppressed.

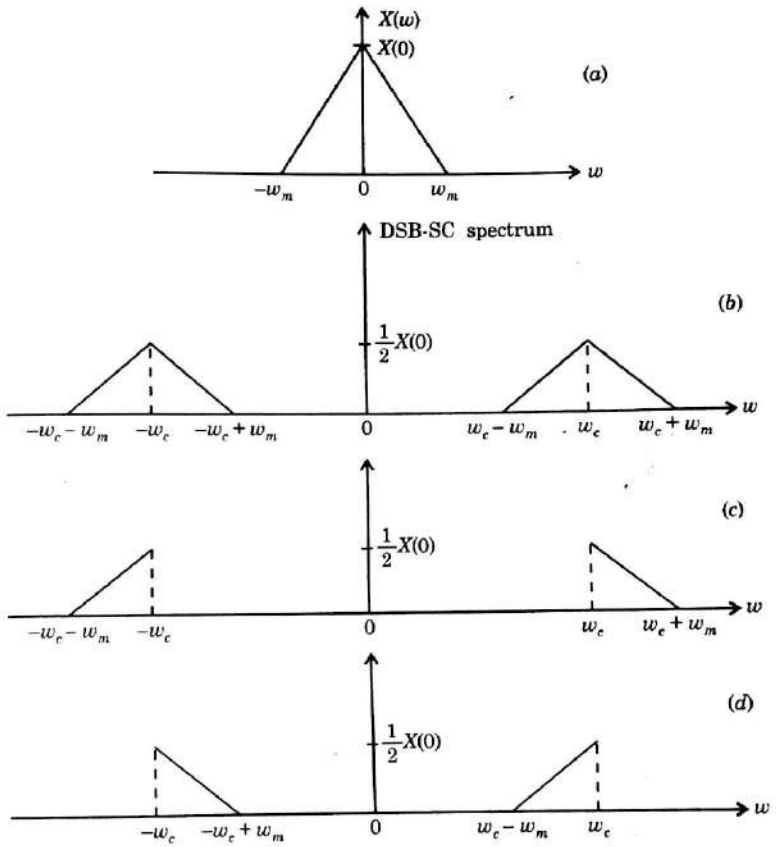


Fig. 3.28. *(a) Spectrum of baseband signal*
(b) Spectrum of DSB-SC wave
(c) Spectrum of SSB-SC wave with the upper sideband transmitted
(d) Spectrum of SSB-SC wave with the lower sideband transmitted.

3.20.1. Time-Domain Description of the SSB-SC Wave

We can derive an expression to represent the SSB wave in time-domain considering a special case of single-tone modulating signal. We can derive a general expression for SSB wave using the concept of a pre-envelope

SSB-SC wave with single-tone modulating signal

Let us consider a single-tone modulating signal as

$$x(t) = \cos w_m t \qquad ...(3.96)$$

The frequency spectrum of this modulating signal consist of two impuises located at $w = \pm w_c$ as shown in figure 3.29(*a*). If this modulating signal, modulates a carrier signal $\cos w_c t$, then the resulting spectrum of DSB-SC signal will be as shown in figure 3.29(*b*).

To get the SSB-SC waveform, we will have to eliminate one of the two sidebands.

Figure 3.29(*c*) shows the spectrum of SSB-SC wave with lower sidebands. From this figure, it is clear that this spectrum corresponds to a time-domain signal $\cos (w_c - w_m)t$, because the frequency

spectrum of cosine function contains two impulses in its frequency domain. This means that a SSB-SC wave with lower sideband may be expressed as

$$\cos(w_c - w_m)t = \cos w_m t \cos w_c t + \sin w_m t \sin w_c t \quad ...(3.97)$$

In the same manner, the exprexssion for the single-tone SSB-SC wave with upper sideband may be expressed as

$$\cos(w_c + w_m)t = \cos w_m t \cos w_c t - \sin w_m t \sin w_c t \quad ...(3.98)$$

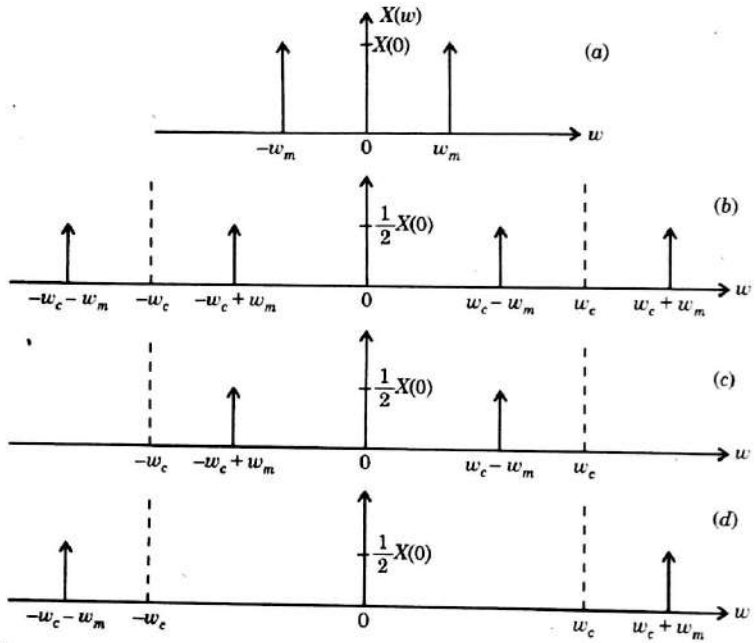


Fig. 3.29. *(a) Frequency spectrum of modulating signal (b) Frequency spectrum of DSB-SC wave (c) Spectrum of SSB-SC with lower sideband, (d) Spectrum of SSB-SC with upper sideband.*

Both these equations (3.97) and (3.98) may be combined as

$$s(t)_{SSB} = \cos w_m t \cos w_c t \pm \sin w_m t \sin w_c t \quad ...(3.99)$$

Here, the (+) sign represents the lower sideband and (–) sign represented the upper sideband. We may write the terms $\sin w_c t$ and $\sin w_c t$ as

$$\sin w_m t = \cos\left(w_c t - \frac{\pi}{2}\right) \quad ...(3.100)$$

$$\sin w_m t = \cos\left(w_c t - \frac{\pi}{2}\right) \quad ...(3.101)$$

This means that the sine terms may be obtained using the corresponding cosine terms simply by giving a phase shift of $(-\pi/2)$.

The expression in equation (3.101) represents the SSB-SC wave for the case of single-tone modulation, however, it provides a way to get a general expression for SSB-SC wave.

In the expression of equation (3.99), the term $\sin w_m t$ is obtained by giving a phase shift of $(-\pi/2)$ to the modulating frequency $\cos w_m t$.

Similarly, in a general modulating signal $x(t)$, if all the frequency components are shifted by $(-\pi/2)$, it may lead to a general expression of SSB-SC signal.

As this point, it may be noted that $x(t)$ may be expressed as a continuous sum of sinusoidal signals.

Hence, expression in equation (3.99) may be extended for a SSB-SC signal modulated by a general modulating signal $x(t)$ as expressed below:

$$s(t)_{SSB} = x(t) \cos w_c t \pm x_h(t) \sin w_c t \quad ...(3.102)$$

where $x_h(t)$ is a signal obtained by shifting the phase of every component present in $x(t)$ by $(-\pi/2)$

Similar to equation (3.99), (+) signal corresponds to the lower sideband and (–) sign corresponds to the upper sideband.

3.21. Hilbert Transform *(Very Important)*

It may be observed that the function $x_h(t)$ obtained by providing $(-\pi/2)$ phase shift to every frequency component present in $x(t)$, actually represents the **Hilbert transform** of $x(t)$. This means that $x_h(t)$ is the Hilbert transform of $x(t)$ defined as

$$x_h(t) = \frac{1}{\pi} x(t) \otimes \frac{1}{t} = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(\tau)}{t-\tau} d\tau \quad ...(3.103)$$

Also, the inverse Hilbert transform is defined as

$$x(t) = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x_h(\tau)}{t-\tau} d\tau \quad ...(3.104)$$

Example 3.7. Show that if every frequency component of a signal $x(t)$ is shifted by an amount $\pi/2$, then the resultant signal $x_h(t)$ is the Hilbert transform of $x(t)$.

Solution: The given situation may be considered as though the signal $x(t)$ is passed through a phase shifting system having transfer function $H(\omega)$ and the output is $x_h(t)$ as shown in figure 3.30.

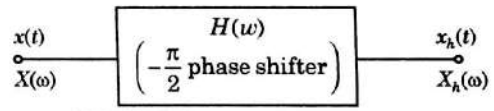


Fig. 3.30. *A phase shifting system.*

The characteristics of this system may be specified as under:

(*i*) The magnitude of the frequency components present in $x(t)$ remains unchanged when it is passed through the system. This means that $H(\omega) = 1$, and

(*ii*) The phase of the positive frequency components is shifted by $-\frac{\pi}{2}$. Now, since the phase spectrum $\theta(\omega)$ has an odd symmetry, the phase of the negative frequency components is shifted by $+\frac{\pi}{2}$. The spectrums $H(\omega)$ and $\theta(\omega)$ have been plotted in figure 3.31.

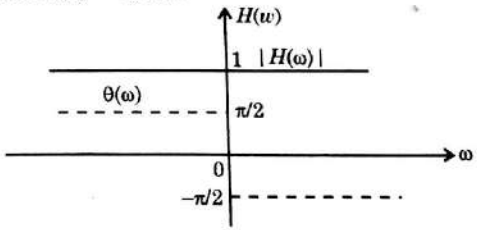


Fig. 3.31. *Transfer function of $-\pi/2$ phase shifter.*

The transfer function is expressed as

$$H(\omega) = |H(\omega)| e^{j\theta(\omega)}$$

or

$$H(\omega) = 1.e^{j\theta(\omega)} \quad ...(3.105)$$

From figure 3.31, it may be observed that

$$\theta(\omega) = \begin{cases} +\frac{\pi}{2} \text{ for } \omega < 0 & (i.e., \text{negtive frequencies}) \\ -\frac{\pi}{2} \text{ for } \omega > 0 & (i.e., \text{positive frequencies}) \end{cases} \quad ...(3.106)$$

Therefore equation (3.105) may be modified as

$$H(\omega) = \begin{cases} e^{j\pi/2} & \text{for } \omega < 0 \\ e^{-j\pi/2} & \text{for } \omega > 0 \end{cases} \quad ...(3.107)$$

Also, we know that $e^{j\frac{\pi}{2}} = \cos\frac{\pi}{2} + j\sin\frac{\pi}{2} = j$

and $e^{-j\frac{\pi}{2}} = \cos\left(-\frac{\pi}{2}\right) + j\sin\left(-\frac{\pi}{2}\right) = -j$

Thus, $H(\omega)$ becomes

$$\frac{H(\omega)}{j} = \begin{cases} 1 & \text{for } \omega < 0 \\ -1 & \text{for } \omega > 0 \end{cases} = -\,sgn\,(\omega)$$

or $$H(\omega) = -j\,sgn\,(\omega) \quad ...(3.108)$$

The response $X_h(\omega)$ of the phase shifting system is related to the input $X(\omega)$ as

$$X_h(\omega) = X(\omega) \cdot H(\omega) \quad ...(3.109)$$

where $x(t) \longleftrightarrow X(\omega)$

and $x_h(t) \longleftrightarrow X_h(\omega)$

Now, substituting the value of $H(\omega)$ in equation (3.109) from equation (3.108), we get

$$X_h(\omega) = -jX(\omega)\,sgn\,(\omega)$$

Taking the inverse Fourier transform of both sides of last equation, we get

$$x_h(t) = F^{-1}[-j.X(\omega)\,sgn(\omega)]$$

Also the time-domain of $sgn(\omega)$ is given as

$$\frac{1}{\pi t} \leftrightarrow sgn(\omega)$$

Using time convolution theorem, we get

$$x_h(t) = \frac{1}{\pi}\left[x(t) \otimes \frac{1}{t}\right] = \frac{1}{\pi}\int_{-\infty}^{\infty}\frac{x(\tau)}{t-\tau}\,d\tau$$

which is the Hilbert transform of $x(t)$.

Some applications of the Hilbert transform may be listed as under:

(*i*) for generation of SSB signals,

(*ii*) for designing of minimum phase type filters,

(*iii*) for representation of bandpass signals.

Properties of the Hilbert Transform

Following are the properties of Hilbert transform:

(*i*) A signal $x(t)$ and its Hilbert transform $x_h(t)$ have the same energy density spectrum.

(*ii*) A signal $x(t)$ and its Hilbert transform $x_h(t)$ have the same autocorrelation function.

(*iii*) A signal $x(t)$ and its Hilbert transform $x_h(t)$ are mutually orthogonal.

Mathematically,

$$\int_{-\infty}^{\infty} x(t)x_h(t)\,dt = 0 \quad ...(3.110)$$

(*iv*) If $x_h(t)$ is a Hilbert transform of $x(t)$, then the Hilbert transform of $x_h(t)$ is $-x(t)$, *i.e.*,

If $H[x(t)] = x_h(t)$

then $H[x_h(t)] = -x(t)$

Here 'H' denotes the Hilbert transform.

The Pre-envelope or Analytic Signal

The concept of pre-envelope, also called as the analytic function is quite useful in deriving the general expression of the SSB-SC signal.

The pre-envelope of a real-valued signal $x(t)$ is defined as

$$x_p(t) = x(t) + jx_h(t) \quad ...(3.111)$$

where $x_h(t)$ is the Hilbert transform of signal $x(t)$. Clearly the pre-envelope $x_p(t)$ is a complex-valued signal. The real part of $x_p(t)$ is $x(t)$, and the imaginary part is its Hilbert transform $x_h(t)$. The complex conjugate of the pre-envelope denoted by $x_p(t)$ is expressed as

$$x_p^* = x(t) - jx_h(t) \quad ...(3.111a)$$

3.22. SSB-SC for a General Modulating Signal

In previous section, we have already derived an expression for a generalized SSB-SC signal by extending the concept of SSB-SC for a signal-tone modulation. In this article let us derive the same expression (*i.e.*, equation) by using the concept of pre-envelope or analytic signal.

As we know the pre-envelope of a function $x(t)$ is defined as

$$x_p(t) = x(t) + jx_h(t)$$

The Fourier transform of $x_p(t)$ is the sum of the Fourier transforms of $x(t)$ and $x_h(t)$ *i.e.*,

$$F[x_p(t)] = F[x(t)] + jF[x_h(t)]$$

or $$X_p(\omega) = X(\omega) + j\,[-j\,X(\omega)\,sgn\,(\omega)]$$

$$= X(\omega) + X(\omega)\,sgn\,(\omega) \quad ...(3.112)$$

We know that

$$sgn\,(\omega) = \begin{cases} 1 & \text{for } \omega > 0 \\ -1 & \text{for } \omega < 0 \end{cases}$$

Therefore, we have

$$X_p(\omega) = \begin{cases} 2X(\omega) & \text{for } \omega > 0 \\ 0 & \text{for } \omega < 0 \end{cases} \quad ...(3.113)$$

Figure 3.32(*a*) and (*b*) show $X(\omega)$ and $X_p(\omega)$ respectively. It is obvious from figure 3.32(*b*) that $X_p(\omega)$ vanishes for negative frequencies.

Similarly, we can find the Fourier transform of $x_p^*(t)$ defined by equation (3.111)

$$X_p^*(\omega) = X(\omega) - j\,[-j \cdot X(\omega)\,sgn\,(\omega)]$$

or $$X_p^*(\omega) = X(\omega) - X(\omega)\,sgn\,(\omega)$$

which may be written as

$$X_p^*(\omega) = \begin{cases} 0 & \text{for } \omega > 0 \\ 2X(\omega) & \text{for } \omega < 0 \end{cases}$$

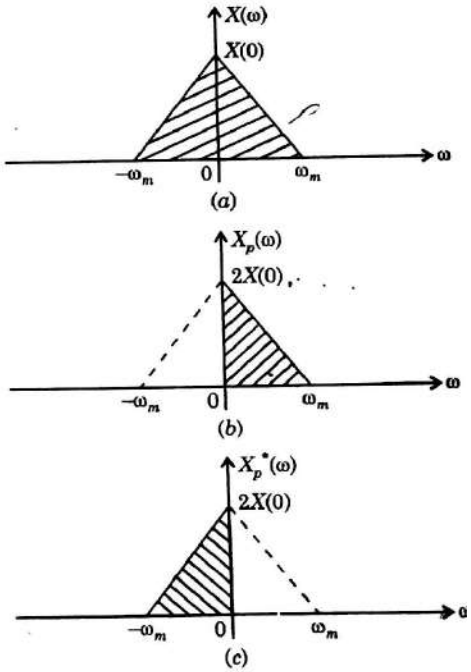


Fig. 3.32.

This has been plotted in figure 3.32(*c*). Clearly, $X_p^*(\omega)$ vanishes for positive frequencies.

Now, let us take an SSB-SC wave consisting of only the lower sidebands of a general modulatin signal $x(t)$ as shown in figure 3.33.

Following points may be observed from figure 3.33.

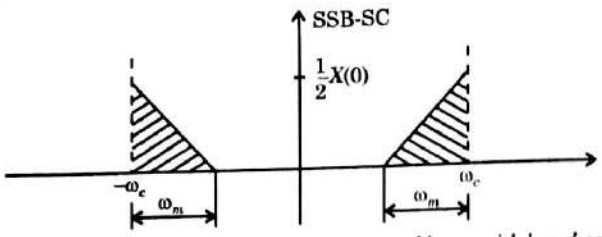


Fig. 3.33. *A SSB-SC spectrum consisting of lower sideband only.*

(*i*) The right-hand portion of the figure represents a spectrum of $\frac{1}{4} x_p^*(t)\, e^{j\omega_c t}$. This is equivalent to shifting the spectrum $X_p^*(\omega)$ towards right by ω_c.

(*ii*) Similarly, the left-hand portion of figure 3.34 represents the spectrum of $\frac{1}{4} x_p(t)\, e^{-j\omega_c t}$.

Thus, figure 3.33 represents the spectrum of a combined signal $x_p(t)\, e^{+j\omega_c t} + x_p(t)\, e^{-j\omega_c t}$. Therefore, the time-domain representation of the SSB-SC spectrum shown in figure 3.34 is expressed as

$$s(t)_{SSB} = \frac{1}{4}[x_p^*(t)\, e^{j\omega_c t} + x_p(t)\, e^{-j\omega_c t}]$$

Substituting in terms of Hilbert transform from equations (3.110) and (3.111), we get

$$s(t)_{SSB} = \frac{1}{4}[x(t) - jx_p(t)]\, e^{j\omega_c t} + \frac{1}{2}[x(t) + jx_h(t)]\, e^{-j\omega_c t}$$

or $$s(t)_{SSB} = \frac{1}{2} x(t) \left[\frac{e^{j\omega_c t} + e^{-j\omega_c t}}{2}\right] + \frac{j}{2} x_h(t) \left[\frac{e^{-j\omega_c t} - e^{j\omega_c t}}{2}\right]$$

or $$s(t)_{SSB} = \frac{1}{2}[x(t) \cos\omega_c t + x_h(t) \sin \omega_c t] \quad ...(3.114)$$

which is the time-domain description of the SSB-SC signal consisting of only the lower sidebands.

Similarly, we can derive an expression for an SSB-SC signal consisting of the upper sidebands as under:

$$s(t)_{SSB} = \frac{1}{2}[x(t) \cos\omega_c t - x_h(t) \sin \omega_c t] \quad ...(3.115)$$

Hence, the time-domain description of the SSB-SC wave is represented by the expression as under:

$$s(t)_{SSB} = x(t) \cos \omega_c t \pm x_h(t) \sin \omega_c t \quad ...(3.116)$$

where '+' and '–' signs correspond to the lower sidebands and upper sidebands respectively.

3.23. Generation of SSB-SC Signal

SSB-SC signals may be generated by two methods as under:

(*i*) Frequency discrimination method or filter method

(*ii*) Phase discrimination method or phase-shift method

3.23.1. Frequency Discrimination Method

In a frequency discrimination method, firstly, a DSB-SC signal is generated simply by using an ordinary product modulator or a balanced modulator. After this, from the DSB-SC, signal one of the

two sidebands is filtered out by a suitable bandpass filter (BPF). The schematic diagram for this method is shown in figure 3.34.

Infact, the design of bandpass filter is quite critical and thus puts same limitations on the modulating or baseband and carrier frequencies

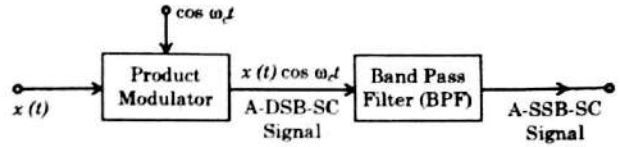


Fig. 3.34. *Frequency-discrimination method for SSB-SC generation.*

Following are the limitations for frequency discrimination method:

(*i*) The frequency-discrimination method is useful only if the baseband signal is restricted at its lower edge due to which the upper and lower sidebands are non-overlapping. For example, the filter method is used for speech communication where lowest spectral component is 70 Hz and it may be taken as 300 Hz without affecting the intelligibility of the speech signal. However, the system is not useful for video communication where the baseband signal starts from d.c.

(*ii*) The another restriction of the frequency discrimination method is that the baseband signal must be appropriately related to the carrier frequency. Infact, the design of the bandpass filter (BPF) becomes difficult if the carrier frequency is quite higher than the bandwidth of the baseband signal.

3.23.2. Phase-Shift Method

The phase-shift method avoids filter. This method makes use of two balanced modulators and two phase-shifting networks as shown in figure 3.35.*

In figure 3.35, one of the modulators, M_1 receives the carrier voltage shifted by 90° and the modulating voltage, whereas another balanced modulator M_2 receives the modulating voltage shifted by 90° and the carrier voltage.

Both balanced modulators produce an output consisting only of sidebands.

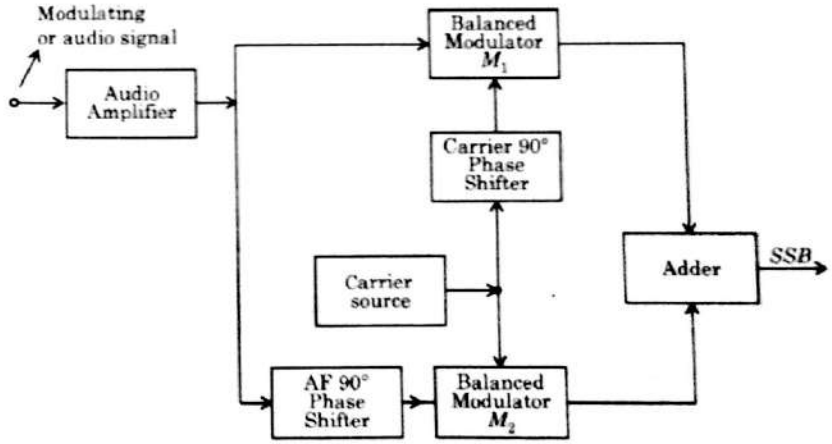


Fig. 3.35. *Phase-shift method for SSB-SC generation.*

* Draw the block diagram of phase cancellation SSB generation and explain how the carrier and the unwanted sideband are suppressed. *(U.P. Tech, Sem. Exam., 2004-05)*

Both the upper sidebands leads the input carrier voltage by 90°. One of the lower sidebands leads the reference voltage by 90° and the other lags it by 90°. The two lower sidebands are thus out of phase, and when combined together in the adder, they cancel each other. The upper sidebands are in phase at the adder and hence they add producing SSB in which the lower sideband has been cancelled.

Mathematical Analysis

Let the expression for carrier be $\sin w_c t$ and that for modulating signal $\sin w_m t$.

Now, the balanced modulator M_1 will receive $\sin w_m t$ and $\sin (w_c t + 90°)$ whereas M_2 will receive $\sin (w_m t + 90°)$ and $\sin w_c t$.

We know that the output of balanced modulator M_1 will contain sum and difference frequencies.

Hence, $$v_1 = \cos\left[(w_c t + 90°) - w_m t\right] - \cos\left[(w_c t + 90°) + w_m t\right]$$

or $$v_1 = \underbrace{\cos (w_c t - w_m t + 90°)}_{\text{(LSB)}} - \underbrace{\cos (w_c t + w_m t + 90°)}_{\text{(USB)}} \quad ...(3.117)$$

Similarly, the output of balanced modulator M_2, will contain

$$v_2 = \cos\left[w_c t - (w_m t + 90°)\right] - \cos\left[(w_c t + (w_m t + 90°)\right]$$

or $$v_2 = \cos (w_c t - w_m t - 90°) - \cos (w_c t + w_m t + 90°) \quad ...(3.118)$$

Therefore, the ouput of the adder will be

$$v_0 = v_1 + v_2 = 2 \cos (w_c t + w_m t + 90°) \quad ...(3.119)$$

which is the expression for SSB-SC.

3.24. Demodulation of SSB-SC Signals

The baseband or modulating signal $x(t)$ can be recovered from the SSB-SC signal by using the synchronous detection technique as already discussed for DSB-SC signals. With the help of synchronous detection method the spectrum of an SSB-SC signal centered about $\omega = \pm \omega_c$, is retranslated to the baseband spectrum which is centered about $\omega = 0$. The process of synchronous detection involves multiplication of the received SSB-SC signal with a locally generated carrier as shown in figure 3.36(*a*).

The output of the multiplier will be

$$e_d(t) = s(t)_{SSB} \cdot \cos \omega_c t$$

or $$e_d(t) = [x(t) \cos \omega_c t \pm x_h(t) \sin \omega_c t] \cos \omega_c t$$

or $$e_d(t) = x(t) \cos^2 \omega_c t \pm x_h(t) \sin \omega_c t \cos \omega_c t$$

or $$e_d(t) = \frac{1}{2} x(t) [1 + \cos 2 \omega_c t] \pm \frac{1}{2} x_h(t) \sin 2 w_c t$$

or $$e_d(t) = \frac{1}{2} x(t) + \frac{1}{2} [x(t) \cos 2 \omega_c t \pm x_h(t) \sin 2 \omega_c t] \quad ...(3.120)$$

Now when $e_d(t)$ is passed through a low-pass filter, then the terms centered about $\pm 2\omega_c$ are filtered out and we get, at the output of detector, signal e_0 which is given as

$$e_0(t) = \frac{1}{2} x(t) \quad ...(3.121)$$

The frequency-domain explanation is evident from figure 3.36. Multiplication of the SSB signal with $\cos \omega_c t$ in the time-domain is equivalent to the convolution of their spectra as shown in figure 3.36(*d*). The component centred around $\pm 2\omega_c$ is filtered out where as the message signal centered at $\omega = 0$ appears at the output. The synchronous detection can be achieved by a ring demodulator circuit or a modulator circuit using non-linear devices.

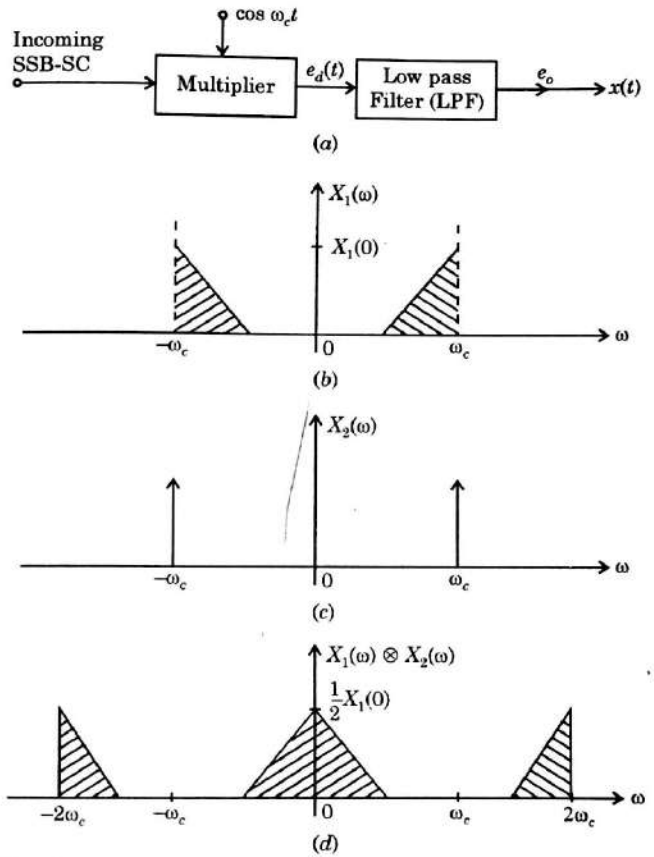


Fig. 3.36. *Demodulation of SSB-SC signal (a) Synchronous detector (b) SSB-SC (lower sideband) (c) Spectrum of cos $\omega_c t$ (locally generated carrier) (d) Convolution of figures (b) and (c).*

3.24.1. Phasor Diagram and Waveform of SSB-SC Signals

In this article, let us consider a single tone modulating signal $V_1 \cos \omega_1 t$ which modulates a carrier signal $V_c \cos \omega_c t$ to generate an SSB-SC signal.

The resulting signal considering lower sideband, will be given as

The signal represents a single sinusoid with an amplitude $\frac{1}{2} V_1 V_C$ and a frequency $(\omega_c - \omega_1)$.

Therefore the phasor diagram consists of a single phasor and there will be no amplitude fluctuation in the modulated wave. This is obvious from the phasor diagram in figure 3.37(*a*) and the waveform drawn in figure 3.37(*b*). It is also clear from the waveform that both amplitude and frequency of the wave are constant. This means that there is no amplitude modulation.

$$s(t)_{SSB} = \frac{1}{2} V_1 V_c \cos (\omega_c - \omega_1) t \quad ...(3.122)$$

Hence, an SSB-SC signal modulated by a signal tone modulating signal does not carry any useful information. However, the baseband signals appearing in actual systems are multiple-tone rather than single tone.

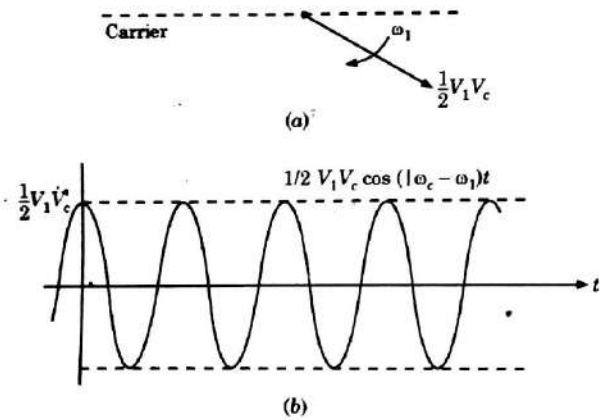


Fig. 3.37. *(a) Phasor diagram of a single tone SSB-SC signal*
(b) Waveform of a single tone SSB-SC signal

Let us consider a two tones baseband signal expressed as

$$x(t) = V_1 \cos \omega_1 t + V_2 \cos \omega_2 t$$

Figure 3.38(*a*) shows the phasor diagram of an SSB-SC signal modulated by the baseband signal $x(t)$. The resultant phasor varies with time as the two phasors rotate. The amplitude variation in the resulting waveform is shown in figure 3.38(*b*). Hence the multiple tones SSB-SC signal shows amplitude variation.

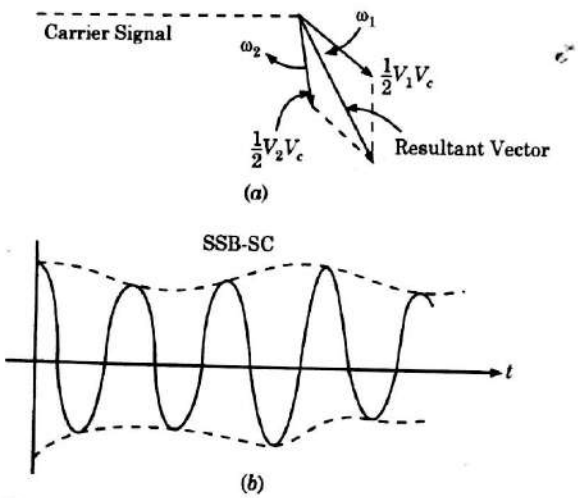


Fig. 3.38. *(a) Phasor diagram of a two-tones SSB-SC signal.*
(b) Waveform of a two-tones SSB-SC signal.

3.24.2. Waveform of SSB-SC with Large Carrier Signal

In this article, let us consider an SSB-SC signal in which a large carrier is also present. This carrier signal can be introduced at the transmitter as in simple AM, or can be added at the receiver end. When a large carrier is added to the SSB-SC signal the demodulation becomes quite easy. When a large carrier signal is present, the SSB-SC waveform shows amplitude variations even for a single-tone modulating signal.

The expression for such a wave may be obtained by adding a large carrier terms in the expression of SSB-SC given in equation (3.122) *i.e.*,

$$s(t)_{SSB} = \frac{1}{2} V_1 V_c \cos(\omega_c - \omega_1)\, t + V_c \cos \omega_c t \text{ for } V_c > V_1$$

The phasor diagram has been shown in figure 3.39. The SSB-SC waveform is formed by the super position of two waveforms (*i.e.*, lower sideband and large carrier term).

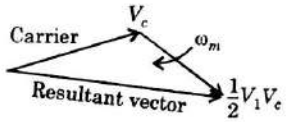


Fig. 3.39. *Phasor diagram of an SSB-SC signal with large carrier.*

3.24.3. Detection of the SSB-SC Signal having Large Carrier

Whenever a large carrier is introduced, a synchronous detection is not essential to recover the modulating signal. In this situation, an envelope detector provides the approximate modulating signal. An envelope detector is quite simpler and cheaper than synchronous detector. Thus, an expression for a SSB-SC signal with a large carrier $A \cos \omega_c t$ may be written as

$$s(t)_{SSB} = x(t) \cos \omega_c t + x_h(t) \sin \omega_c t + A \cos \omega_c t \qquad \text{...(3.12}$$

Here, it may be noted that the carrier is added to an SSB-SC signal (*i.e.*, not multiplied with it as in synchronous detection). This technique is also known as **carrier re-insertion technique.**

Equation (3.123) may be written as

$$s(t)_{SSB} = [A + x(t)] \cos \omega_c t + x_h(t) \sin \omega_c t$$

or

$$s(t)_{SSB} = e(t) \cos(\omega_c t + \theta)$$

where

$$e(t) = \sqrt{[A + x(t)]^2 + x_h^2(t)}$$

is the envelope of the wave and

$$\theta = -\tan^{-1}\left[\frac{x_h(t)}{A + x(t)}\right]$$

is the phase of the wave. When this SSB-SC signal is applied to an envelope detector, the output of the detector will be the envelope $e(t)$ *i.e.*,

$$e(t) = \left[\{A + x(t)\}^2 + x_h^2(t)\right]^{\frac{1}{2}}$$

Simplifying, we get

$$e(t) = A\left[1 + \frac{2x(t)}{A} + \frac{x^2(t)}{A^2} + \frac{x_h^2(t)}{A^2}\right]^{\frac{1}{2}}$$

which may be expanded in the form of a Binomial series, gives as

$$e(t) = A\left[1 + \frac{2x(t)}{2A} + \text{higher terms}\right]$$

But since $A >> |x(t)|$ and $|x_h(t)|$ and higher terms may be neglected, therefore

$$e(t) \equiv A + x(t)$$

Hence, the output of the envelope detector is close to the desired modulating signal $x(t)$. The carrier re-insertion technique is also useful for the detection of DSB-SC signals as discussed already.

This large carrier may be added either at the transmitter or at the receiver. When the carrier signal is added at the receiver, the synchronization problem remains as it is and this technique has no advantage. When carrier is added at the transmitter than this mode has the advantage of both SSB-SC and AM systems. It requires only half the bandwidth, and at the same time, detection becomes quite simple.

3.24.4. Compatible Single Sideband Signal

An SSB-SC signal may be generated in which the carrier is suppressed, even then it can be detected with an envelope detector (*i.e.*, simple diode detector). Such a signal is compatible for reception using a commercial AM radio receiver. The signal with this characteristics is termed as **compatible single sideband (CSSB)**. However, such a signal involves complex signal processing, and the system is presently impractical for commercial applications.

3.25. Vestigial Sideband (VSB) Modulation Systems *(Important)*

A VSB modulation system is actually a compromise between DSB-SC and SSB modulation system. In other words, we can say that it is an optimum choice in which the advantages of DSB-SC and SSB modulation systems have been exploited.

As a matter of fact, the generation of VSB modulation signals is easier than other modulated signals such as conventional AM, DSB-SC and SSB signals. Its bandwidth is only slightly higher (approximately 25%) than SSB signals but considerable less than DSB-SC signals. SSB modulation is rather most suited for the transmission of voice signals because of the energy gap that exists in the frequency spectrum of the vioce signals between zero and few hundred hertz. On the other hand, when signals contain frequency components at extremely low frequencies (as in telegraph signals) the USB and LSB of the translated signal tend to meet at the carrier frequency. Under such circumstances, it becomes very difficult to isolate one sideband from the other. Hence, SSB scheme becomes unsuitable for handling such types of signals.

This difficulty has been overcome in a scheme known as **vestigial sideband (VSB) modulation**. In VSB modulation instead of rejecting one sideband completely as in SSB modulation scheme, a gradual cut-off of one sideband is allowed. This gradual cut is compensated by a vestige or portion of the other sideband.

This technique has been illustrated in figure 3.40 which shows the frequency spectrum of the modulating signal and corresponding DSB-SC, SSB and VSB signals.

Obviously, the bandwidth of VSB signal is given by

$$BW = f_c + f_v - f_c + f_m$$

$$BW = f_m + f_v$$

where f_m = Bandwidth of the message signal, and
f_v = width of the *VSB*.

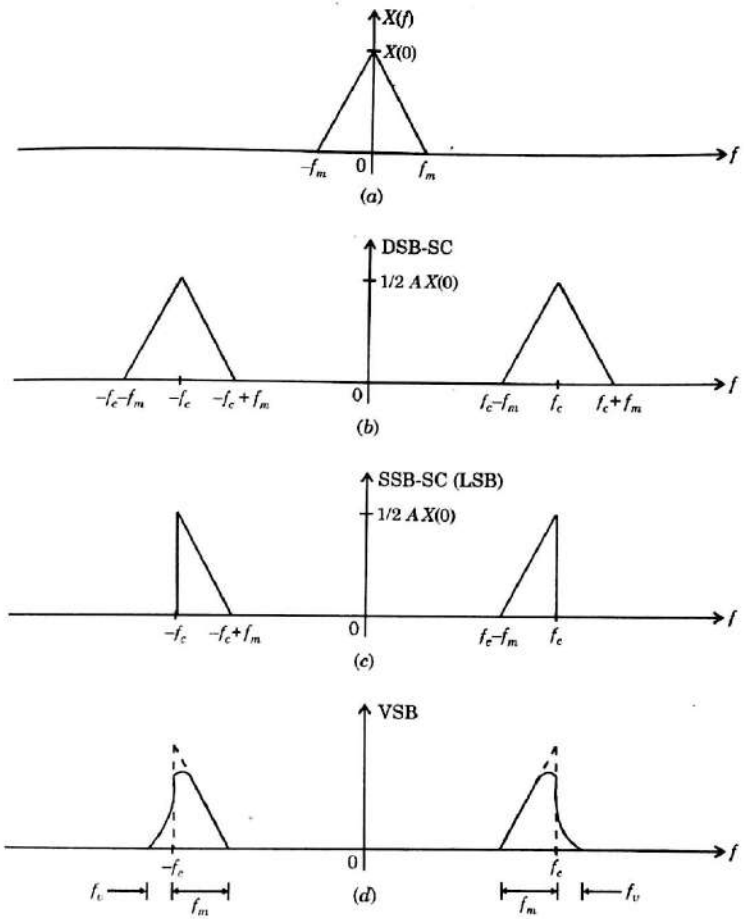


Fig. 3.40. *Illustration of frequency spectrum of VSB signal.*
(a) Frequency spectrum of Baseband signal x(t)
(b) Frequency spectrum of DSB-SC signal
(c) Frequency spectrum of SSB signal
(d) Frequency spectrum of VSB signal.

3.26. Generation of VSB Signals

Basically, VSB signal can be generated by passing a DSB-SC signal through an appropriate filter having transfer function $H(f)$. It has been shown in figure 3.41. The frequency spectrum of VSB signal, $S(f)$ is therefore, given by

$$S(f) = \text{Fourier transform } \{s(t)\}$$

or

$$S(f) = \frac{A}{2}[X(f-f_c) + X(f+f_c)]\,H(f) \qquad \text{...(3.124)}$$

where $X(f)$ is the Fourier transform of modulating signal $x(t)$.

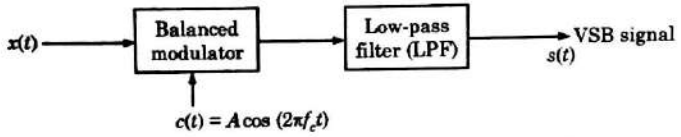


Fig. 3.41. *Scheme for generation of VSB signal.*

Now, let us find the specification of transfer function $H(f)$ of low-pass filter (LPF) correspond to the spectrum $S(f)$ of the VSB signal $s(t)$. This can be established by passing $s(t)$ through a coherent detector without distortion.

We shall, now, derive the expression for transfer function $H(f)$ by determining the necessary conditions for the coherent detection output to provide an undistorted version of the original modulating signal $x(t)$. Thus, multiplying $s(t)$ by a locally generated sine wave $A' \cos 2\pi f_c t)$, which is synchronous with the carrier wave $A \cos(2\pi f_c t)$ in both frequency and phase. It has been shown in figure 3.42.

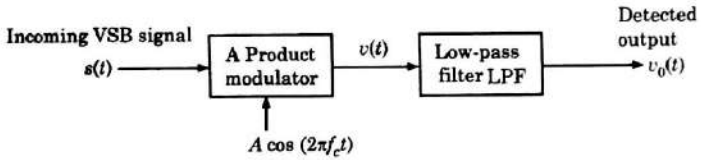


Fig. 3.42. *Block diagram of VSB demodulator.*

The output of the product modulator is given by

$$v(t) = s(t)\, A' \cos(2\pi f_c t)$$

or

$$v(t) = A' s(t) \cos(2\pi f_c t) \qquad ...(3.125)$$

Taking Fourier Transform, we have

$$V(f) = \frac{A'}{2}[S(f-f_c) + S(f+f_c)] \qquad ...(3.126)$$

Substituting equation (3.124) in equation (3.126), we obtain

$$V(f) = \frac{A\,A'}{4}[X(f-2f_c) + X(f)]\,H(f-f_c) + \frac{A\,A'}{4}[X(f) + X(f+2f_c)]\,H(f+f_c)$$

or

$$V(f) = \frac{A\,A'}{4}[H(f-f_c) + H(f+f_c)]\,X(f) + \frac{A\,A'}{4}[X(f-2f_c)\,H(f-f_c) + X(f+2f_c)\,H(f+f_c)] \qquad ...(3.127)$$

The spectrum of $V(f)$ has been shown in figure 3.43. The first term of equation (3.127) corresponds to the frequency spectrum of the modulating signal. The second term of $V(f)$ corresponds to the frequency spectrum of the VSB signal having carrier frequency $2f_c$. The second term can be removed by using low-pass filter (LPF).

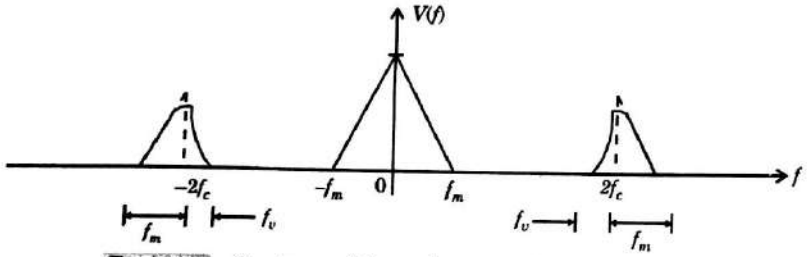


Fig. 3.43. *Spectrum of the product modulator output v(t)*

The frequency spectrum of the signal $v_0(t)$ at the output of the LPF is given by

$$V_0(f) = \frac{A\,A'}{4}[H(f-f_c) + H(f+f_c)]\,X(f) \qquad ...(3.128)$$

For a distortionless reproduction of the original modulating signal $x(t)$ at the output of the coherent detector, we must have $V_0(f)$ as a scaled version of $X(f)$. It has been shown in figure 3.44.

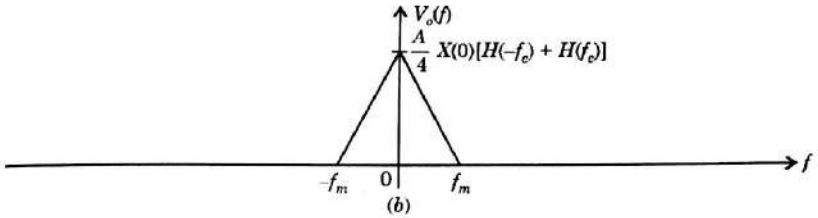


Fig. 3.44. *Spectrum of the demodulated signal $v_0(t)$.*

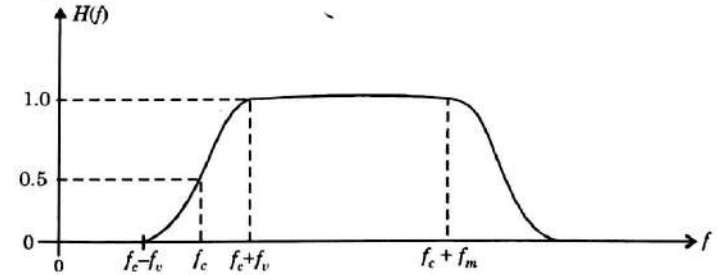
Thus, for distortionless reception by the VSB modulation scheme, it is necessary that the transfer function of the filter $H(f)$ must satisfy the condition given as under:

$$H(f-f_c) + H(f+f_c) = 2H(f_c) \qquad ...(3.129)$$

where $H(f_c)$ is the transfer function of LPF at carrier frequency f_c. It is a constant value.

Now, when the spectrum $X(f)$ is zero outside the interval $-f_m \le f \le f_m$, equation (3.141) is required to be satisfied only in the above range. This requirement can be fulfilled by using a filter having a normalised amplitude response as shown in figure 3.45 for positive frequencies. The magnitude of normalised $|H(f)|$ should be one-half at the carrier frequency f_c. The sum of the values of $|H(f)|$ at any two frequencies equally spaced above and below f_c, should be unity in the transition band $f_c - f_v \le f \le f_c + f_v$. The design of such a filter is much more simpler than the design of filters needed in generation of SSB signals.

In order to preserve the frequency spectrum of $x(t)$ properly, the phase spectrum of the signal should be linear in the interval $f_c - f_v \le |f| \le f_c + f_m$ and its value should be equal to zero at f_c, or an integral multiple of 2π.



@sTc

Fig. 3.45. *Illustration of amplitude response of VSB filter.*

3.26.1. The Time-Domain Representation of the VSB Signal

The time-domain representation of the VSB signal can be obtained by using the canonical form of the bandpass signal given by

$$s(t) = s_c(t) \cos(2\pi f_c t) - s_s(t) \sin(2\pi f_c t) \qquad ...(3.129a)$$

Frequency spectrum of VSB signal will be

$$S(f) = \frac{A}{2}[X(f-f_c) + X(f+f_c)]\,H(f) \qquad ...(3.129b)$$

Fourier transform of the in phase component $s_c(t)$ will be given by

$$S_c(f) = \begin{cases} S(f-f_c) + S(f+f_c), \text{ for } -f_m \le f \le f_m \\ 0, \qquad \text{otherwise} \end{cases} \qquad ...(3.130)$$

Substituting equation (3.129b) in equation (3.130), we obtain

$$S_c(f) = \frac{1}{2}A[H(f-f_c) + H(f+f_c)]X(f) \qquad ...(3.131)$$

Here we have assumed that the transfer function $H(f)$ of the VSB filter satisfies the equation (3.129) with $H(f_c) = 0.5$

Thus using equation (3.131), we obtain

$$S_c(f) = \frac{1}{2}A\,X(f) \qquad ...(3.132)$$

Taking the inverse Fourier transform of both sides of equation (3.132), we get the in-phase component of the VSB signal as under:

$$s_c(t) = \frac{1}{2}A\,x(t) \qquad ...(3.133)$$

The Fourier transform of the quadrature components $s_s(t)$ can be obtained by substituting equation (3.131) in equation given below:

$$S_s(f) = \begin{cases} j[S(f-f_c) + S(f+f_c)], \text{ for } -f_m \le f \le f_m \\ 0, \qquad \text{otherwise} \end{cases}$$

$$S_s(f) = \frac{j}{2}A[H(f-f_c) - H(f+f_c)]X(f) \qquad ...(3.134)$$

From equation (3.134), it may be observed that we can generate the quadrature component $s_s(t)$, except for a scaling factor, by passing the message signal through a filter whose transfer function is given by

$$H_s(f) = j[H(f-f_c) - H(f+f_c)] \qquad ...(3.135)$$

The frequency response of the filter has been shown in figure 3.46, assuming that $H(f)$ is defined as in figure 3.46. Let $x_s(t)$ be the output of this filter in response to $x(t)$. Hence, we can express the quadrature component of the VSB signal as

$$s_s(t) = \frac{1}{2}A.x_s(t) \qquad ...(3.136)$$

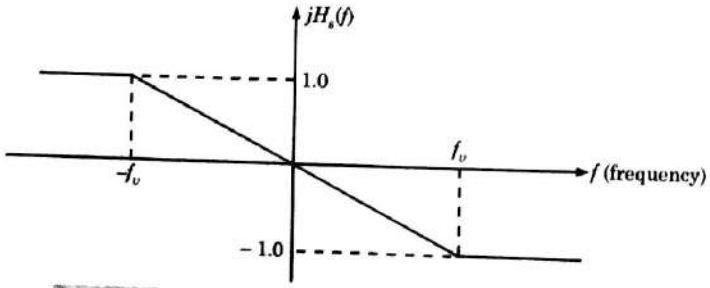


Fig. 3.46. *Frequency response of the filter to produce $s_s(t)$.*

Substituting equations (3.133) and (3.136) in equation (3.129a), we get the time-domain representation of VSB signal as under:

$$s(t) = \frac{A}{2}[x(t)\cos(2\pi f_c t) - x_s(t)\sin(2\pi f_c t)] \qquad ...(3.137)$$

A scheme for generating VSB signal on the basis of the time-domain representation has been illustrated in figure 3.47.

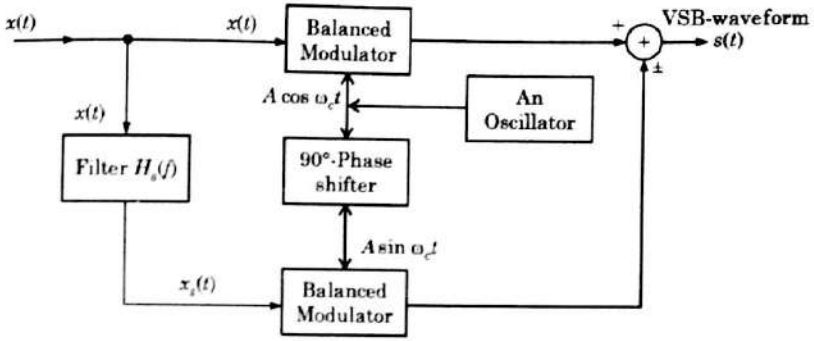


Fig. 3.47. *Illustration of Block diagram of phase discrimination method for generating VSB signals.*

Note: *By following a procedure similar to the above, the time-domain representation of VSB signal containing a vestige of the LSB may also be obtained in a similar form as in equation (3.137) except that a plus sign appears on the RHS in place of minus sign. Both types of VSB can be obtained by using the generalised scheme shown in figure 3.46, taking care of adding and subtracting the outputs of the two product modulators.*

3.27. Demodulation of VSB Signals *(Expected)*

VSB signals can be detected using coherent detection or synchronous detection technique. However, we use another scheme in which an envelope detector is used for demodulating VSB signal containing a large carrier. In commercial television broadcasting, a sizeable carrier is usually transmitted with the video signal which is sent using VSB. We can represent VSB + carrier using equation given as follows:

$$s(t) = \frac{A}{2}[x(t)\cos(2\pi f_c t) - x_s(t)\sin(2\pi f_c t)] \qquad ...(3.138)$$

as
$$s(t) = \frac{A}{2}\left[1 + \frac{k_a}{2}x(t)\right]\cos(2\pi f_c t) - \frac{k_a x_s(t)}{2}\sin(2\pi f_c t)$$

where k_a is the scaling factor.*

The last equation can be written as under:

$$s(t) = e(t)\cos(2\pi f_c t + \theta) \qquad ...(3.139)$$

where
$$e(t) = \sqrt{A^2\left[1 + \frac{k_a x(t)}{2}\right]^2 + A^2\left[\frac{k_a x_s(t)}{2}\right]^2} \qquad ...(3.140)$$

* Scaling factor k_a determines the percentage modulation.

and $$\theta = \tan^{-1}\left[\frac{\frac{1}{2}k_a x_s(t)}{1+\frac{1}{2}k_a x(t)}\right] \quad ...(3.141)$$

The envelope of $s(t)$ can thus be written as under:

$$e(t) = A\left[1+\frac{1}{2}k_a x(t)\right]\left[1+\left\{\frac{\frac{1}{2}k_a x_s(t)}{1+\frac{1}{2}k_a x(t)}\right\}^2\right]^{\frac{1}{2}} \quad ...(3.142)$$

Note: *It can be observed from equation (3.142) that the envelope consists of the modulating signal $x(t)$ and hence can be detected using an envelope detector.*
However, distortion is introduced by the quadrature component $x_s(t)$ of the incoming VSB signal. This distortion can be reduced by reducing the modulation index and by increasing the width of VSB and thus reducing $x_s(t)$.

3.28. Comparison of Various AM Systems

In this section, let us compare conventional AM system, DSB-SC system, SSB system and VSB system of Amplitude modulation. We know that the conventional AM signal contains the two sidebands as well as carrier. In suppressed carrier schemes, the carrier is removed before transmission process. The relative merits and demerits of various forms of AM can be summarized below:

(*i*) The demodulation or detection of AM signal is simpler than DSB-SC and SSB systems. The conventional AM can be demodulated by rectifier or envelope detector. Detection of DSB-SC and SSB is rather difficult and expensive alsop. Furthermore, it is quite easier to generate conventional AM signals at high power levels as compared to DSB-SC and SSB signals. For this reason, conventional AM systems are used for broadcasting purpose.

(*ii*) The advantage of DSB-SC and SSB systems over conventional AM system is that the former requires lesser power to transmit the same information. For sinusoidal modulation the carrier consumes about 2/3 of the total power for 100% modulation in conventional AM. However, only 1/3 of the total power is carried by the sidebands which carry the information. This makes AM transmitters less coefficient. On the other hand, the receivers of DSB-SC and SSB systmes, though efficient, are much more complex and expensive too. Due to this reason, DSB-SC and SSB systems only find applications in point-to-point communication. In point-to-point communication only a few receives and one transmitter are needed. In public broadcast system one transmitter caters to millions of receivers. It is obviously needed to the simpler and cheaper.

(*iii*) SSB scheme needs only one-half of the bandwidth required in DSB-SC system and less than that required in VSB also. Thus, we can say that SSB modulation scheme is the most-efficient scheme among DSB-SC and VSB schemes. SSB modulation scheme is used for long distance transmission of voice signals because it allows longer spacing between repeaters.

3.29. Frequency Translation and Mixing

In the processing of signals in communication systems, it is often required to translate or shift the modulated signal to a new frequency band. As an example, in most commercial AM radio receivers, the received radio-frequency (RF) signal [540 to 1640 kilohertz (kHz)] is shifted to the intermediate-

frequency (IF) (455 kHz) band for processing. The received signal, now translated to a fixed IF, can easily be amplified, filtered, and demodulated.

A device that performs the frequency translation of a modulated signal is known as a ***frequency mixer.*** The operation is often called ***frequency mixing, frequency conversion,*** or ***heterodyning.***

A common problem associated with frequency mixing is the presence of the *image frequency*. For example, in an AM superheterodyne receiver, the locally generated frequency is chosen to be 455 kHz higher than the incoming signal. Suppose that the reception of an AM station at 600 kHz is desired. The locally generated signal is at 1055 kHz. Now if there is another station at 1510 kHz, it will also be received (note that 1510 kHz − 155 kHz = 455 kHz). The second frequency, 1510 kHz = 600 kHz + 2(455 kHz), is called the image frequency of the first, and after the heterodyning operation it is impossible to distinguish the two. Note that the image frequency is separated from the desired signal by exactly twice the IF. Usually, the image frequency signal is attenuated by a selective RF amplifier placed before the mixer.

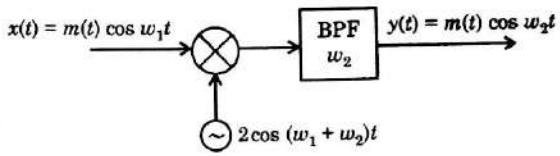


Fig. 3.48. *A Frequency mixer.*

MISCELLANEOUS SOLVED EXAMPLES

Example 3.8. An audio signal given as 15 sin 2 π(1500*t*) amplitude modulates a carrier given as 60 sin 2 π (100,000*t*) determine the following:

(*a*) Sketch the audio signal.
(*b*) Sketch the carrier signal.
(*c*) Construct the modulated wave.
(*d*) Determine the modulation index and per cent modulation.
(*e*) What are the frequencies of the audio signal and the carrier?
(*f*) What frequencies would present in a spectrum analysis of the modulated wave?

Solution: Given that Audio signal = 15 sin 2 π (1500*t*)
Carrier = 60 sin 2 π (100,000*t*)

(*a*) audio signal or modulating signal is sketched in figure 3.49

(*b*) The carrier signal is sketched in figure 3.49

(*c*) To construct the modulated wave, first let us develop the envelope of the modulated wave in the following two steps:

(*i*) Locate the amplitude of the carrier (dashed line).

(*ii*) Using the amplitude of the carrier as an axis, lay in the audio signal.

Now that the envelope has been determined, a signal having an amplitude defined by the envelope found above and having a frequency of the carrier is laid in within the envelope as shown in figure 3.49 (*c*).

(*d*) The modulation index is given by

$$m_a = \frac{\text{Maximum audio amplitude}}{\text{Maximum carrier amplitude}} = \frac{15}{60} = \frac{1}{4} = 0.25 \quad \textbf{Ans.}$$

Further, converting modulation index to per cent modulation, we have

$$M = m_a \times 100 = 0.25 \times 100 = 25\% \quad \textbf{Ans.}$$

(*e*) Since the expression for audio signal is given by

$$v_m = V_m \sin 2\pi f_m t = 15 \sin 2\pi (1500t)$$

Hence $f_m = 1500$ Hz **Ans.**

Since the expression for carrier signal is given by

$$v_c = V_c \sin 2\pi f_c t = 60 \sin 2\pi (100{,}000)$$

Therefore, $f_c = 100{,}000$ Hz **Ans.**

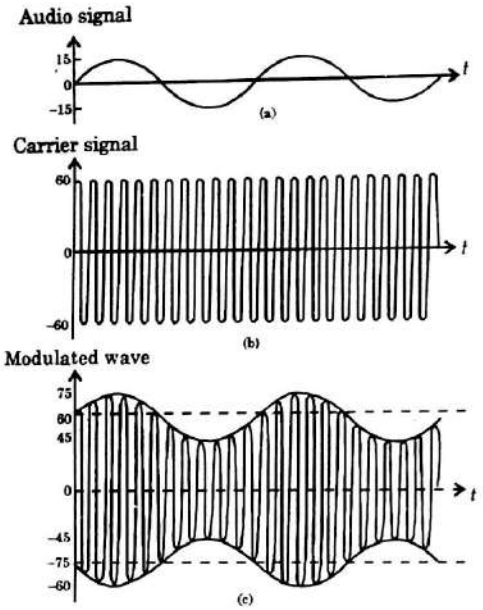


Fig. 3.49. *Waveforms for example 3.8.*

(*f*) We know that the frequency spectrum of an amplitude-modulated wave consists of

$$f_c,\ f_c + f_m \text{ and } f_c - f_m$$

$$f_c = 100{,}000 \text{ Hz}$$

$$f_c + f_m = 100{,}000 + 1500 = 101{,}500 \text{ Hz}$$

$$f_c - f_n = 100{,}00 - 1500 = 98{,}500 \text{ Hz}$$

Therefore, the frequency content of the modulated wave will be

100,000 Hz
101,500 Hz
98,500 Hz **Ans.**

Example 3.9. The maximum power efficiency of an AM modulator is

(*a*) 25% **(*b*) 50%**
(*c*) 75% **(*d*) 100%** *(GATE Examination-1998)*

Solution: We know that $P_t = P_c\left(1 + \frac{m^2}{2}\right) = \frac{A_c^2}{R}\left(1 + \frac{m^2}{2}\right)$

Also, useful power $P_u = \frac{m^2}{2} \cdot \frac{A_c^2}{R}$

and $P_c = \frac{A_c^2}{R}$

Hence, efficiency $\eta = \frac{P_u}{P_c} = \frac{m^2}{2}$

But $m \le 1$, therefore $\eta_{max} = \frac{1}{2}$ *i.e.*, 50%

Thus, option (*b*) is correct.

Example 3.10. A 75 MHz carrier signal having an amplitude of 50 V is modulated by a 3 kHz audio signal having an amplitude of 20 V.

(*a*) Sketch the audio signal.
(*b*) Sketch the carrier signal
(*c*) Construct the modulated wave.
(*d*) Determine the modulation index and per cent modulation.
(*e*) What frequencies would be there in a spectrum analysis of the modulated wave?
(*f*) Write trigonometric equation for the carrier and the modulating waves.

Solution: Given that:

$$f_c = 75 \text{ MHz};\quad V_c = 50 \text{ V};\quad f_m = 3 \text{ kHz};\quad V_m = 2\ V$$

(*a*) Audio signal or modulating signal is sketched in figure 3.50(*a*).

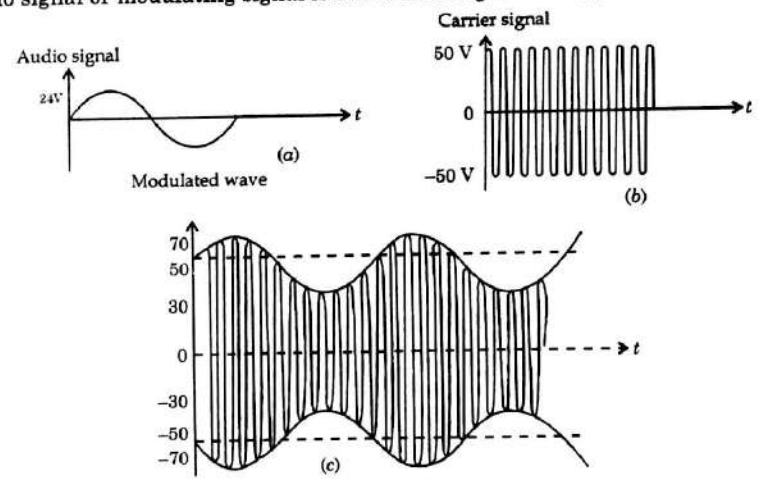


Fig. 3.50.

(*b*) The carrier signal is sketched in figure 3.50(*b*)

(*c*) The envelope of the carrier is fist developed by drawing the horizontal dashed line at the unmodulated carrier amplitude, both positive and negative. The audio signal is now sketched around the dashed line, providing the envelope within which the radio-frequency signal can be laid as shown in figure 3.50(*c*)

(*d*) Modulation index is given as

$$m_a = \frac{V_m}{V_c} = \frac{20}{50} = 0.4 \qquad \textbf{Ans.}$$

Per cent modulation may now be determined by multiplying the modulation index by 100, *i.e.*,

$$M = m_a \times 100 = 0.4 \times 100 = 40\% \qquad \textbf{Ans.}$$

(*e*) The frequency content of an AM signal consists of the carrier frequency and the sideband frequencies which result from adding the audio frequency to the carrier and from subtracting the audio frequency from the carrier frequency as under:

$$f_c = 75 \text{ MHz}$$
$$f_c + f_m = 75 \text{ MHz} + 3 \text{ kHz}$$
$$= 75{,}000 \text{ kHz} + 3 \text{ kHz} = 75{,}003 \text{ kHz}$$
$$f_c - f_m = 75{,}000 \text{ kHz} - 3 \text{ kHz} = 74{,}997 \text{ kHz}$$

Thus, the frequency content of the AM wave will be

75.000 MHz
75.003 MHz
75.997 MHz **Ans.**

(*f*) The expression for modulating signal is given by

$$v_m = V_m \sin 2\pi f_m t$$

where $V_m = 20$ V and $f_m = 3000$ Hz. **Ans.**

Therefore, $v_m = 20 \sin 2\pi (3000)t = 20 \sin 6000\pi t$

The expression for carrier signal is given by

$$v_c = V_c \sin 2\pi f_c t$$

where $V_c = 50$ V and $f_c = 75$ MHz

$$v_c = 50 \sin 2\pi (75 \times 10^6)t = 50 \sin 150 \times 10^6 \pi t \quad \textbf{Ans.}$$

Remember that v_c represents the maximum amplitude of the carrier signal and v_m the maximum amplitude of the modulating signal.

Example 3.11. How many AM broadcast stations can be accommodated in 100 kHz bandwidth if the highest frequency modulating a carrier is 5 kHz?

Solution: Given that:

$$\text{Total BW} = 100 \text{ kHz}$$
$$f_{m_{max}} = 5 \text{ kHz}$$

We know that, any station being modulated by a 5 kHz signal will produce an upper-side frequency 5 kHz above its carrier and a lower-side frequency 5 kHz below its carrier, thereby requiring a bandwidth of 10 kHz. Thus, we have

Number of stations accommodated

$$= \frac{\text{Total BW}}{\text{BW per station}} = \frac{100 \times 10^3}{10 \times 10^3}$$

Hence, number of stations accommodated = 10 stations. **Ans.**

Example 3.12. A bandwidth of 20 MHz is to be considered for the transmission of AM signals. If the highest audio frequencies used to modulate the carriers are not to exceed 3 kHz, how many station could broadcast within this band simultaneously without interfering with one another?

Solution: Given that:

$$\text{Total BW} = 20 \text{ MHz}$$
$$f_{m_{max}} = 3 \text{ kHz}$$

We know that the maximum bandwith of each AM station is determined by the maximum frequency of the modulating signal.

$$\text{Station BW} = 2f_{m_{max}}$$
$$= 2 \times 3 \times 10^3 = 6 \times 10^3 = 6 \text{ kHz}$$

Thus, the number of stations that can be broadcasted simultaneously without interfering with one another will be

$$\frac{20 \times 10^6}{6 \times 10^3} = 3.333 \times 10^3$$

Number of station = 3333 **Ans.**

Example 3.13. The total power content of an AM signal is 1000 W. Determine the power being transmitted at the carrier frequency and at each of the sidebands when the per cent modulation is 100%.

Solution: Given that

$$P_t = 1000 \text{ W}$$
$$m = 100\% = 1$$

We know that the total power consists of the power at the carrier frequency, at the upper sideband frequency and at the lower sideband frequency, *i.e.*,

$$P_t = P_c + P_{USB} + P_{LSB}$$

Form the equation for total power, we have

$$P_t = P_c + \frac{m_a^2 P_c}{4} + \frac{m_a^2 P_c}{4} = P_c + \frac{m^2 P_c}{2}$$

Substituting, given values, we get

$$1000 = P_c + \frac{(1.0)^2 P_c}{2} = P_c + 0.5\, P_c = 1.5 P_c$$

or $$\frac{1000}{1.5} = P_c$$

Solving, we get

$$P_c = 666.67 \text{ W} \quad \textbf{Ans.}$$

The leaves 1000 – 666.67 = 333.33 watts to be shared equally between upper and lower sidebands.

i.e, $P_{USB} + P_{LSB} = 333.33$ W

But, $P_{USB} = P_{LSB}$

Thus, $2P_{LSB} = 333.33$

or $$P_{LSB} = P_{USB} = \frac{333.33}{2} = 166.66 \quad \textbf{Ans.}$$

Example 3.14. Determine the power content of the carrier and each of the sidebands for an AM signal having a per cent modulation of 80% and a total power of 2500 W.

Solution: Given that

$$M = 80\%; \quad m_a = 0.8$$
$$P_t = 2500 \text{ W}$$

We have to find P_c, P_{USB}, and P_{LSB}.

We know that the total power of an AM signal is the sum of the power at the carrier frequency and the power contained in the sidebands, *i.e.*

$$P_t = P_c + P_{USB} + P_{LSB}$$

Using the equation for total power, we write

$$P_t = P_c + \frac{m_a^2 P_c}{4} + \frac{m_a^2 P_c}{4}$$

But $$\frac{m_a^2 P_c}{4} + \frac{m_a^2 P_c}{4} = \frac{m_a^2 P_c}{2}$$

So, $$P_t = P_c + \frac{m_a^2 P_c}{2}$$

$$2500 = P_c + \frac{(0.8)^2 P_c}{2} = P_c + \frac{0.64}{2} P_c = 1.32 P_c = 1.32\ P_c$$

$$P_c = \frac{2500}{1.32} = 1893.9 \text{ W} \quad \textbf{Ans.}$$

The power in the two sidebands is the difference between the total power and the carrier power, *i.e.*,

$$P_{USB} + P_{LSB} = 2500 - 1893.9$$

or $$P_{USB} + P_{LSB} = 606.1 \text{ W}$$

or $$P_{USB} = P_{LSB} = \frac{606.1}{2} \text{W} = 303.50 \text{ W} \quad \textbf{Ans.}$$

Example 3.15. The power content of the carrier of an AM wave is 5 kilowatts. Determine the power content of each of the sidebands and the total power transmitted when the carrier is modulated upto 75%.

Solution: Given that

$$P_c = 5 \text{ kW}$$
$$M = 75\%;\ m_a = 0.75$$

We have to find P_{USB}, P_{LSB}, and P_t

Since, in an AM wave, the power in each of the sidebands is equal, therefore, we have

$$P_{USB} = P_{LSB} = \frac{m_a^2 P_c}{4} = \frac{(0.75)^2 (5000)}{4}$$
$$P_{USB} = P_{LSB} = 703.13 \text{ W} \quad \textbf{Ans.}$$

Now, the total power is the sum of the carrier power and the power in the two sidebands.

Thus, we have

$$P_t = P_c + P_{USB} + P_{LSB}$$

or $$P_t = 5000 + 703.13 + 703.13 = 6406.26 \text{ W} \quad \textbf{Ans.}$$

Example 3.16. Which of the following demodulators can be used for demodulating the signal $x(t) = 5(1 + 2 \cos 2000\ \pi t) \cos 2000\ \pi t$

(*a*) Envelope demodulator

(*b*) Square-law demodulator

(*c*) Synchronous demodulator

(*d*) None of the above. *(GATE Examination-1999)*

Solution: The given signal is

$$x(t) = 5(1 + 2 \cos 2000\pi t) \cos 2000\ \pi t$$

Note that this is essentially an amplitude modulation (AM) double-sideband signal which can be demodulated by any of the 3 demodulators (*a*), (*b*) or (*c*), though envelope detection using a square law device (such as a diode) is most commonly used.

Hence, answers (*a*), (*b*) and (*c*) are correct.

Example 3.17. An amplitude-modulated wave has a power content of 800 W at its carrier frequency. Determine the power content of each of the sidebands for a 90% modulation.

Solution: Given that

$$P_c = 800 \text{ W};\quad M = 90\%;\quad m = 0.90$$

We have to find P_{USB}, P_{USB}

Since, the power in each of the sidebands is equal to $\frac{m_a^2 P_c}{4}$, therefore, we get

$$P_{LSB} = P_{USB} = \frac{m_a^2 P_c}{4} = \frac{(0.9)^2 800}{4}$$

or $$P_{LSB} = P_{USB} = 162 \text{ W} \quad \textbf{Ans.}$$

Example 3.18. Determine the per cent modulation of an amplitude-modulated wave which has a power content at the carrier of 8 kW and 2 kW in each of its sidebands when the carrier is modulated by a simple audio tone.

Solution: Given that

$$P_c = 8 \text{ kW};\quad P_{USB} = P_{LSB} = 2 \text{ kW}$$

We have to find M.

Knowing the power content of the sidebands and the carrier, the relationship of sideband power can be used to determine the modulation index. Once the modulation index is known, merely multiplying it by 100 provides per cent modulation.

$$P_{USB} = P_{LSB} = \frac{m_a^2 P_c}{4}$$

$$2 \times 10^3 = \frac{m_a^2 (8 \times 10^3)}{4}$$

or $$m_a^2 = \frac{4 \times 2 \times 10^3}{8 \times 10^3} = 1.0$$

or $$m_a = 1.0$$

Also, $$M = m_a \times 100 = 100\% \quad \textbf{Ans.}$$

Example 3.19. The total power content of an AM wave is 600 W. Determine the per cent modulation of the signal if each of the sidebands contains 75 W.

Solution: Given that

$$P_t = 600 \text{ W};\quad P_{USB} = P_{LSB} = 75 \text{ W}$$

We have to find M.

In order to determine the per cent modulation, the power contained at the carrier frequency is first determined. Once P_c is known, the relationship between P_c and the sideband power will provide a means of determining the modulation index, from which the per cent modulation is easily found.

Carrier power can be determined as follows:

$$P_t = P_c + P_{USB} + P_{LSB}$$
$$600 = P_c + 75 + 75$$

or $$P_c = 6 - 150 = 450 \text{ watts} \quad \textbf{Ans.}$$

Now, using the relationship between sideband power and carrier power, we have

$$P_{USB} = P_{LSB} = \frac{m_a^2 P_c}{4}$$

$$75 = \frac{m_a^2 (450)}{4}$$

or $$m_a^2 = \frac{4(75)}{450} = 0.667$$

or $$m_a = 0.816 \quad \textbf{Ans.}$$

@sTc

Converting modulation index to per cent modulation, we get

$$M = m_a \times 100 = 0.816 \times 100$$

or $$M = 81.6\% \quad \textbf{Ans.}$$

Example 3.20. Find the per cent modulation of an AM wave whose total power content is 2500 W and whose sidebands each contains 400 W.

Solution: Given that

$$P_t = 2500 \text{ W}; \quad P_{USB} = P_{LSB} = 400 \text{ W}$$

We have to find M.

First find the power contained at the carrier frequency. Then use the relationship between sideband power and carrier power to determine the modulation index. Once the modulation index is known, the per cent modulation can easily be found merely by multiplying by 100.

The power at the carrier frequency can be determined as follows:

$$P_t = P_c + P_{USB} + P_{LSB}$$

$$2500 = P_c + 400 + 400$$

$$P_c = 2500 - 800 = 1700 \text{ W}$$

$$P_{USB} = P_{LSB} = \frac{m_a^2 P_c}{4}$$

$$400 = \frac{m_a^2(1700)}{4}$$

or $$m_a^2 = \frac{400(4)}{1700} = \frac{1600}{1700} = 0.941$$

or $$m_a = 0.970$$

Thus, $$M = 0.970 \times 100 = 97\% \quad \textbf{Ans.}$$

Example 3.21. Determine the power content of each of the sidebands and of the carrier of an AM signal that has a per cent modulation of 85% and contains 1200 W of total power.

Solution: Given that

$$M = 85\%; \quad m_a = 0.85; \quad P_t = 1200 \text{ W}$$

We have to find P_c, P_{USB}, P_{LSB}

Using the expression which relates the total power to carrier power, we have

$$P_t = P_c\left(1 + \frac{m_a^2}{2}\right)$$

Substituting, $$1200 = P_c\left[1 + \frac{(0.85)^2}{2}\right] = P_c\left[1 + \frac{0.7225}{2}\right] = P_c\,[1 + 0.3613] = 1200 = 1.363P_c$$

or $$P_c = \frac{1200}{1.3613} = 881.5 \text{ W} \quad \textbf{Ans.}$$

The sum of carrier power and sideband power is equal to total power, *i.e.*,

$$P_c + P_{SB} = P_t$$

$$881.5 + P_{SB} = 1200$$

$$P_{SB} = 1200 - 881.5 = 318.5 \text{ watts}$$

The total sideband power is made up equally of upper sideband power and lower sideband power *i.e.*,

$$P_{USB} = P_{LSB} = \frac{P_{SB}}{2} = \frac{318.5}{2} = 159.25 \quad \textbf{Ans.}$$

Example 3.22. An AM signal in which the carrier is modulated upto 70% contains 1500 W at the carrier frequency. Determine the power content of the upper and lower sidebands for this per cent modulation. Calculate the power at the carrier and the power content of each of the sidebands when the per cent modulation drops to 50%.

Solution: Given that

$$M = 70\%; \quad m_a = 0.70; \quad P_{c70} = 1500 \text{ W}$$

We have to find P_{USB70}, P_{LSB70}, P_{c50}, P_{USB50}, P_{LSB50}

We know that the power content of each of the sidebands is equal to $m_a^2 P_c/4$, thus, we have

$$P_{USB70} = P_{LSB70} = \frac{m_a^2 P_c}{4} = \frac{(0.7)^2 1500}{4} = 183.75$$

$$P_{USB70} = P_{LSB70} = 183.75 \text{ W} \quad \textbf{Ans.}$$

Since in standard AM transmission, carrier power remains the same, regardless of per cent modulation, thus, we have

$$P_{c50} = P_{c70} = 1500 \text{ W}$$

$$P_{USB50} = P_{LSB50} = \frac{m_a^2 P_c}{4} = \frac{(0.5)^2 1500}{4}$$

$$P_{c50} = 93.75 \text{ W} \quad \textbf{Ans.}$$

Example 3.23. The per cent modulation of an AM wave changes from 40% to 60%. Originally, the power content at the carrier frequency was 900 W. Determine the power content at the carrier frequency and within each of the sidebands after the per cent modulation has risen to 60%.

Solution: Given that

$$M_1 = 40\%; \quad m_1 = 0.40; \quad M_2 = 60\%; \quad m_2 = 0.60$$

$$= P_{c40} = 900 \text{ watts}$$

We have to find P_{c60}, P_{USB60}, P_{LSB60}.

The power content of the carrier of an AM signal remains the same regardless of per cent modulation. Thus, we have

$$P_{c60} = P_{c40} = 900 \text{ W}$$

The power content of each of the sidebands is equal to $m_a^2 P_c/4$.

Thus, we have $$P_{USB60} = P_{LSB40} = \frac{m^2 P_c}{4} = \frac{(0.60)^2(900)}{4}$$

$$P_{USB60} = P_{LSB60} = 81.0\text{W} \quad \textbf{Ans.}$$

Example 3.24. A signal sideband (SSB) signal contains 1 kW. How much power is contained in the sidebands and how much at the carrier frequency?

Solution: Given that

$$P_{SSB} = 1 \text{ kW}$$

We have to find P_{SB}, P_c

We know that in a signal sideband transmission, the carrier and one of the two sidebands are eliminated. Therefore, all the transmitted power is transmitted at one of the sidebands regardless of per cent modulation.

Thus, $$P_{SB} = 1 \text{ kW}$$

and $$P_c = 0 \text{ W} \quad \textbf{Ans.}$$

Example 3.25. An SSB transmission contains 10 kW. This transmission is to be replaced by a standard amplitude modulated signal with the same power content. Determine the power content of the carrier and each of the sidebands when the per cent modulation is 80%.

Solution: Given that

$$P_{SSB} = 10 \text{ kW}; \quad M = 80\%; \; m_a = 0.80$$

We have to find P_c, P_{LSB}, P_{USB}

Since the total power content of the new AM signal is to be the same as the total power content of the SSB signal, therefore, we have

$$P_t = P_{SSB} = 10 \text{ kW}$$

Solving for power contained at the carrier frequency, we get

$$P_t = P_c + P_{LSB} + P_{USB} = P_c + \frac{m_a^2 P_c}{4} + \frac{m_a^2 P_c}{4}$$

$$10{,}000 = P_c + \frac{(0.8)^2 P_c}{4} + \frac{(0.8)^2 P_c}{4} = P_c + \frac{0.64 P_c}{2} = 1.32\, P_c$$

$$\frac{10{,}000}{1.32} = P_c$$

$$P_c = 7575.76 \text{ W} \qquad \textbf{Ans.}$$

The power content of the sidebands is equal to the difference between the total power and the carrier power, *i.e.*,

$$P_{SB} = P_t - P_c$$

The power content of the upper and the lower sidebands is equal, *i.e.*,

$$P_{LSB} + P_{USB} = 10{,}000 - 7575.76 = 242.24$$

$$P_{LSB} = P_{LSB} = \frac{2424.24}{2} = 1212.12 \text{ W}$$

Thus,

$$P_c = 7575.76 \text{ W}$$

$$P_{LSB} = P_{USB} = 1212.12 \text{ W} \qquad \textbf{Ans.}$$

Example 3.26. Determine the modulation index and per cent modulation of the signal shown in figure 3.51.

Solution: Given that an AM signal which is shown in figure 3.51.

We have to find m_a and M

Using the equation relating maximum peak-to-peak amplitude and minimum peak-to-peak amplitude to modulation index, we have

$$m = \frac{\max p\text{-}p - \min p\text{-}p}{\max p\text{-}p + \min p\text{-}p}$$

From, figure 3.51, we get

$$\max p\text{-}p = 2(80) = 160$$

$$\min p\text{-}p = 2(20) = 40$$

Thus,

$$m_a = \frac{160-40}{160+40} = \frac{120}{200} = 0.6$$

$$m_a = 0.6$$

Hence,

$$M = m_a \times 100 = 0.6 \times 100$$

$$M = 60\% \qquad \textbf{Ans.}$$

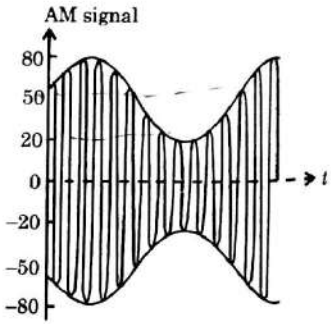


Fig. 3.51.

Example 3.27. Find the modulation index and the per cent modulation of the signal shown in figure 3.52.

Solution: Given that AM signal as shown in figure 3.52.

We have to find m_a and M

Using

$$m_a = \frac{\max p\text{-}p - \min p\text{-}p}{\max p\text{-}p + \min p\text{-}p}$$

and values from figure 3.52, we get

$$\max p\text{-}p = 2(50) = 100$$

$$\min p\text{-}p = 2(15) = 30$$

Thus,

$$m_a = \frac{100-30}{100+30} = \frac{70}{130} = 0.538$$

Hence,

$$M = m_a \times 100 = 53.8\% \qquad \textbf{Ans.}$$

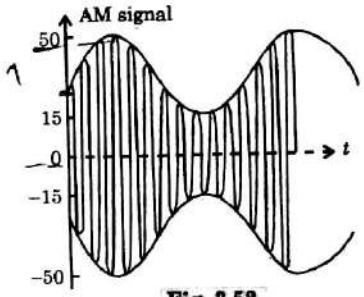


Fig. 3.52.

Example 3.28. Verify that the message signal $m(t)$ is recovered from a modulated DSB signal by first multiplying it by a local sinusoidal carrier and then passing the resultant signal through a low-pass filter (LPF), as shown in figure 3.53 (*a*) in the time domain and (*b*) in the frequency domain.

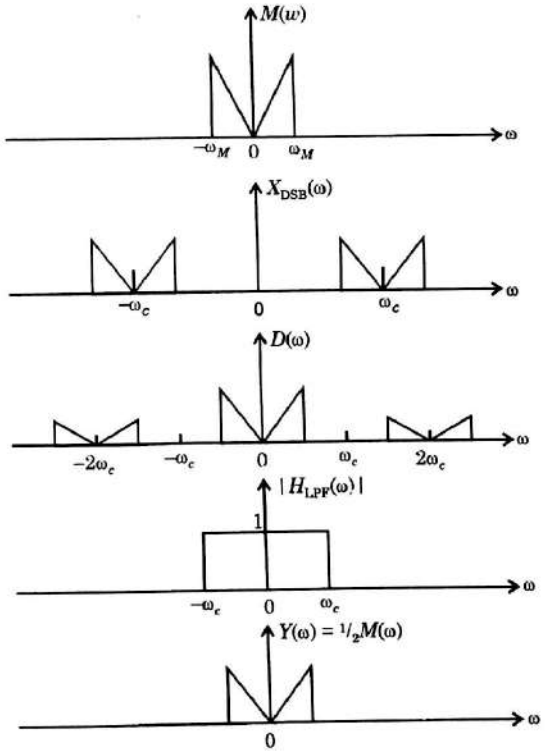


Fig. 3.53. *Demodulation of a DSB signal.*

Solution: The output of the multiplier is given by

$$d(t) = x_{DSB}(t) \cos w_c t = [m(t) \cos w_c t] \cos w_c t$$

or $$d(t) = m(t)\cos^2 w_c t$$

or $$d(t) = \frac{1}{2}m(t) + \frac{1}{2}m(t)\cos 2w_c t$$

After low-pass filtering of signal $d(t)$, we get

$$y(t) = \frac{1}{2}m(t)$$

Thus, by proper amplification (multiplying by 2), we can recover the message signal $m(t)$.

(*b*) Demodulation of signal $x_{DSB}(t)$ in frequency domain has been illustrated in figure 3.53.

Example 3.29. Evaluate the effect of a phase error in the local oscillator on synchronous DSB demodulation.

Solution: Let the phase error of the local oscillator be ϕ. Then the local carrier is expressed as $\cos(\omega_c t + \phi)$. Now, we have

$$x_{DSB}(t) = m(t)\cos w_c t$$

and $$d(t) = [m(t)\cos\omega_c t]\cos(\omega_c t + \phi)$$

$$= \frac{1}{2}m(t)[\cos\phi + \cos(2\omega_c t + \phi)]$$

or $$d(t) = \frac{1}{2}m(t)\cos\phi + \frac{1}{2}m(t)\cos(2\omega_c t + \phi)$$

The second term on the right-hand side is filtered out by the low-pass filter (LPF), and thus we obtain

$$y(t) = \frac{1}{2}m(t)\cos\phi$$

This output is proportional to $m(t)$ when ϕ is constant. The output is completely lost when $\phi = \pm\pi/2$. Thus, the phase error in the local carrier cause attenuation of the output signal without any distortion as long as ϕ is constant and not equal to $\pm\pi/2$. If the phase error ϕ varies randomly with time, then the output also will vary randomly and is undersirable.

Example 3.30. A given AM broadcast station transmits a total power of 50 kW when the carrier is modulated by a sinusoidal signal with a modulation index of 0.707. Calculate:

(*i*) the carrier power

(*ii*) the transmission efficiency, and

(*iii*) the peak amplitude of the carrier assuming the antenna to be represented by a (50 + *j*0) Ω load. *(GATE Examination, 1998)*

Solution: (*i*) The total power transmitted by the AM broadcast station is given by

$$P_c(1 + P_x) = 50 \text{ kW}$$

where P_c is the carrier power and

$$P_x = \frac{1}{2}\times(0.707)^2 = 0.25$$

Thus, we write

$$P_c(1 + 0.25) = 50 \text{ kW}$$

or $$P_c = \frac{50}{1.25} = 40 \text{ kW} \quad \textbf{Ans.}$$

(*ii*) We know that transmission efficiency, η, is given by

$$\eta = \frac{P_c \cdot P_x}{P_c + P_c P_x} = \frac{P_x}{1 + P_x}$$

or $$\eta = \frac{0.25}{1.25} = 0.2 \text{ or } 20\%$$

(*iii*) Also, carrier power, $$P_c = \frac{A_c^2}{2\times 50} = 40\times 10^3 \text{ W}$$

Therefore, peak carrier amplitude,

$$A_c = \sqrt{2\times 50\times 40\times 10^3} = 2000 \text{ Volts or } 2 \text{ kV} \quad \textbf{Ans.}$$

Example 3.31. Evaluate the effect of a small frequency error in the local oscillator on synchronous DSB demodulation.

Solution: Let the frequency error of the local oscillator be $\Delta\omega$. The local carrier is then expressed as $\cos(\omega_c + \Delta\omega)t$.

Then, we have

$$d(t) = m(t)\cos\omega_c t\ \cos(\omega_c + \Delta\omega)t$$

or $$d(t) = \frac{1}{2}m(t)\cos(\Delta\omega)t + \frac{1}{2}m(t)\cos 2\omega_c t$$

Thus, $$y(t) = \frac{1}{2}m(t)\cos(\Delta\omega)t$$

Note: *The output is the signal m (t) multiplied by a low-frequency sinsuoid. This is a* ***"beating"*** *effect and is a very undersirable distortion.*

Example 3.32. Sketch the ordinary AM signal for a single-tone modulation with modulation indices of $m_a = 0.5$ and $m_a = 1$.

Solution: For a single-tone modulation, we have

$$m(t) = a_m\cos\omega_m t$$

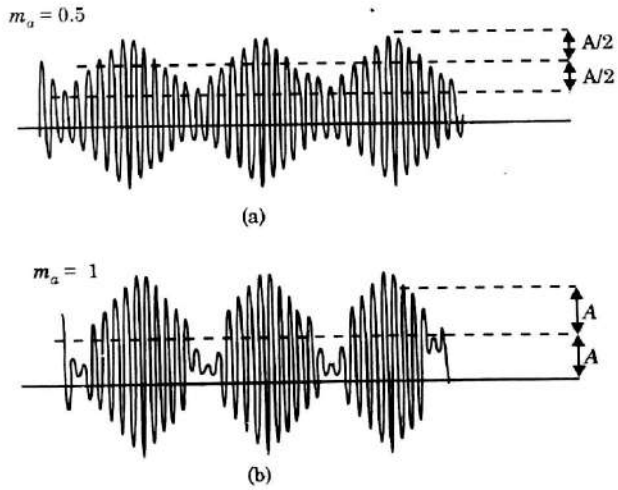


Fig. 3.54.

Then, the modulation index is

$$m_a = \frac{|\max\{m(t)\}|}{A} = \frac{a_m}{A}$$

Hence, $$m(t) = a_m\cos\omega_m t = m_a A\cos\omega_m t$$

Also, $$x_{AM}(t) = [A + m(t)]\cos\omega_c t = A[1 + m_a\cos\omega_m t]\cos\omega_c t$$

Figure 3.54(*a*) and (*b*) show the ordinary AM signal corresponding to $m_a = 0.5$ and $m_a = 1$, respectively.

Example 3.33. The efficiency η of ordinary AM is defined as the per centage of the total power carried by the sidebands, that is,

$$\eta = \frac{P_s}{P_t} \times 100\%$$

where P_s is the power carried by the sidebands and P_t is the total power of the AM signal.
(*i*) Find η for $m_a = 0.5$ (50 per cent modulation).
(*ii*) Show that for a single-tone AM, η_{max} is 33.3 percent at $m_a = 1$.

Solution: A single-tone AM signal can be expressed as under:

$$x_{AM}(t) = A \cos \omega_c t + m_a A \cos \omega_m t \cos \omega_c t$$

$$= A \cos \omega_c t + \frac{1}{2} m_a A \cos (\omega_c - \omega_m)t + \frac{1}{2} m_a A \cos (\omega_c + \omega_m)t$$

$$P_c = \text{carrier power} = \frac{1}{2} A^2$$

$$P_s = \text{sideband power}$$

$$P_s = \frac{1}{2}\left[\left(\frac{1}{2} m_a A\right)^2 + \left(\frac{1}{2} m_a A\right)^2\right] = \frac{1}{4} m_a^2 A^2 \qquad ...(i)$$

The total power P_t is given by

$$P_t = P_c + P_s = \frac{1}{2}A^2 + \frac{1}{4} m_a^2 A^2$$

$$P_t = \frac{1}{2}\left(1 + \frac{1}{2} m_a^2\right) A^2 \qquad ...(ii)$$

Thus $$\eta = \frac{P_s}{P_t} \times 100\% \qquad ...(iii)$$

Substituting values of P_s and P_t in equation (*iii*), we get

or $$\eta = \frac{\frac{1}{4} m_a^2 A^2}{\left(\frac{1}{2} + \frac{1}{4} m_a^2\right) A^2} \times 100\% = \frac{m_a^2}{2 + m_a^2} 100\%$$

with the condition that $m_a \le 1$.

(*i*) For $m_a = 0.5$, we have

$$\eta = \frac{(0.5)^2}{2 + (0.5)^2} \times 100\% = 11.1\%$$

(*ii*) Since $m_a \le 1$, it can be observed that η_{max} occurs at $m_a = 1$ and is given by

$$\eta = \frac{1}{3} \times 100\% = 33.3\% \qquad \textbf{Ans.}$$

Example 3.34. Show that a synchronous demodulator shown in figure 3.55(*a*) can demodulate an AM signal $x_{AM}(t) = [A + m(t)] \cos \omega_c t$ regardless of the value of A.

From figure 3.55(*a*), we have

$$d(t) = x_{AM}(t) \cos \omega_c t$$

But given that

$$x_{AM}(t) = [A + m(t)] \cos \omega_c t$$

Therefore, we write

$$d(t) = [A + m_a(t)] \cos^2 \omega_c t$$

Applying $\cos^2 \theta = \frac{1}{2}[1 + \cos 2\theta]$

$$d(t) = \frac{1}{2}[A + m(t)] + \frac{1}{2}[A + m(t)] \cos^2 \omega_c t$$

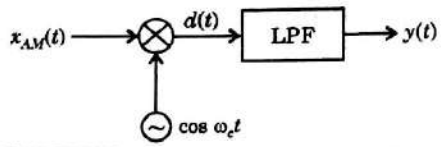


Fig. 3.55. *(a) Synchronous detector for Example 3.34.*

Now, this signal is passed through a low-pass filter (LPF).

Hence, after low-pass filtering (LPF), we obtain

$$y(t) = \frac{1}{2}[A + m(t)] = \frac{1}{2} m(t) + \frac{1}{2} A$$

A blocking capacitor will suppress the direct-current (dc) term $\frac{1}{2}A$, *yielding the output* $\frac{1}{2} m(t)$.

Hence Proved

Example 3.35. Show that an AM signal with large carrier can be demodulated by squaring it and then passing the resulting signal through a low-pass filter (LPF), as shown in figure 3.55(*b*). This type of detector is known as a square-law detector.

We know that AM signal $x_{AM}(t)$ is given by

$$x_{AM}(t) = [A + m(t)] \cos \omega_c t$$

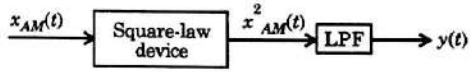


Fig. 3.55. *(b) A square-law detector.*

Squaring the above signal, we get

$$x^2_{AM}(t) = [A + m(t)]^2 \cos^2 \omega_c t$$

Simplifying, we get

$$x^2_{AM}(t) = \frac{1}{2}[A^2 + 2.A.m(t) + m^2(t)] (1 + \cos 2\omega_c t)$$

The low-pass filter output $y(t)$ will be

$$y(t) = \frac{A^2}{2}\left\{1 + 2\frac{m(t)}{A} + \left[\frac{m(t)}{A}\right]^2\right\}$$

Now, with a large carrier, the term $[m(t)/A]^2$ can be neglected and we shall have

$$y(t) = \frac{A^2}{2} + Am(t)$$

Further, a blocking capacitor will suppress the d.c. term. $A^2/2$, yielding the output $Am(t)$.

Hence Proved

Example 3.36. The input to an envelope detector is a single-tone AM signal $x_{AM}(t) = A(1 + m_a \cos \omega_m t) \cos \omega_c t$, where m_a is a constant, $0 < m_a < 1$, and $\omega_c >> \omega_m$.

(*i*) Show that if the detector output is to follow the envelope of $x_{AM}(t)$, it is required that at any time t_0

$$\frac{1}{RC} \ge \omega_m \left(\frac{m_a \sin \omega_m t_0}{1 + m_a \cos \omega_m t_0}\right) \qquad ...(i)$$

(*ii*) Show that if the detector output is to follow the envelope at all times, it is required that

$$RC \le \frac{1}{\omega_m} \frac{\sqrt{1 - m_a^2}}{m_a} \qquad ...(ii)$$

Solution: (*i*) Figure 3.56 shows the envelope of $x_{AM}(t)$ and the output of the detector (*i.e.*, the voltage across the capacitor). Let us assume that the capacitor discharges from the peak value $E_0 = A(1+m_a\cos\omega_m t_0)$ at $t_0 = 0$. Then the voltage $v_c(t)$ across the capacitor can be given by

$$v_c(t) = E_0 e^{-t/(RC)} \qquad ...(iii)$$

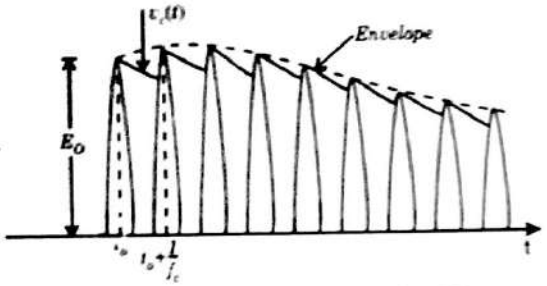


Fig. 3.56. *Waveform for example 3.36.*

The interval between two successive carrier peaks is $1/f_c = 2\pi/\omega_c$ and $RC >> 1/\omega_c$. This means that the time constant RC is much larger than the interval between two successive carrier peaks.

Therefore, $v_c(t)$ can be approximated as

$$v_c(t) \cong E_0\left(1-\frac{t}{RC}\right) \qquad ...(iv)$$

Thus, if the voltage $v_c(t)$ is to follow the envelope of $x_{AM}(t)$, it is required that at any time t_0

$$(1+m_a\cos\omega_m t_0)\left(1-\frac{1}{RCf_c}\right) \le 1+m_a\cos\omega_m\left(t_0+\frac{1}{f_c}\right)$$

Now if $\omega_m << \omega_c$, then, we have

$$1+m_a\cos\omega_m\left(t_0+\frac{1}{f_c}\right) = 1+m_a\cos\left(\omega_m t_0+\frac{\omega_m}{f_c}\right)$$

$$= 1+m_a\cos\omega_m t_0\frac{\omega_m}{f_c} - m_a\sin\omega_m t_0\sin\frac{\omega_m}{f_c}$$

or $$1+m_a\cos\omega_m\left(t_0+\frac{1}{f_c}\right) \cong 1+m_a\cos\omega_m t_0 - m_a\frac{\omega_m}{f_c}\sin\omega_m t_0$$

Hence $$(1+m_a\cos\omega_m t_0)\left(\frac{1}{RCf_c}\right) \ge \frac{m_a\omega_m}{f_c}\sin\omega_m t_0$$

or $$\frac{1}{RC} \ge \omega_m\left(\frac{m_a\sin\omega_m t_0}{1+m_a\cos\omega_m t_0}\right)$$ **Hence Proved**

(*ii*) Now, rewriting equation (*i*), we have

$$\frac{1}{RC}+\frac{m_a}{RC}\cos\omega_m t_0 \ge m_a\omega_m\sin\omega_m t_0$$

or $$m_a\left(\omega_m\sin\omega_m t_0-\frac{1}{RC}\cos\omega_m t_0\right) \le \frac{1}{RC}$$

or $$m_a\sqrt{\omega^2{}_m+\left(\frac{1}{RC}\right)^2}\sin\left(\omega_m t_0-\tan^{-1}\frac{1}{\omega_m RC}\right) \le \frac{1}{RC}$$

Since this inequality must hold for every t_0, we must have

$$m_a\sqrt{\omega^2{}_m+\left(\frac{1}{RC}\right)^2} \le \frac{1}{RC}$$

or $$m_a^2\left[\omega^2{}_m+\left(\frac{1}{RC}\right)^2\right] \le \left(\frac{1}{RC}\right)^2$$

From which we obtain

$$RC \le \frac{1}{\omega_m}\frac{\sqrt{1-m_a^2}}{m_a}$$ **Hence Proved**

Example 3.37. Using the single-tone modulating signal $\cos\omega_m t$, verify that the output of the SSB generator is indeed an SSB signal, and show that an upper-sideband (USB) or a lower-sideband (LSB) signal results from subraction or addition at the summation junction.

Solution: Referring to figure 3.57, we have

$$m(t) = \cos\omega_m t$$

$$\cos\left(\omega_c t-\frac{\pi}{2}\right) = \sin\omega_c t$$

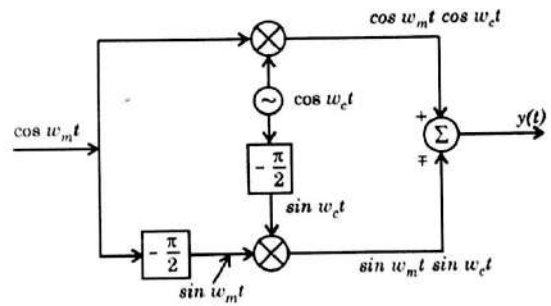


Fig. 3.57. *Block diagram for example 3.37.*

Therefore, $$\hat{m}(t) = \cos\left(\omega_m t-\frac{\pi}{2}\right) = \sin\omega_m t$$

Hence, we have $$y(t) = \cos\omega_m t\cos\omega_c t \mp \sin\omega_m t\sin\omega_c t = \cos(\omega_c \pm \omega_m)t$$

Thus, with subtraction, we have

$$y(t) = x_{USB}(t) = \cos(\omega_c+\omega_m)t$$

and with addition, we shall have

$$y(t) = x_{LSB}(t) = \cos(\omega_c-\omega_m)t$$ **Hence Proved**

Example 3.38. Show that if the output of a phase-shift modulator is an SSB signal, (*i*) the difference of the signals at the summing junction produces the upper-sideband SSB signal and

(*ii*) the sum produces the lower-sideband SSB signal. That is,

$$x_c(t) = x_{USB}(t) = m(t)\cos w_c t - \hat{m}(t)\sin w_c t \qquad ...(i)$$

is an upper-sideband SSB signal, and

$$x_c(t) = x_{LSB}(t) = m(t)\cos w_c t + \hat{m}(t)\sin w_c t \qquad ...(ii)$$

is a lower-sideband SSB signal.

Solution: (*i*) Let $m(t) \leftrightarrow M(w)$

and $\hat{m}(t) \leftrightarrow \hat{M}(w)$

Now, let us apply the modulation theorem or the frequency-shifting property of Fourier transform, we have

$$m(t)\cos w_c t \leftrightarrow \frac{1}{2}M(w-w_c) + \frac{1}{2}M(w+w_c)$$

and
$$\hat{m}(t)\sin w_c t \leftrightarrow \frac{1}{2j}\hat{M}(w-w_c) - \frac{1}{2j}\hat{M}(w+w_c)$$

Taking the Fourier transform of equation (*i*), we have

$$X_c(w) = \frac{1}{2}M(w-w_c) + \frac{1}{2}M(w+w) - \left[\frac{1}{2j}\hat{M}(w-w_c) - \frac{1}{2j}\hat{M}(w+w_c)\right]$$

we also have

$$\hat{M}(w-w_c) = -j\,sgn\,(w-w_c)\,M(w-w_c)$$

and
$$\hat{M}(w+w_c) = -j\,sgn\,(w+w_c)\,M(w+w_c)$$

Thus
$$X_c(w) = \frac{1}{2}M(w-w_c) + \frac{1}{2}M(w+w_c) - \left[-\frac{1}{2}sgn(w-w_c)M(w-w_c) + \frac{1}{2}sgn(w+w_c)M(w+w_c)\right] \quad ...(iii)$$

or
$$X_c(w) = \frac{1}{2}M(w-w_c)[1+sgn(w-w_c)] + \frac{1}{2}M(w+w_c)[1-sgn(w+w_c)]$$

Since
$$1+sgn\,(w-w_c) = \begin{cases} 2 & \text{for } w > w_c \\ 0 & \text{for } w < w_c \end{cases}$$

and
$$1-sgn\,(w-w_c) = \begin{cases} 2 & \text{for } w < -w_c \\ 0 & \text{for } w > -w_c \end{cases}$$

we have
$$X_c(w) = \begin{cases} 0 & \text{for } |w| < w_c \\ M(w+w_c) & \text{for } w < -w_c \\ M(w-w_c) & \text{for } w > w_c \end{cases} \quad ...(iv)$$

which has been sketched in figure 3.58(*b*). We see that $x_c(t)$ is an upper sideband SSB signal.

Hence Proved

(*ii*) In a similar manner, taking the Fourier transform of equation (*ii*), we have

$$X_c(w) = \frac{1}{2}M(w-w_c)[1-sgn\,(w-w_c)] + \frac{1}{2}M(w+w_c)[1+sgn\,(w+w_c)] \quad ...(v)$$

Now, since
$$1-sgn\,(w-w_c) = \begin{cases} 2 & \text{for } w < w_c \\ 0 & \text{for } w > w_c \end{cases}$$

and
$$1+sgn(w+w_c) = \begin{cases} 2 & \text{for } w > -w_c \\ 0 & \text{for } w < -w_c \end{cases}$$

we have
$$X_c(w) = \begin{cases} 0 & \text{for } |w| > w_c \\ M(w-w_c) & \text{for } w < w_c \\ M(w+w_c) & \text{for } w > -w_c \end{cases} \quad ...(vi)$$

which has been sketched in figure 3.58(*c*). We see that $x_c(t)$ is a lower-sideband SSB signal.

Hence Proved

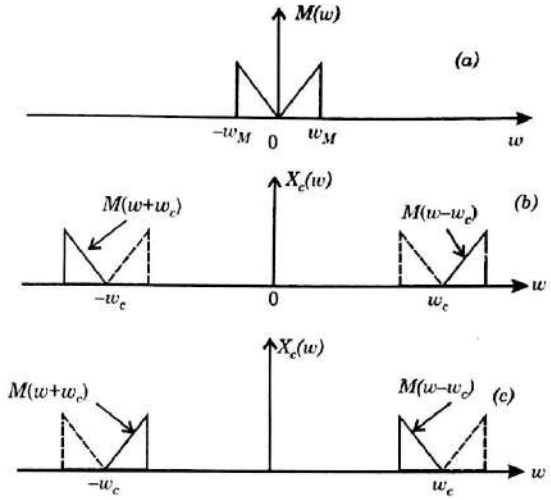


Fig. 3.58. *Waveforms for example 3.38.*

Example 3.39. Show that an SSB signal can be demodulated by the synchronous detector of figure 3.59 by sketching the spectrum of the signal at each point and (*ii*) by the time-domain expression of the signals at each point.

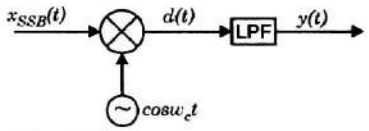


Fig. 3.59. *Synchronous detector for example 3.39.*

Solution: (*i*) Let $M(w)$, the spectrum of the message $m(t)$, be as shown in figure 3.60(*a*). Also let us assume that $x_{SSB}(t)$ is a lower-sideband SSB signal and its spectrum is $X_{SSB}(\omega)$, as shown in figure 3.60(*b*). Multiplication by $\cos \omega_c t$ shifts the spectrum $X_{SSB}(\omega)$ to $\pm\omega_c$ and we obtain $D(\omega)$, the spectrum of $d(t)$ as shown in figure 3.60 (*d*) After low-pass filtering, we obtain $Y(w) = \frac{1}{2}M(w)$, the spectrum of $y(t)$ as shown in figure 3.60(*c*).

Thus, we obtain $y(t) = \frac{1}{2}m(t)$ which is proportional to message signal $m(t)$. **Hence Proved.**

(*ii*) We know that $x_{SSB}(t)$ can be expressed as under:

$$x_{SSB}(t) = m(t)\cos w_c t \mp \hat{m}(t)\sin w_c t \quad ...(i)$$

Thus,
$$d(t) = x_{SSB}(t)\cos w_c t$$

Using equation (*i*), we get

$$d(t) = m(t)\cos^2 w_c t \mp \hat{m}(t)\sin w_c t\cos w_c t$$

$$= \frac{1}{2}m(t)(1+\cos 2w_c t) \mp \frac{1}{2}\hat{m}(t)\sin 2w_c t$$

or
$$d(t) = \frac{1}{2}m(t) + \frac{1}{2}m(t)\cos 2w_c t \mp \frac{1}{2}\hat{m}(t)\sin 2w_c t$$

Hence, after low-pass filtering, we obtain

$$y(t) = \frac{1}{2}m(t)$$

Hence Proved.

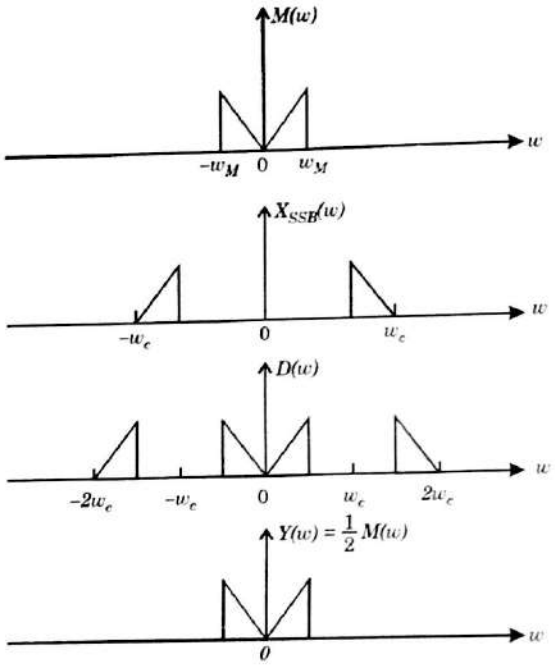


Fig. 3.60. *Waveforms for example 3.39.*

Example 3.40. A carrier signal of 1.0 Volt amplitude and a sinusoidal modulating signal of 0.5 V, put in series, are applied to a square law modulator of characteristics,

$$i'_o = 10 + kV_i + k'V_i^2 \text{ mA}$$

where V_i is input in volts,

$$k = 2 \text{ mA/V and } k' = 0.2 \text{ mA/V}^2$$

considering only the frequency components of the AM signal corresponding to the carrier frequency, find the depth of modulation in the resulting AM signal. *(GATE Examination-1999)*

Solution: We have

$$V_i(t) = \cos(w_c t) + 0.5\cos(w_m t)$$

and $$i_o = 10 + kV_i + k'V_i^2$$

or $$i_o = 10 + 2\times10^{-3}[\cos(w_c t) + 0.5\cos(w_m t)] + 0.2\times10^{-3}[\cos(w_c t) + 0.5\cos(w_m t)]^2$$

Further, considering only the carrier terms, we have

$$i_o = 2\times10^{-3}\cos(w_c t) + \frac{0.2\times10^{-3}}{0.5}\times 0.5\cos(w_c t)\cos(w_m t)$$

Now, modulation depth will be

$$m = \frac{0.2\times10^{-3}}{0.5} = 0.4\times10^{-3} \qquad \textbf{Ans.}$$

Example 3.41. Show that the system shown in figure 3.61 can be used to demodulate an SSB signal.

Solution: We know that the upper-sideband SSB signal $x_c(t)$ is given by

$$x_c(t) = m(t)\cos w_c t - \hat{m}(t)\sin w_c t$$

we also have, $$\hat{x}_c(t) = m(t)\sin w_c t + \hat{m}(t)\cos w_c t$$

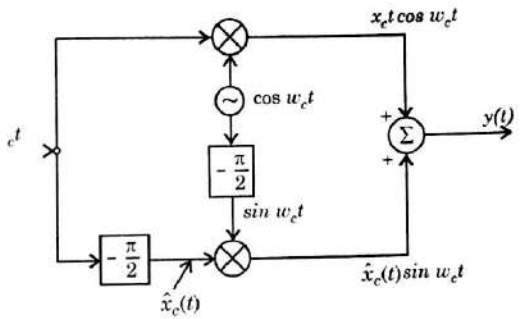


Fig. 3.61. *Phase-shift SSB demodulator for example 3.41.*

From figure 3.61, we have

$$y(t) = x_c(t)\cos w_c t + \hat{x}_c(t)\sin w_c t$$

or $$y(t) = m(t)(\cos^2 w_c t + \sin^2 w_c t) = m(t)$$

In a similar manner, the lower-sideband SSB signal $x_c(t)$ is given by

$$x_c(t) = m(t)\cos w_c t + \hat{m}(t)\sin w_c t$$

and $$\hat{x}_c(t) = m(t)\sin w_c t - \hat{m}(t)\cos w_c t$$

Again, from figure 3.63, we obtain

$$y(t) = x_c(t)\cos w_c t + \hat{x}_c(t)\sin w_c t$$
$$= m(t)(\cos^2 w_c t + \sin^2 w_c t) = m(t)$$

Hence Proved

Example 3.42. Determine the per centage power saving when the carrier wave and one of the sidebands are suppressed in an AM wave modulated to a depth of (*i*) 100% and (*ii*) 50%.

Solution: (*i*) Total power in *AM* wave is given by

$$P_t = P_c\left(1 + \frac{m^2}{2}\right) \qquad ...(i)$$

Given that modulation index $m = 100\% = 1.0$

Equation (*i*) can be written as

$$P_t = P_c\left(1 + \frac{(1)^2}{2}\right) = 1.5P_c \qquad ...(ii)$$

The power in one sideband is given by

$$P_{SB} = P_c\frac{m^2}{4} = P_c\frac{(1)^2}{4} = \frac{P_c}{4} = 0.25P_c \qquad ...(iii)$$

$$\text{Power saving} = \frac{P_t - P_{SB}}{P_t} = \frac{1.5P_c - 0.25P_c}{1.5P_c} = \frac{1.25}{1.50} = \frac{5}{6} = 0.833 = 83.3\%$$

(*ii*) Given that $m = 50\% = 0.50$ and $P_t = P_c\left(1 + \frac{m^2}{2}\right)$

Putting the values of m, we have

$$P_t = P_c\left(1 + \frac{(0.5)^2}{2}\right) = 1.125\,P_c \qquad ...(iv)$$

$$P_{SB} = P_c \frac{m^2}{4} = P_c \frac{(0.5)^2}{4} = 0.0625\, P_c \qquad ...(v)$$

Hence, power saving $= \dfrac{P_t - P_{SB}}{P_t} = \dfrac{1.125P_c - 0.0625P_c}{1.125P_c} = 0.944 = 94.4\%$ **Ans.**

Example 3.43. An AM signal $s(t) = [1 + 0.2 \cos\{2\pi(f_m/3)t\}] \cos 2\pi f_c t$ is detected using a square-law demodulator. This square-law modulator has the characteristic $v_0(t) = [s(t) + 2]^2$. The output is then filtered by an ideal low-pass filter (LPF) with a cut-off frequency at f_m, Hertz. Obtain the frequency spectrum of the demodulated signal. Also plot the frequency spectrum in the frequency range $-f_m \le f \le f_m$.

Solution: The output of the demodulator is given as

$$v_0(t) = [s(t) + 2]^2 = s^2(t) + 4 + 4s(t) \qquad ...(i)$$

here $s(t)$ is input to the modulator.

Given an *AM* modulated signal as under :

$$s(t) = \left[1 + 0.2\cos\left\{2\pi\left(\frac{f_m}{3}\right)t\right\}\right]\cos(2\pi f_c t) \qquad ...(ii)$$

Substituting equation (*ii*) in equation (*i*), we get

$$v_0(t) = s^2(t) + 4 + 4s(t)$$

or $$v_0(t) = \left\{\left[1 + 0.2\cos\left\{2\pi\left(\frac{f_m}{3}\right)t\right\}\right]\cos(2\pi f_c t)\right\}^2 + 4\left[1 + 0.2\cos\left\{2\pi\left(\frac{f_m}{3}\right)t\right\}\right]\cos(2\pi f_c t)$$

or $$v_0(t) = \left[1 + 0.2\cos\left\{2\pi\left(\frac{f_m}{3}\right)t\right\}\right]^2\cos^2(2\pi f_c t) + 4\left[1 + 0.2\cos\left\{2\pi\left(\frac{f_m}{3}\right)t\right\}\right]\cos(2\pi f_c t) \qquad ...(iii)$$

Simplifying last equation, we get

$$v_0(t) = 4.51 + 0.2\cos\left[2\pi\left(\frac{f_m}{3}\right)t\right] + 0.01\cos\left[2\pi\left(\frac{2f_m}{3}\right)t\right] + 1.02\cos(4\pi f_c t)$$

$$+0.2\cos\left[2\pi\left(2f_c + \frac{f_m}{3}\right)t\right]\left[2\pi\left(2f_c - \frac{f_m}{3}\right)t\right] + 0.01\cos\left[2\pi\left(2f_c + \frac{2f_m}{3}\right)t\right]$$

$$+0.01\cos\left[2\pi\left(2f_c - \frac{2f_m}{3}\right)t\right] + 4\cos(2\pi f_c t) + 0.4\cos\left[2\pi\left(f_c + \frac{f_m}{3}\right)t\right] + 0.4\cos\left[2\pi\left(f_c - \frac{f_m}{3}\right)t\right] \qquad ...(iv)$$

Also, the frequency spectrum of the signal $v_0(t)$ can be obtained by taking Fourier transform of both sides of equation (*iv*), as under:

$$V_o(f) = 4.51\,\delta(f) + 0.1\left[\delta\left(f - \frac{f_m}{3}\right) + \delta\left(f + \frac{f_m}{3}\right)\right] + 0.005\left[\delta\left(f - \frac{2f_m}{3}\right) + \delta\left(f + \frac{2f_m}{3}\right)\right]$$

$$+ 0.51\,[\delta(f - 2f_c) + \delta(f + 2f_c)] + 0.01\left[\delta\left(f - 2f_c - \frac{f_m}{3}\right) + \delta\left(f + 2f_c + \frac{f_m}{3}\right)\right] + 0.005$$

$$+ 0.01\left[\delta\left(f - 2f_c - \frac{2f_m}{3}\right) + \delta\left(f + 2f_c + \frac{2f_m}{3}\right)\right] + 0.005\left[\delta\left(f - 2f_c + \frac{2f_m}{3}\right) + \delta\left(f + 2f_c - \frac{2f_m}{3}\right)\right]$$

$$+ 2[\delta(f - f_c) + \delta(f + f_c)] + 0.2\left[\delta\left(f - f_c - \frac{f_m}{3}\right) + \delta\left(f + f_c + \frac{f_m}{3}\right)\right] + 0.2\left[\delta\left(f - f_c + \frac{f_m}{3}\right) + \delta\left(f + f_c - \frac{f_m}{3}\right)\right]$$

The spectrum of the low-pass filter LPF output has been shown in figure 3.62.

The low-pass filter with a cut-off frequency equal to f_m, will only allow the d.c. component and the components having frequencies $f_m/3$ and $2f_m/3$ and reject all other frequency components present in the output $V_o(f)$ of the square-law modulator.

Hence, the spectrum of the filter output is given by

$$S_{LPF}(f) = 4.51\,\delta(f) + 0.1\left[\delta\left(f - \frac{f_m}{3}\right) + \delta\left(f + \frac{f_m}{3}\right)\right]$$

$$+ 0.005\left[\delta\left(f - \frac{2f_m}{3}\right) + \delta\left(f + \frac{2f_m}{3}\right)\right] \quad \textbf{Ans.}$$

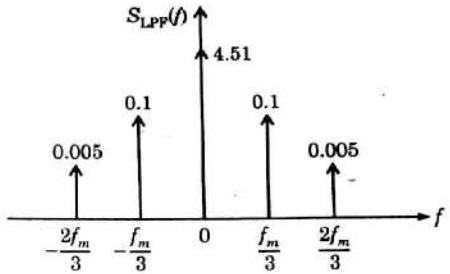


Fig. 3.62. *Showing spectrum of the LPF output.*

Example 3.44. The amplitude modulated (AM) waveform $s(t) = [1 + m \cos(2\pi f_m t)] \cos(2\pi f_c t)$ is to be demodulated by an envelope detector. Find the maximum permissible value of the modulation index m for which such demodulation is possible. Also sketch the output of the envelope detector for $m = 2$.

Solution: The envelope of the *AM* signal $s(t) = [1 + m\cos(2\pi f_m t)]\cos(2\pi f_c t)$ corresponds to the modulating, *i.e.*, message signal so long as

$$|m \cos(2\pi f_m t)| < 1 \qquad ...(i)$$

This means that the per centage of modulation is less than 100%. Hence, it is possible to recover the message signal $x(t)$ from *AM* signal $s(t)$ by an envelope detection so long as per centage modulation is less than 100%.

For modulation index, $m = 2$, the *AM* wave or signal will appear like one shown in figure 3.63(*a*). The corresponding demodulated signal is shown in figure 3.63(*b*) Here, it can be noted that the demodulator cannot detect the modulating signal $x(t)$ properly since it can follow only the positive cycles.

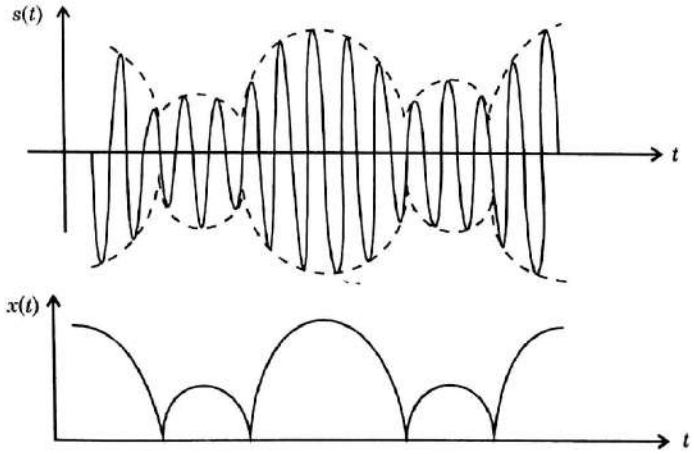


Fig. 3.63. *(a) AM signal for m = 2, (b) Corresponding demodulated signal.*

Example 3.45. A signal $v(t) = [1 + m(t)] \cos(w_c t)$ is detected using a square law detector, having the characteristic $v_o = v^2$. If the Fourier transform of $m(t)$ is constant, M_o, extending from $-f_m$ to $+f_m$, sketch the Fourier transform of $v_o(t)$ in the frequency range $-f_m < f < f_m$.

(GATE Examination-1998)

Solution: Given

$$v(t) = [1 + m(t)] \cos w_c t$$

$$v_o = v^2 = [1 + m^2(t) + 2m(t)] \cos^2 w_c t = \frac{1}{2}[1+m^2+2m][1+\cos 2 w_c t]$$

$$v_o = \frac{1}{2}[1+m^2+2m+\cos 2w_c t + m^2 \cos 2w_c t + 2m \cos 2w_c t]$$

or $$v_o(t) = \frac{1}{2}[1+m^2+2m+\cos 2 w_c t + m^2 \cos 2 w_c t + 2m \cos 2w_c t]$$

Also, $$V_o(f) = \frac{1}{2}[\delta(f) + M(f) \otimes M(f) + 2M(f) + \frac{\partial(f-2f_c)}{2} + \frac{\partial(f+2f_c)}{2}$$

$$\frac{M_c(f-2f_c)}{2} + \frac{M_c^2(f+2f_c)}{2} + M(f-2f_c) + M(f+2f_c)$$

where $M_c(f) = M(f) \otimes M(f)$

In the frequency range, $-f_m \le f \le f_m$, we have

$$V_m(f) = \frac{1}{2}[\delta(f) + M(f) \otimes M(f) + 2M(f)]$$

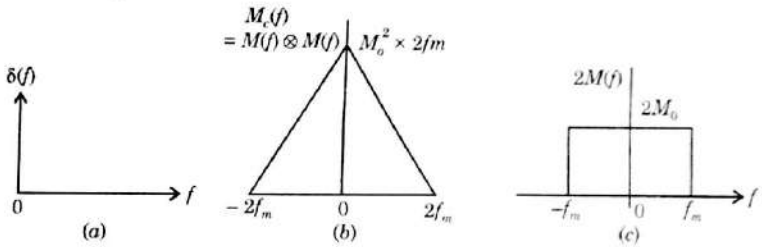


Fig. 3.64.

Hence $V_m(f)$ will be as shown in figure 3.65.

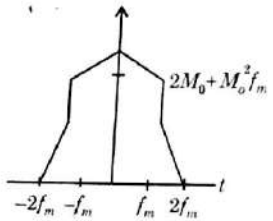


Fig. 3.65.

Exmaple 3.46. The modulating signal $x(t)$ is recovered from the DSB-SC signal $s(t) = x(t) \cos (\omega_c t)$ by multiplying $s(t)$ by a locally generated carrier $c'(t) = \cos (\omega_c t + \phi)$, where $\omega_c = 2\pi f_c$, is the angular carrier frequency. The product of $s(t)$ $c'(t)$ is passed through a low-pass filter (LPF) which rejects the double frequency signal. Determine the maximum allowable value for the phase angle if the recovered signal is to be 95% of the maximum possible output. If the modulating signal is band limited to 10 kHz, determine the minimum value of carrier frequency f_c for which $x(t)$ can be recovered by filtering.

Solution: The product signal is given by

$$v(t) = s(t) \cos (2\pi f_c t + \phi)$$

or $$v(t) = x(t) \cos (2\pi f_c t) \cos (2\pi f_c t + \phi)$$

or $$v(t) = x(t)\left[\frac{\cos(4\pi f_c t + \phi) + \cos\phi}{2}\right]$$

or $$v(t) = \frac{x(t)}{2} \cos (4\pi f_c t + \phi) + \frac{x(t)}{2} \cos\phi$$

The low pass filter (LPF) will reject the first term.

The maximum allowable value of phase angle ϕ can be found as under:

$$\cos (\phi_{max}) = \frac{\left[\frac{x(t)}{2} \cos \phi\right]}{\max\left[\frac{x(t)}{2} \cos\phi\right]} = 0.95$$

i.e., $$\phi_{max} = \cos^{-1}(0.95) = 18°.19$$

In order to recover $x(t)$ from $v(t)$ using filter method, it is essential that the lowest frequency contained in the first term of $v(t)$ must be greater than the highest frequency contained in the second term, *i.e.*,

$$2f_c - 10 \text{ kHz} > 10 \text{ kHz}$$

or $$f_c > 10 \text{ kHz}$$

Hence, the minimum value of f_c will be

$$f_c = 10 \text{ kHz} \quad \textbf{Ans.}$$

Example 3.47. (*i*) The modulating signal $x(t)$ is to be recovered from the DSB-SC signal given by $s(t) = x(t) \cos (2\pi f_c t)$. The available local carrier is a periodic signal $p(t)$ with a time period $1/f_c$. Prove that the modulating signal $x(t)$ can be recovered from the product $v(t) = s(t)\, p(t)$ by using appropriate filter.

(*ii*) Show that $x(t)$ can be recovered as well if the periodic waveform has a period n/f_c, where n is an integer.

(*iii*) If modulating signal $x(t)$ is band-limited to the frequency range from 0 to 5 kHz, find the largest value of n which will allow $x(t)$ to recover.

Given f_c = 1 MHz.

Solution : (*i*) The periodic waveform can be expanded in exponential Fourier series as under:

$$p(t) = \sum_{n=-\infty}^{\infty} C_n e^{(j2\pi n f_c t)} \qquad ...(i)$$

where C_n's are complex Fourier coefficient.

Thus, $C_{-n} = C_n^*$

The product $v(t)$ is expressed as

$$v(t) = s(t)\, p(t) \qquad ...(ii)$$

Substituting equation (*i*) and $s(t) = x(t) \cos (2\pi f_c t)$ in equation (*ii*), we get

$$v(t) = p(t)\, v(t) \qquad ...(iii)$$

or $$v(t) = \sum_{n=-\infty}^{\infty} C_n e^{(j2\pi n f_c t)}\, x(t) \cos (2\pi f_c t) = \sum_{n=-\infty}^{\infty} C_n x(t) \cos (2\pi f_c t)\, e^{(j2\pi n f_c t)}$$

or $$v(t) = x(t) \sum_{n=-\infty}^{\infty} \left[C_n \left\{\frac{e^{j2\pi f_c t} + e^{-j2\pi n f_c t}}{2}\right\} e^{(j2\pi n f_c t)}\right]$$

or $$v(t) = \frac{x(t)}{2} \sum_{n=-\infty}^{\infty} C_n e^{[j2\pi(n+1) f_c t]} + \frac{x(t)}{2} \sum_{n=-\infty}^{\infty} C_n e^{[j2\pi(n-1) f_c t]} \qquad ...(iv)$$

The signal $x(t)$ can be recovered from product $v(t)$ by passing it through a low-pass filter (LPF) which rejects double frequency signal. The output of the filter is then given by

$$v_0(t) = \frac{x(t)}{2}(C_1 + C_{-1}) = \frac{x(t)}{2}[2R_e(C_1)]$$

or $$v_0(t) = [R_e(C_1)]\,x(t) \qquad ...(v)$$

(*ii*) Substituting f_c/n in place of f_c in Fourier series expansion of $p(t)$, we obtain the product signal $v(t)$ as

$$v(t) = \frac{x(t)}{2}\sum_{n=-\infty}^{\infty} C_n e^{\left[j2\pi\left(\frac{1}{n}+1\right)f_c t\right]} + \sum_{n=-\infty}^{\infty} C_n e^{\left[j2\pi\left(\frac{1}{n}-1\right)f_c t\right]}$$

The original modulating signal $x(t)$ can be recovered from $v(t)$ only when the largest frequency contained in $x(t)$ is less than f_c/n.

Under this condition, we have

$$v_0(t) = [R_e(C_n)]\,x(t)$$

(*iii*) For the given band-limited modulating signal $x(t)$, highest frequency component contained will be

$$f_m = 5 \text{ kHz}$$

The carrier frequency $f_c = 1$ MHz (given)

The largest value of n can be found from

$$\frac{f_c}{n_{\max}} = f_m$$

i.e., $$n_{max} = \frac{f_c}{f_m} = \frac{1\times10^6 \text{ Hz}}{5\times10^3 \text{ kHz}} = 200 \quad \textbf{Ans.}$$

Example 3.48. (*i*) Prove that the signal

$$s(t) = \sum_{i=1}^{N}[\cos(\omega_c t)\cos(\omega_i t+\phi_i) - \sin(\omega_c t)\sin(\omega_i t+\phi_i)]$$

in an SSB signal ($f_c >> f_N$), where $\omega_c = 2\pi f_c$, carrier angular frequency and $\omega_i = 2\pi f_i$. Identify the sideband.

(*ii*) Obtain an expression for missing sideband.

(*iii*) Obtain an expression for the total DSB-SC signal.

Solution: (*i*) The given signal can be modified using trigonometric expression as

$$s(t) = \sum_{i=1}^{N}[\cos(\omega_c t)\cos(\omega_i t+\phi_i) - \sin(\omega_c t)\sin(\omega_i t+\phi_i)]$$

or $$s(t) = \sum_{i=1}^{N}\left[\frac{1}{2}\cos\{(\omega_c+\omega_i)t+\phi_i\} + \frac{1}{2}\cos\{\omega_c-\omega_i)t-\phi_i\}\right]$$

$$-\sum_{i=1}^{N}\left[\frac{1}{2}\cos\{\omega_c-\omega_i)t-\phi_i\} + \frac{1}{2}\cos\{\omega_c+\omega_i)t+\phi_i\}\right]$$

or $$s(t) = \sum_{i=1}^{N}\cos\{(\omega_c+\omega_i)t+\phi_i\} = s_u(t) \quad \text{[Let us say]} \qquad ...(i)$$

Hence, $s(t)$ contains frequency components higher than the carrier frequency f_c and thus SSB signal with upper sideband (USB).

(*ii*) The lower sideband (LSB) will be expressed as

$$s(t) = \sum_{i=1}^{N}[\cos(\omega_c t)\cos(\omega_i t+\phi_i) + \sin(\omega_c t)\sin(\omega_i t+\phi_i)] \qquad ...(ii)$$

Equation (*ii*) can be modified using trigonometric expression and we obtain

$$s(t) = \sum_{i=1}^{N}\cos\{(\omega_c-\omega_i)t-\phi_i\} = s_L(t) \quad \text{[Let us say]} \qquad ...(iii)$$

(*iii*) The total DSB-SC signal can be written as under:

$$s_{\text{DSB-SC}}(t) = s_U(t) + s_L(t)$$

Substituting equations (*i*) and (*iii*) in equation (*iv*), we get

$$s_{\text{DSB-SC}}(t) = \sum_{i=1}^{N}\cos\{(\omega_c+\omega_i)t+\phi_i\} + \sum_{i=1}^{N}\cos\{(\omega_c-\omega_i)t-\phi_i\}$$

or $$s_{\text{DSB-SC}}(t) = \sum_{i=1}^{N}\cos\{(w_c+w_i)t+\phi_i\} + \cos\{(\omega_c-\omega_i)t-\phi_i)$$

On modifying last expression, we get

$$s_{\text{DSB-SC}}(t) = 2\sum_{i=1}^{N}\cos(\omega_c t)\cos(\omega_i t+\phi_i)$$

or $$s_{\text{DSB-SC}}(t) = \left[2\sum_{i=1}^{N}\cos(\omega_i t+\phi_i)\right]\cos(\omega_c t) \quad \textbf{Ans.}$$

Exmaple 3.49. In SSB-SC signal generation using phase discrimination method, the carrier phase shift network produces a phase shift which differs from $\pi/2$ by a small angle α. Obtain the output waveform and also verify that the output no longer corresponds to SSB-SC signal. The modulating signal $x(t)$ may be considered to be a single tone sinusoidal signal 1.0 cos ($2\pi f_m t$).

Solution: Given that single-tone sinusoidal signal is

$$x(t) = 1.0\cos(2\pi f_m t) \quad \text{and} \quad \hat{x}(t) = 1.0\sin(2\pi f_m t+\alpha) \qquad ...(i)$$

The output of the SSB generating system will be

$$s(t) = x(t)\cos(2\pi f_c t) \pm \hat{x}(t)\sin(2\pi f_c t+\alpha) \qquad ...(ii)$$

Substituting equation (*i*) in equation (*ii*), we get

$$s(t) = 1.0\cos(2\pi f_m t)\cos(2\pi f_c t) \pm 1.0\sin(2\pi f_m t+\alpha)\sin(2\pi f_c t+\alpha)$$

or $$s(t) = \cos(2\pi f_m t)\cos(2\pi f_c t) \pm \sin(2\pi f_m t+\alpha)\sin(2\pi f_c t+\alpha) \qquad ...(iii)$$

Modifying equation (*iii*), we get

$$s(t) = \frac{1}{2}[\cos\{2\pi(f_c+f_m)t\} + \cos\{2\pi(f_c-f_m)t\}]$$

$$\pm\frac{1}{2}[\cos\{2\pi(f_c-f_m)t-\alpha\} - \cos\{2\pi(f_c+f_m)t+\alpha\}]$$

or $$s(t) = \frac{1}{2}[\cos\{2\pi(f_c+f_m)t\} \mp \cos\{2\pi(f_c+f_m)t+\alpha\}]$$

$$+[\cos\{2\pi(f_c+f_m)t\} \pm \cos\{2\pi(f_c-f_m)t-\alpha\}]$$

Hence, we can say that the output waveform is no longer an SSB-SC waveform.

Hence Proved

Example 3.50. A synchronous detection of SSB-SC signal shows phase and frequency discrepancy. Consider the synchronous detection of SSB signal given by

$$s(t) = \sum_{i=1}^{N}[\cos(\omega_c t)\cos(\omega_i t+\phi_i) - \sin(\omega_c t)\sin(\omega_i t+\phi_i)]$$

The signal is multiplied by the locally generated carrier $\cos\omega_c t$ and then low-pass filtered.

(i) Prove that the modulating signal can be completely recovered if the cut-off frequency of the filter is $f_N < f_o < 2f_c$.

(ii) Determine the recovered signal when the multiplying signal is $\cos[\omega_c t + \phi]$.

(iii) Determine the recovered signal when the multiplying signal is $\cos[(\omega_c + \Delta\omega)\, t]$.

Given $\Delta f << f_i$ where $\omega_c = 2\pi f_c$, $\Delta\omega = 2\pi\Delta f$

Solution: (i) The product signal is

$$v(t) = s(t)\cos(\omega_c t) \qquad ...(i)$$

Substituting value of $s(t)$ in equation (i), we get

$$v(t) = \sum_{i=1}^{N}[\cos(\omega_c t)\cos(\omega_i t+\phi_i) - \sin(\omega_c t)\sin(\omega_i t+\phi_i)]\cos(\omega_c t)$$

$$= \sum_{i=1}^{N}\cos[\omega_c + \omega_i)\, t + \phi_i]\cos(\omega_c t) \qquad ...(ii)$$

Modifying equation (ii), we get

$$s(t) = \sum_{i=1}^{N}\left[\frac{1}{2}\cos\{(2\omega_c + \omega_i)\, t + \phi_i\} + \frac{1}{2}\cos(\omega_i t + \phi_i)\right]$$

or $$s(t) = \frac{1}{2}\sum_{i=1}^{N}\cos[(\omega_c + \omega_i)\, t + \phi_i] + \frac{1}{2}\sum_{i=1}^{N}\cos(\omega_i t + \phi_i) \qquad ...(iii)$$

The output waveform of the low-pass filter (LPF) is thus given by

$$v_0(t) = \frac{1}{2}\sum_{i=1}^{N}\cos(\omega_i t + \phi_i) \qquad ...(iv)$$

when the cut-off frequency, f_o of the LPF is in the range $f_N < f_o < 2f_c$.

(ii) when the multiplying signal is $\cos(\omega_c t + \phi)$, the product signal will be given by

$$v(t) = s(t)\cos(\omega_c t + \phi) \qquad ...(v)$$

Substituting the value of $s(t)$ in equation (v), we get

$$v(t) = \sum_{i=1}^{N}[\cos(\omega_c t)\cos(\omega_i t+\phi_i) - \sin(\omega_c t)\sin(\omega_i t+\phi)]\cos(\omega_c t+\phi)$$

or $$v(t) = \sum_{i=1}^{N}\cos[(\omega_c + \omega_i)\,\phi_i]\cos(\omega_c t+\phi) \qquad ...(vi)$$

Modifying equation (vi), we get

$$v(t) = \sum_{i=1}^{N}\left[\frac{1}{2}\cos\{(2\omega_c + \omega_i)\, t + (\phi_i + \phi)\} + \frac{1}{2}\cos\{\omega_i t + (\phi_i - \phi)\}\right]$$

$$v(t) = \frac{1}{2}\sum_{i=1}^{N}\left[\frac{1}{2}\cos\{(2\omega_c + \omega_i)\, t + (\phi_i + \phi)\} + \frac{1}{2}\sum_{i=1}^{N}\cos\{\omega_i t + (\phi_i - \phi)\}\right] \qquad ...(vii)$$

The output of the low-pass filter (LPF) will be

$$v_0(t) = \frac{1}{2}\sum_{i=1}^{N}\cos[\omega_i t + (\phi_i - \phi)] \qquad ...(viii)$$

Therefore, the demodulated signal shows phase discrepancy.

(iii) when we multiply $s(t)$ by $\cos[(\omega_c + \Delta\omega)\, t]$ then the product signal will be expressed as

$$v(t) = s(t)\cos[(\omega_c + \Delta\omega)t] \qquad ...(ix)$$

Substituting the value of $s(t)$ in equation (ix), we get

$$v(t) = s(t)\cos[(\omega_c + \Delta\omega)\, t]$$

or $$v(t) = \sum_{i=1}^{N}[\cos(\omega_c t)\cos(\omega_i t+\phi_i) - \sin(\omega_c t)\sin(\omega_i t+\phi_i)]\cos[(\omega_c + \Delta\omega)t]$$

or $$v(t) = \sum_{i=1}^{N}\cos[(\omega_c + \omega_i)\, t + \phi_i]\cos[(\omega_c + \Delta\omega)\, t] \qquad ...(x)$$

On modifying equation (x), we get

$$v(t) = \sum_{i=1}^{N}\left[\frac{1}{2}\cos\{(2\omega_c + \omega_i + \Delta\omega)\, t + \phi_i\} + \frac{1}{2}\cos\{(\omega_i - \Delta\omega)\, t\right]$$

or $$v(t) = \frac{1}{2}\sum_{i=1}^{N}\cos\{(2\omega_c + \omega_i + \Delta\omega)\, t + \phi_i\} + \frac{1}{2}\sum_{i=1}^{N}\cos\{(\omega_i - \Delta\omega)\, t\}$$

The output of the low-pass filter (LPF) will then be given by

$$v_0(t) = \frac{1}{2}\sum_{i=1}^{N}\cos\{(\omega_i - \Delta\omega)\, t\}$$

Hence, the demodulated signal shows frequency error. **Ans.**

Example 3.51. A received single-tone sinusoidally modulated SSB-SC signal $\cos\{(\omega_c + \omega_m)t\}$ has a normalised power of 0.5 volt². The signal is to be detected by carrier reinsertion technique. Find the amplitude of the carrier to be reinserted so that the power in the recovered signal at the demodulator output is 90% of the normalised power. The d.c. component can be neglected and $\omega_c = 2\pi f_c$ and $\omega_n = 2\pi f_m$.

Solution: For a single-tone SSB-SC signal, the waveform after carrier insertion becomes

$$s'(t) = s(t) + c(t)$$

or $$s'(t) = \cos\{(\omega_c + \omega_m)t\} + A\cos(\omega_c t)$$

Using trigonometric relation, we get

$$s'(t) = \cos(\omega_c t)\cos(\omega_m t) - \sin(\omega_c t)\sin(\omega_m t) + A\cos(\omega_c t)$$

or $$s'(t) = [A + \cos(\omega_m t)]\cos(\omega_c t) - \sin(\omega_c t)\sin(\omega_m t) \qquad ...(i)$$

The output of the demodulator is given by

$$v(t) = \sqrt{[A + \cos(\omega_m t)]^2 + [\sin(\omega_m t)]^2}$$

or $$v(t) = \sqrt{A^2 + \cos^2(w_m t) + 2A\cos(w_m t) + \sin^2(w_m t)}$$

or $$v(t) = \sqrt{A^2 + 1 + 2A\cos(\omega_m t)} = [A^2 + 1 + 2A\cos(\omega_m t)]^{1/2}$$

or $$v(t) = \sqrt{A^2+1}\left[1 + \frac{2A}{A^2+1}\cos(\omega_m t)\right]^{1/2} \cong \sqrt{A^2+1}\left[1 + \frac{A}{A^2+1}\cos(\omega_m t)\right]$$

or $$v(t) \cong \sqrt{A^2+1} + \frac{A}{\sqrt{A^2+1}}\cos(\omega_m t) \qquad ...(ii)$$

Neglecting the d.c. component, the normalised power of the detected signal will be

$$P_d = \frac{1}{2}\left(\frac{A^2}{A^2+1}\right)$$

For the recovered signal at the demodulator output to be 90% of this normalised power, we must have

Detector output = 90% of normalised power of SSB modulated waveform

i.e., $P_d = 90\% \times P_n$

Substituting values, we get

$$\frac{1}{2}\left(\frac{A^2}{A^2+1}\right) = 0.90 \times 0.5 = 0.45$$

or $A = 0.9\,(A^2 + 1)$

or $0.9\,A^2 - A^2 + 1 = 0$

or $-0.1\,A^2 = -1$

or $A^2 = 10$ or $A \cong 3$

Thus, amplitude of the reinserted carrier must be almost 3 volt. **Ans.**

Example 3.52. A sinusoidally modulated ordinary AM waveform is shown in figure 3.66.

(*i*) Determine modulation index.

(*ii*) Compute the efficiency.

(*iii*) Also, find the amplitude of the carrier which must be added to attain a modulation index of 0.1.

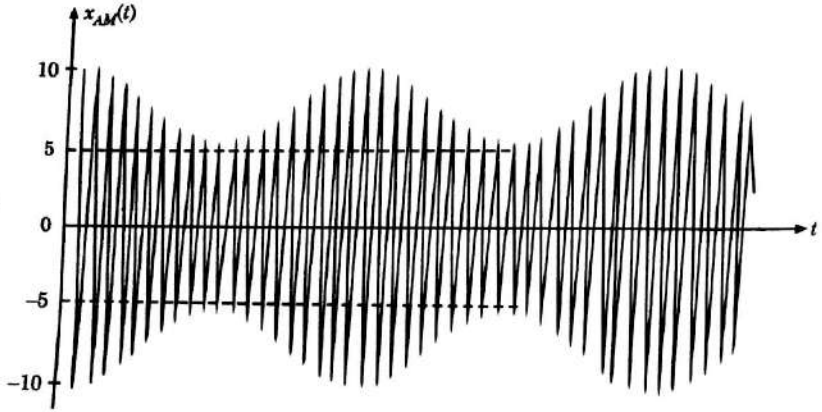


Fig. 3.66.

Solution: (*i*) Modulation index is given by

$$m_a = \frac{A_{max} - A_{min}}{A_{max} + A_{min}}$$

From figure, on substituting values, we get

$$m_a = \frac{10-5}{10+5} = \frac{1}{3}$$

(*ii*) Modulated AM signal is expressed as,

$$x_{AM}(t) = A\cos\omega_c t + m_a A\cos\omega_m t\cos\omega_c t \qquad ...(i)$$

or $$x_{AM}(t) = A\cos\omega_c t + \frac{1}{2}m_a A[\cos(\omega_c - \omega_m)t + \cos(\omega_c + \omega_m)t]$$

Now, P_c = carrier power $= \frac{1}{2}A^2$

P_s = sideband power

or $$P_s = \frac{1}{2}\left[\left(\frac{1}{2}m_a A\right)^2 + \left(\frac{1}{2}m_a A\right)^2\right] = \frac{1}{4}m_a^2 A^2$$

Total power, $$P_t = P_c + P_s = \frac{1}{2}\left(1 + \frac{m_a^2}{2}\right)A^2$$

Thus, efficiency $$\eta = \frac{P_s}{P_t} \times 100\% = \frac{\frac{1}{4}m_a^2}{\left[\frac{1}{2} + \frac{m_a^2}{4}\right]A^2} \times 100\% = \frac{m_a^2}{2 + m_a^2} \times 100\%$$

Substituting $m_a = \frac{1}{3}$, we get

Efficiency, $$\eta = \frac{\left(\frac{1}{3}\right)^2}{2 + \left(\frac{1}{3}\right)^2} \times 100 = 5.26 \text{ per cent}$$

(*iii*) From equation (*i*), we have

$A_{max} = A_1 + m_1 A_1$ [∵ Initially amplitude of carrier wave is A_1]

and $A_{min} = A_1 - m_1 A_1$

then $$m_1 A_1 = \frac{A_{max} - A_{min}}{2}$$

or $$A_1 = \frac{5}{2 \times \frac{1}{3}} = 7.5 \text{ V}$$

Let, new amplitude of the carrier be A_2.

Then, we have

www.dahiyastech.blogspot.com

$$m_2 A_2 = \frac{A_{max} - A_{min}}{2}$$

$$A_2 = \frac{5}{2 \times 0.1} = 25 \text{ V}$$

Thus, amplitude to be added $= A_2 - A_1 = (25 - 7.5)$ V $= 17.5$ V **Ans.**

Example 3.53. Given a real signal $m(t)$, define a signal

$$M_+(t) = m(t) + j\hat{m}(t)$$

where $\hat{m}(t)$ is the Hilbert transform of $m(t)$. And $m_+(t)$ is called an analytic signal.

(i) Show that

$$F[m_+(t)] = M_+(\omega) = \begin{cases} 2M(\omega) & \text{for } \omega > 0 \\ 0 & \text{for } \omega < 0 \end{cases}$$

(ii) Show that $Re[m_+(t)\, e^{j\omega t}]$ is an upper-sideband SSB signal and $Re[m_+(t)\ e^{-j\omega_c t}]$ is a lower lower-sideband SSB signal.

Solution: (i) Given, $M_+(t) = m(t) + j\hat{m}(t)$

Taking Fourier transform of both sides, we get

$$F[m_+(t)] = F[m(t)] = jF[\hat{m}(t)]$$
$$= M(\omega) + j[-j\, sgn(\omega)\, M(\omega)] = M(\omega) + sgn(\omega)\, M(\omega)$$
$$= \begin{matrix} 2M(\omega), & \omega > 0 \\ 0 \quad , & \omega < 0 \end{matrix} \quad \left[\begin{matrix} \text{Using} \\ \text{sgn}(\omega) = \ \ 1, \omega > 0 \\ \qquad\quad = -1, \omega < 0 \end{matrix}\right]$$

(ii) $$\text{Re}[m_+(t)\, e^{j\omega_c t}] = \text{Re}[\{m(t) + j\hat{m}(t)\}\,(e^{j\omega_c t})]$$
$$= \text{Re}[(m(t) + j\,\hat{m}(t)][(\cos\omega_c t + j\sin\omega_c t)]$$
$$= m(t)\cos\omega_c t - \hat{m}(t)\sin\omega_c t$$

which is an **upper** SSB-signal.

Now, $\text{Re}[m_+(t)\, e^{-j\omega_c t}]$

$$= \text{Re}[\{m(t) + j\,\hat{m}(t)\}.e^{-j\omega_c t})]$$
$$= \text{Re}[\{m(t) + j\,\hat{m}(t)\}.(\cos\omega_c t - j\sin\omega_c t)]$$
$$= m(t)\cos\omega_c t + \hat{m}(t).\sin\omega_c(t)$$

which is an **lower** SSB-signal.

Example 3.54. The frequency multiplier is a nonlinear device followed by a bandpass filter, as shown in the figure 3.67. Suppose that the nonlinear device is an ideal square-law device with input-output characteristics:

$$e_o(t) = a\, e_i^2(t)$$

Determine the output $y(t)$ if the input is an FM signal given by

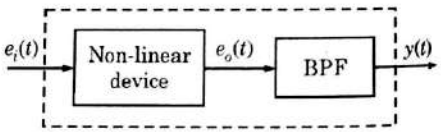


Fig. 3.67.

$$e_i(t) = A\cos(\omega_c t + \beta\sin\omega_m t)$$

Solution: $y(t) = A'\cos(2\omega_c t + 2\beta\sin\omega_m t)$

where $A' = 1/2aA^2$

This result shows that a square-law device can be used as a frequency doubler.

Now, output of the non-linear device will be

$$e_0(t) = ae_i^2(t) = a\,A^2\cos^2(\omega_c t + \beta\sin\omega_m t) = \frac{aA^2}{2}[1 + \cos(2\omega_c t + 2\beta\sin\omega_m t)]$$

After bandpass filtering, the output will be

$$y(t) = \frac{aA^2}{2}\cos(2\omega_c t + 2\beta\sin\omega_m t) = A'\cos(2\omega_c t + 2\beta\sin\omega_m t)$$

where $A' = \dfrac{aA^2}{2}$

Hence Proved.

Example 3.55. The given figure 3.68 shows the circuit diagram of a square-law modulator. The signal applied to the nonlinear device is relatively weak, such that it can be represented by the following square law:

$$v_2(t) = a_1 v_1(t) + a_2 v^2{}_1(t)$$

where a_1 and a_2 are constants, $v_1(t)$ is the input voltage, and $v_2(t)$ is the output voltage. The input voltage is defined as

$$v_1(t) = A_c\cos(2\pi f_c t) + m(t)$$

where $m(t)$ is a message signal and $A_c\cos(2\pi f_c t)$ is the carrier wave.

(i) Evaluate the output voltage $v_2(t)$

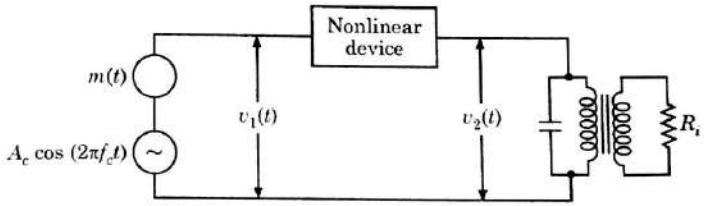


Fig. 3.68.

(ii) Specify the frequency response that the tuned circuit in the given figure must satisfy in order to generate and AM signal with f_c as the carrier frequency.

(iii) What will be the amplitude sensitivity of this AM signal?

Solution: The input voltage $v_1(t)$ is given by

$$v_1(t) = A_c\cos(2\pi f_c t) + m(t) \qquad ..(i)$$

Then, the output voltage will be

$$v_2(t) = a_1[A_c\cos(2\pi f_c t) + m(t)] + a_2[A_c\cos(2\pi f_c t) + m(t)]^2 \qquad ..(ii)$$

(i) $$v_2(t) = a_1 A_c\left[1 + \frac{2a_2}{a_1}m(t)\right]\cos(2\pi f_c t) + a_1 m(t) + a_2 m^2(t) + a_2 A_c^2\cos^2(2\pi f_c t) \qquad ...(iii)$$

Note that the required term for the AM signal is term $a_1 A_c\left[1 + \dfrac{2a_2}{a_1}m(t)\right]\cos(2\pi f_c t)$.

Then, the amplitude sensitivity, will be $k_a = \dfrac{2a_2}{a_1}$.

(ii) The remaining terms after AM signal can be removed by filtering. For this tuned filter, at the modulator output, to be designed such that it has a mid-band frequency f_c and bandwidth $2f_m$.

Example 3.56. Given the AM signal

$$s(t) = A_c[1 + m_a\cos(2\pi f_m t)]\cos(2\pi f_c t)$$

produced by a sinusoidal modulating signal of frequency f_m. Assume that the modulation factor is $m_a = 2$, and the carrier frequency f_c is much greater than f_m. The AM signal $s(t)$ is applied to an ideal envelope detector, producing the output $v(t)$.

(i) Find the Fourier series representation of $v(t)$.

(ii) What will be the ratio of second-harmonic amplitude of fundamental amplitude in $v(t)$

Solution: (i) Envelope detector output is given by

$$v(t) = A_c[1 + m_a\cos(2\pi f_m t)]$$

which has been illustrated in figure 3.69, for the case when $m_a = 2$.

From the figure, it may be observed that, $v(t)$ is periodic with a period equal to f_m, and eve function of t, and so only cosine terms will be present in its Fourier series representation *i.e.*,

$$v(t) = a_0 + 2\sum_{n=1}^{\infty} a_n \cos(2\pi f_m t) \qquad ...(i)$$

where $a_0 = 2f_m \int_0^{\frac{1}{2f_m}} v(t)\,dt$

or $a_0 = 2A_c f_m \int_0^{\frac{1}{3f_m}} [1 + 2\cos(2n\pi f_m t)]\,dt + 2A_c f_m \int_{\frac{1}{3f_m}}^{\frac{1}{2f_m}} [-1 - 2\cos(2\pi f_m t)]\,dt$

or $a_0 = \frac{A_c}{3} + \frac{4A_c}{\pi}\sin\left(\frac{2\pi}{3}\right)$

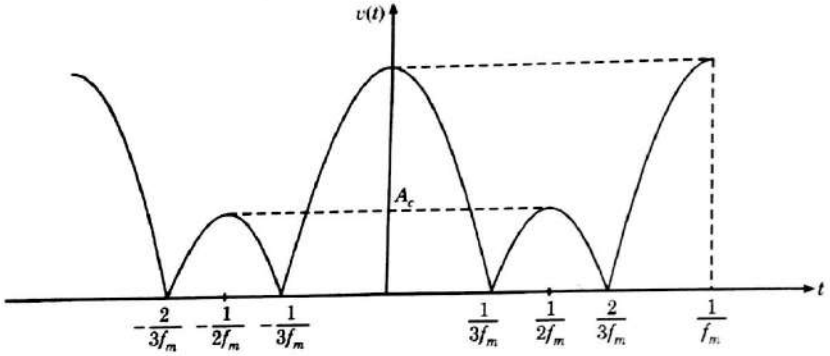


Fig. 3.69.

Now, we have

$$a_n = 2f_m \int_0^{\frac{1}{2f_m}} v(t)\cos(2n\pi f_m t)\,dt$$

or $a_n = 2A_c f_m \int_0^{\frac{1}{3f_m}} [1 + 2\cos(2n\pi f_m t)\cos(2\pi f_m t)\,dt + 2A_c f_m \int_{\frac{1}{3f_m}}^{\frac{1}{2f_m}} [1 - 2\cos(2n\pi f_m t)\cos(2n\pi f_m t)\,dt$

or $a_n = \frac{A_c}{n\pi}\left[2\sin\left(\frac{2n\pi}{3}\right) - \sin(n\pi)\right] + \frac{A_c}{(n+1)\pi}\left\{2\sin\left[\frac{2\pi}{3}(n+1)\right] - \sin[\pi(n+1)]\right\}$

$$+ \frac{A_c}{(n-1)\pi}\left\{2\sin\left[\frac{2\pi}{3}(n-1)\right] - \sin[\pi(n-1)]\right\} \qquad ...(ii)$$

For $n = 0$, equation (*ii*) reduces to that shown in equation (*i*).

(*ii*) For $n = 1$, equation (*ii*) yields

$$a_1 = A_c\left(\frac{\sqrt{3}}{2\pi} + \frac{1}{3}\right)$$

For $n = 2$, it yields

$$a_2 = \frac{A_c\sqrt{3}}{2\pi}$$

Therefore, the required ratio of second-harmonic amplitude to fundamental amplitude in $v(t)$ is

$$\frac{a_2}{a_1} = \frac{3\sqrt{3}}{2\pi + 3\sqrt{3}} = 0.452 \qquad \textbf{Ans.}$$

Example 3.57. Given a square-law detector, using a nonlinear device whose transfer characteristic is defined by

$$v_2(t) = a_1 v_1(t) + a_2 v_1^2(t)$$

where a_1 and a_2 are constants, $v_1(t)$ is the input, and $v_2(t)$ is the output. The input consists of the AM wave

$$v_1(t) = A_c[1 + k_a\, m(t)]\cos(2\pi f_c t)$$

(*i*) Compute the output $v_2(t)$.

(*ii*) Determine the conditions for which the message signal $m(t)$ may be recovered from output $v_2(t)$.

Solution: Given,

$$v_1(t) = A_c\,[1 + k_a m(t)]\cos 2\pi f_c t$$

Then, the output of the square-law device will be

(*i*) $v_2(t) = a_1 v_1(t) + a_2 v_1^2(t)$

$$= a_1 A_c[1 + k_a m(t)]\cos 2\pi f_c t + \frac{1}{2}a_2 A_c^2[1 + k_a m(t) + a + k_a^2 m^2(t)][1 + \cos(4\pi f_c t)] \qquad \textbf{Ans.}$$

(*ii*) The required signal $aA_c^2 k_a m(t)$ can be extracted by a low-pass filter. Also the term $\frac{1}{2}a_2 A_c^2\, k_a^2\, m^2(t)$ will give rise to plurality of similar frequency components within the baseband spectrum. Thus, ratio of wanted signal to distortion is $\frac{2}{k_a m(t)}$. To make this ratio large, $k_a m(t)$ should be small compared to unity. **Ans.**

Example 3.58. The following AM signal

$$s(t) = A_c\,[1 + k_a m(t)]\cos(2\pi f_c t)$$

is applied to the system in the figure 3.70. Assuming that $|k_a m(t)| < 1$ for all t and message signal $m(t)$ is limited to the interval $-f_m \le f \le f_m$, and that the carrier frequency $f_c > 2f_m$, show that $m(t)$ can be obtained from the square-rooter output $v_3(t)$.

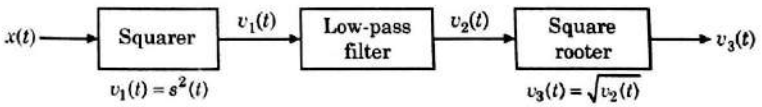


Fig. 3.70.

Solution: Squarer output is given by

$$v_1(t) = s_1^2(t)$$

or $v_1(t) = A_c^2\,[1 + k_a m(t)]^2\cos^2(2\pi f_c t)$

or $v_1(t) = \frac{A_c^2}{2}[1 + 2k_a m(t)] + m^2(t)\,[1 + \cos(4\pi f_c t)]$

The amplitude spectrum of $v_1(t)$ is, therefore, as shown in figure 3.71, assuming that $m(t)$ is limited to the interval $-f_m \le f \le f_m$:

Since $f_c > 2f_m$, then $2f_c - 2f_m > 2f_m$. Therefore, by choosing the cutoff frequency of the low-pass filter greater than $2f_m$, but less than $2f_c - 2f_m$. The output will be

$$v_2(t) = \frac{A_c^2}{2}[1 + k_a m(t)]^2$$

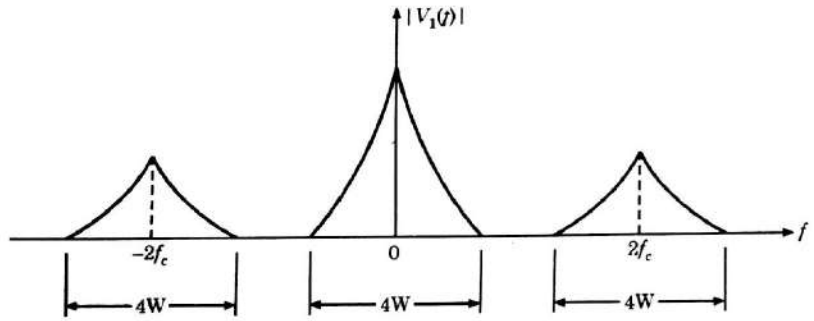


Fig. 3.71.

∴ Output of square rooter is

$$v_3(t) = \sqrt{v_2(t)} = \frac{A_c}{\sqrt{2}}[1 + k_a m(t)]$$

Now, the blocking capacitor will suppress the dc term present, thus, we have

$$v_3(t) = \frac{A_c}{\sqrt{2}} k_a m(t)$$

Hence, the output is proportional to the message signal $m(t)$. **Hence Proved.**

Example 3.59. A DSB-SC modulated signal is demodulated by applying it to a coherent detector.

(*i*) Compute the effect of a frequency error Δf in the local carrier frequency of the detector, measured with respect to the carrier frequency of the incoming DSB-SC signal.

(*ii*) For the case a sinusoidal modulating wave, show that because of this frequency error, the demodulated signal exhibits beats at the error frequency. Give your answer with a sketch of this demodulated signal.

Solution: (*i*) We know that the output of a product modulator is given by

$$v(t) = A_c m(t) \cos(2\pi f_c t) \cos[2\pi(f_c + \Delta f)t]$$

$$= \frac{A_c}{2} m(t) \{\cos(2\pi \Delta f_c t) + \cos[2\pi(2f_c + \Delta f)t]\}$$

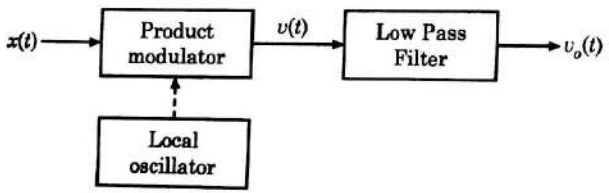


Fig. 3.72.

Then after low pass filtering, we have

$$v_0(t) = \frac{A_c}{2} m(t) \cos(2\pi f_c t)$$

Hence, the output signal is the message signal modulated by a sinusoid of frequency Δf.

(*ii*) If $\quad m(t) = \cos(2\pi f_m t)$

then $\quad s_2(t) = \frac{A_c}{2} \cos(2\pi f_m t) \cos(2\pi \Delta f t)$

or $\quad s_2(t) = \frac{A_c}{4} \{\cos[2\pi(f_m + \Delta f)t] + \cos[2\pi(f_m - \Delta f)t]\}$

the waveform of $s_2(t)$ will be as shown in figure 3.73.

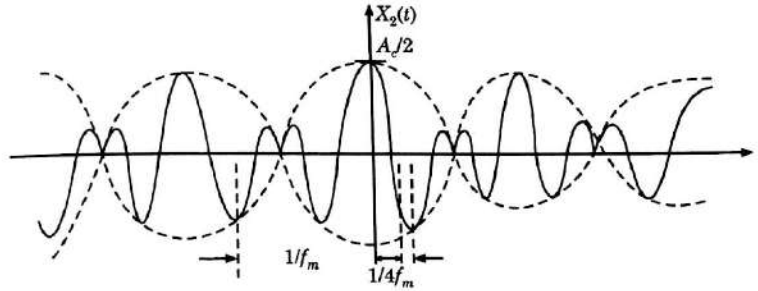


Fig. 3.73.

Example 3.60. The signal tone modulating $m(t) = A_m \cos(2\pi f_m t)$ is used to generate the following VSB signal:

$$s(t) = \frac{1}{2} a A_m A_c \cos 2\pi(f_c + f_m)t] + \frac{1}{2} A_m A_c (1 - a) \cos[2\pi(f_c - f_m)t]$$

where a is constant, less than unity, representing the attenuation of the upper side frequency.

(*i*) Determine the quadrature component of the VSB signal $s(t)$.

(*ii*) The VSB signal, plus the carrier $A_c \cos(2\pi f_c t)$, is passed through an envelope detector. Determine the distortion produced by the quadrature component.

(*iii*) What will be the value of constant a for which this distortion reaches its worst possible condition ? *(GATE Examination, 1991)*

Solution: (*i*) We know that any signal $s(t)$ can be expressed in terms of in phase and quardative components as under:

$$s(t) = s_I(t) \cos(2\pi f_c t) - s_Q(t) \sin(2\pi f_c t) \quad \text{...(i)}$$

Now, expanding the given signal $s(t)$, we have

$$s(t) = \frac{1}{2} a A_m A_c \cos(2\pi f_c t) \cos(2\pi f_m t) - \frac{1}{2} a A_m A_c \sin(2\pi f_c t) \sin(2\pi f_m t)$$

$$+ \frac{1}{2}(1 - a) A_c A_m \cos(2\pi f_c t) \cos(2\pi f_m) t + \frac{1}{2}(1 - a) A_m A_c \sin(2\pi f_c t) \sin(2\pi f_m t)$$

or $$s(t) = \frac{1}{2} A_m A_c \cos(2\pi f_c t) \cos(2\pi f_m t) + \frac{1}{2} A_m A_c (1 - 2a) \sin(2\pi f_c t) \sin(2\pi f_m t) \quad \text{...(ii)}$$

Now, comparing equation (*ii*) with equation (*i*), we have the quadrative component,

$$s_Q(t) = -\frac{1}{2} A_c A_m (1-2a)\sin(2\pi f_m t)$$

(*ii*) Further, the VSB signal after adding carrier signal $A_c \cos(2\pi f_c t)$, will be

$$s(t) = A_c\left[1+\frac{A_m}{2}\cos(2\pi f_m t)\right]\cos(2\pi f_c t) + \frac{1}{2}A_c A_m(1-2a)\sin(2\pi f_m t)\sin(2\pi f_c t)$$

Then, envelope will be equal to

$$a(t) = \sqrt{s_Q^2(t)+s_I^2(t)} = A_c\sqrt{\left[1+\frac{1}{2}A_m\cos(2\pi f_m t)\right]^2 + \left[\frac{1}{2}A_m(1-2a)\sin(2\pi f_m t)\right]^2}$$

or $$a(t) = A_c\left[1+\frac{1}{2}A_m\cos(2\pi f_m t)\right]\sqrt{1+\left[\frac{\frac{1}{2}A_m(1-2a)\sin(2\pi f_m t)}{1+\frac{1}{2}A_m\cos(2\pi f_m t)}\right]^2}$$

or $$a(t) = A_c\left[1+\tfrac{1}{2}A_m\cos(2\pi f_m t)\right]d(t)$$

where distortion is given by

$$d(t) = \sqrt{1+\left[\frac{\frac{1}{2}A_m(1-2a)\sin(2\pi f_m t)}{1+\frac{1}{2}A_m\cos(2\pi f_m t)}\right]^2} \quad \text{...(iii)}$$

(*iii*) From equation (*iii*), it is clear that the distortion will be maximum when, $a = 0$. **Ans.**

Example 3.61. Given the following modulated wave

$$s(t) = A_c\cos(2\pi f_c t) + m(t)\cos(2\pi f_c t) - \hat{m}(t)\sin(2\pi f_c t)$$

which represents a carrier plus an SSB signal, with $m(t)$ denoting the message signal and $\hat{m}(t)$ its Hilbert transform. Find the conditions for which an ideal envelope detector, with $s(t)$ as input, would produce a good approximation to the message signal $m(t)$.

Solution: Given modulated wave is

$$s(t) = [A_c + m(t)]\cos 2\pi f_c(t) - \hat{m}(t)\sin(2\pi f_c t)$$

then, in phase component

$$s_I(t) = A_c + m(t)$$

and, quadrative component

$$s_Q(t) = +\hat{m}(t)$$

then, envelope detector output will be

$$a(t) = \sqrt{s_I^2(t)+s_Q^2(t)} = \sqrt{[A_c+m(t)]^2+[\hat{m}(t)]^2}$$

If carrier amplitude is such that,

$A_c >> |m(t)|$ and $A_c >> |\hat{m}(t)|$, for all values of t,

then, we have $$a(t) \cong \sqrt{A_c^2 + 2A_c m(t)} \cong A_c\sqrt{1+\frac{2}{A_c}m(t)}$$

or $$a(t) \cong A_c\left[1+\frac{m(t)}{A_c}\right] = A_c + m(t)$$

It may be observed that except for dc bias A_c, the envelope $a(t)$ is approximately equal to $m(t)$. **Ans.**

Example 3.62. Given the SSB wave

$$s(t) = m(t)\cos(2\pi f_c t) - \hat{m}(t)\sin(2\pi f_c t)$$

where f_c is the carrier frequency, $m(t)$ is the message signal and $\hat{m}(t)$ it its Hilbert transform. This modulated wave is applied to a square-law device characterized by

$$y(t) = s^2(t)$$

Prove that the output $y(t)$ contains a frequency components at twice the carrier frequency but that it has a time-varying phase, which makes it impractical to recover the carrier by squaring.

Solution: The square law device output, with SSB wave $s(t)$ as input, is given by

$$y(t) = s^2(t) = [m(t)\cos(2\pi f_c t) - \hat{m}(t)\sin(2\pi f_c t)]^2$$

$$= m^2(t)\cos^2(2\pi f_c t) + \hat{m}^2(t)\sin^2(2\pi f_c t) - 2m(t)\hat{m}(t)\cos(2\pi f_c t)\sin(2\pi f_c t)$$

$$= \frac{1}{2}m^2(t)[1+\cos(4\pi f_c t)] + \frac{1}{2}\hat{m}^2(t)[1-\cos(4\pi f_c t)] - m(t)\hat{m}(t)\sin(4\pi f_c t)$$

$$y(t) = \frac{1}{2}[m^2(t)+\hat{m}^2(t)] + \left\{\frac{1}{4}[m^2(t)+\hat{m}^2(t)]^2 + m^2(t)\hat{m}^2(t)\right\}^{1/2} - \cos[4\pi f_c t + Q(t)]$$

where $$Q(t) = \tan^{-1}\left[\frac{2m(t)\hat{m}(t)}{m^2(t)+\hat{m}^2(t)}\right]$$

Since, $y(t)$ is a mixed combination of amplitude and frequency modulation. The AM can be eliminated by filtering but the phase modulation is unaffected by such an operation, which makes it impractical to recover the carrier wave by squaring. **Hence Proved.**

Example 3.63. Single-sideband modulation can be viewed as a hybrid form of amplitude modulation and frequency modulation. Compute the envelope and instantaneous frequency of an SSB wave for the following two cases:

(*i*) When only the upper sideband is transmitted.

(*ii*) When only the lower sideband is transmitted.

Solution: (*i*) For SSB modulation, the modulated wave is given by

$$s(t) = \frac{A_c}{2}[m(t)\cos(2\pi f_c t) \pm \hat{m}(t)\sin(2\pi f_c t)]$$

where +ve and –ve sign corresponds to upper and lower sideband of the modulated wave respectively. Now, the envelope for the SSB wave will be given by,

$$a(t) = \frac{A_c}{2}\sqrt{m^2(t)+\hat{m}^2(t)}$$

which is same for both transmitted upper sideband and lower sideband.

For upper sideband transmission, the angle, is given by

$$Q_i(t) = 2\pi f_c t + \tan^{-1}\frac{\hat{m}(t)}{m(t)}$$

and the instantaneous frequency, will be

$$f_i(t) = \frac{1}{2\pi}\frac{d\theta_i(t)}{dt}$$

Thus, we have

or $$f_i(t) = f_c + \frac{m(t)\hat{m}(t) - \hat{m}(t)m'(t)}{2\pi[m^2(t)+\hat{m}^2(t)]}$$

here ' ' denotes time derivative.

(*ii*) For lower sideband transmission, the angle is given by

$$\theta_i(t) = 2\pi f_c t + \tan^{-1}\left(-\frac{\hat{m}(t)}{m(t)}\right)$$

and, the instantaneous frequency will be

$$f_i(t) = f_c + \frac{\hat{m}(t)\, m'(t) - m(t)\hat{m}'(t)}{2\pi[m^2(t) + \hat{m}^2(t)]} \quad \textbf{Ans.}$$

Example 3.64. An SSB signal is demodulated by using a synchronous demodulator. However, the locally arranged carrier has a phase error θ. Determine the effect of the error on demodulation. What will be the effect of this error if the input is DSB-SC in place of SSB?

(GATE Examination, 1998)

Solution: We know that an SSB signal is expressed by,

$$x_{SSB}(t) = m(t).\cos(\omega_c t) \mp \hat{m}(t).\sin(\omega_c t)$$

where $\hat{m}(t)$ is the Hilbert transform of message signal $m(t)$.

In a synchronous demodulator shown in figure 3.74, an SSB signal is present at the input.

Then, we have $\quad d(t) = [m(t).\cos(\omega_c t) \mp \hat{m}(t).\sin(\omega_c t)]\cos|\omega_c t + \theta|$

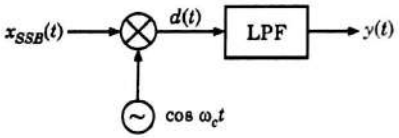


Fig. 3.74.

or $$d(t) = \frac{m(t)}{2}.[\cos(2\omega_c t + \theta) + \cos\theta] \mp \hat{m}(t).[\sin(2\omega_c t + \theta) - \sin\theta)]$$

After low-pass filtering, we get

Output $$y(t) = \frac{m(t)}{2}.\cos\theta \mp \hat{m}(t).\sin\theta$$

Hence, there is no attenuation but only phase distortion by angle θ.

Now, when the input is DSB-SC

then, we have $\quad x_{DSB}(t) = m(t).\cos(\omega_c t + \theta)$

and, $$d(t) = [m(t).\cos(\omega_c t)].\cos\{\omega_c t + \theta] = \frac{1}{2}m(t)[\cos\theta + \cos(2\omega_c t + \theta)]$$

after low-pass filtering, we get the output

$$y(t) = \frac{1}{2}m(t).\cos\theta$$

Hence, output is proportional to $m(t)$, when θ is is constant. But if θ varies randomly with time, then the output will also vary which is undesirable.

Example 3.65. Derive an expression for the signal $v_3(t)$ in figure 3.75 for $v_1(t) = 10\cos(2000\pi t) + 4\sin(200\pi t)$. Assume that $v_2(t) = v_1(t) + 0.1\, v_1^2(t)$ and that the BPF is an ideal unity gain filter with passband from 800 Hz to 1200 Hz.

(GATE Examination-1998)

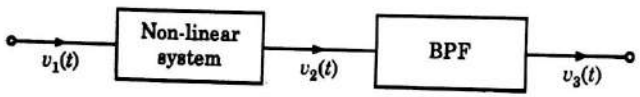


Fig. 3.75.

Solution: From the given figure, we obtain

Output at non-linear system,

$$v_2(t) = v_1(t) + 0.1 v_1^2(t)$$

or $$v_2(t) = [10\cos(2000\pi t) + 4\sin(200\pi t)] + [10\cos(2000\pi t) + 4\sin(200\pi t)]^2 \times 0.1$$

or $$v_2(t) = 10\cos(2000\pi t) + 4\sin(200\pi t) + 0.1 \times [100\cos^2(200\pi t) + 16\sin^2(200\pi t) + 80\cos(2000\pi t)\sin(200\pi t)]$$

or $$v_2(t) = 10\cos(2000\pi t) + 4\sin(200\pi t) + \left[100\left(\frac{1+\cos(4000\pi t}{2}\right) + \frac{16}{2}\frac{[1-\cos|900\pi t|}{2}\right] \times 0.1 + 40[\sin(2200\pi t) - \sin(1800\pi t)] \times 0.1$$

or $$v_2(t) = 10.\cos(2000\pi t) + 0.5\cos(4000\pi t) - 0.8\cos(400\pi t) + 4\sin(200\pi t) + 4\sin(2200\pi t) - 4\sin(1800\pi t) + 5.8$$

Because, BPF has a pass-band from 800 Hz-1200 Hz, therefore the output after filtering will be

$$v_3(t) = 10\cos(2000\pi t) + 4\sin(2200\pi t) - 4\sin(1800\pi t) \quad \textbf{Ans.}$$

Example 3.66. A given AM broadcast station transmits a total power of 5 kW when the carrier is modulated by a sinusoidal signal with a modulation index of 0.7071. Compute

(*i*) the carrier power

(*ii*) the transmission efficiency, and

(*iii*) the peak amplitude of the carrier assuming the antenna to be represented by a (50 + *j*0) Ω load.

Solution: (*i*) An amplitude modulated signal can be expressed as under:

$$x_{AM}(t) = A\cos\omega_c t + mA\cos\omega_m t\cos\omega_c t$$

or $$x_{AM}(t) = A\cos\omega_c t + \frac{1}{2}mA\cos(\omega_c - \omega_m)t + \frac{mA}{2}.\cos(\omega_c + \omega_m)t \quad ...(i)$$

Now, $$P_c = \text{carrier power} = \frac{A^2}{2}$$

$$P_s = \text{sideband power} = \frac{1}{2}\left[\left(\frac{1}{2}mA\right)^2 + \left(\frac{mA}{2}\right)^2\right] = \frac{m^2A^2}{4} \quad ...(ii)$$

We know that the total power is given by

$$P_t = \frac{A^2}{2}\left(1 + \frac{m^2}{2}\right) = P_c\left(1 + \frac{m^2}{2}\right)$$

Given $\quad P_t = 50 \times 10^3$ watt,

then, $$P_c = \frac{P_t}{1 + \frac{m^2}{2}} = \frac{50 \times 10^3}{1.25} = \mathbf{40\ kW.}\ \textbf{Ans.}$$

(*ii*) Further, we know that the transmission efficiency is given by

$$\eta = \frac{P_s}{P_t} \times 100\% = \frac{\frac{1}{4}m^2A^2}{\left(\frac{1}{2} + \frac{m^2}{4}\right)A^2} \times 100\%$$

or $$\eta = \frac{m^2}{2 + m^2} \times 100\% = \frac{0.50}{2 + 0.50} \times 100\% = 20\% \quad \textbf{Ans.}$$

(*iii*) Power delivered to the load will be

or $$A_c = \frac{A_c^2}{2R} = 40 \times 10^3 = \sqrt{2 \times 40 \times 10^3 \times 50} = 2 \text{ kV} \quad \textbf{Ans.}$$

Example 3.67. Show that in quadrature amplitude moduation (QAM) each signal can be recovered by synchronous detection of the received signal by using two local carriers of same frequency but in phase quadrature.

Solution: A QAM scheme has been shown in figure 3.76.

In the figure, the boxes labelled $\frac{-\pi}{2}$ are phase shifters. If the two baseband signals to be transmitted are $m_1(t)$ and $m_2(t)$, then, we will get

$$x_{QAM}(t) = m_1(t) \cos \omega_c t + m_2(t). \sin \omega_c t$$

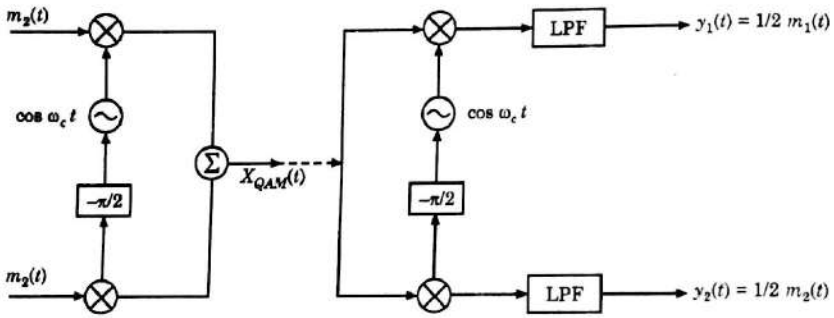


Fig. 3.76.

The output $x_1(t)$ of the upper multiplier will be

$$x_1(t) = 2x_{QAM}(t). \cos \omega_c t = 2[m_1(t) \cos \omega_c t + m_2(t) . \sin \omega_c t]. \cos \omega_c t$$
$$= m_1(t) + m_1(t) \cos 2\omega_c t + m_2(t). \sin 2\omega_c t$$

After low-pass filtering, we have

$$x_1(t) = m_1(t)$$

Now, in a similar fashion, we get

$$x_2(t) = 2x_{QAM}(t). \sin \omega_c t$$
$$= 2[m_1(t). \cos \omega_c t + m_2(t).\sin \omega_c t]. \sin \omega_c t$$

or $$x_2(t) = m_1(t) \sin 2\omega_c t + m_2(t) - m_2(t). \cos 2\omega_c t$$

Again, low-pass filtering yields

$$x_2(t) = m_2(t)$$

It may noted that Q_{AM} is an efficient method of transmitting two message signals within the same bandwidth.

Example 3.68. Explain the Quadrature null effect in a coherent detector for DSB-SC signals.

Solution: A coherent detector has been depicted in figure 3.77. Here, it is assumed that the local oscillator signal is exactly coherent or synchronized, in both frequency and phase, with the carrier wave. However, the more general demodulation process is using a local oscillator having same frequency but arbitrary phase difference ϕ with respect to carrier wave. Hence, expressing local oscillator signal as, $A_c' \cos (2\pi f_c t + \phi)$, the product modulator output will be

$$v_0(t) = A_c' \cos (2\pi f_c t + \phi) . s(t) = A_c' A_c \cos (2\pi f_c t + \phi). \cos (2\pi f_c t)\ m(t) \left[A_c \cos 2\pi = \frac{1}{2}\right]$$

[Here, $s(t) = m(t) = m(t). A_c \cos 2\pi f_c t$]

or $$v_0(t) = \frac{1}{2} A_c A'_c (\cos 4\pi f_c t + \phi(t)\ m(t) + \frac{1}{2} A_c A'_c \cos \phi\ m(t)$$

The first term is removed by low-pass filtering having cut-off frequency of this filter is greater than W but less than $2f_c - W$.

Then, filter output would be

$$v_0(t) = \frac{1}{2} A_c A'_c \cos \phi\ m(t)$$

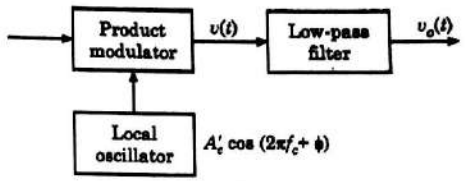


Fig. 3.77.

The demodulated signal $v_0(t)$ will therefore proportional to $m(t)$ if ϕ is constant.

The amplitude goes maximum for $\phi = 0$, and minimum for $\phi = \pm\frac{\pi}{2}$. The zero demodulator signal occurring at $\phi = \pm\frac{\pi}{2}$ represents the **quadrature null effect** of synchronous demodulator or **coherent detector.**

Example 3.69. In an AM system, the modulating signal is sinusoidal with frequency f_m Hz. If 80% modulation is used, then find the ratio of total sideband power in the modulation signal, to the total power.

Solution: The AM signal is expressed as

$$x_{AM}(t) = A \cos \omega_c t + mA \cos \omega_m t \cos \omega_c t$$
$$= A \cos \omega_c t + \frac{mA}{2} [\cos(\omega_c - \omega_m)t + \cos(\omega_c + \omega_m)t]$$

$$P_c \text{ (carrier power)} = \frac{1}{2} A^2$$

$$P_s \text{ (sideband power)} = \frac{1}{2}\left[\left(\frac{1}{2} mA\right)^2 + \left(\frac{1}{2} mA\right)^2\right] = \frac{1}{4} m^2 A^2$$

Total power is given by

$$P_s = P_c + P_s = \frac{1}{2}\left(1 + \frac{m^2}{2}\right) A^2$$

$$\text{Ratio} = \frac{\text{Sideband power}}{\text{Total power}} = \frac{1/4 m^2 A^2}{1/4(2 + m^2) A^2} = \frac{m^2}{2 + m^2}$$

$$= \frac{(0.80)^2}{2 + (0.80)^2} = \frac{0.64}{2.64} = \mathbf{0.242.} \quad \textbf{Ans.}$$

Example 3.70. In an envelope detector, the input is an AM signal which is expressed as

$$x_{AM}(t) = A(1 + m \cos \omega_m t). \cos \omega_c t$$

where m is a constant, $0 < m < 1$, and $\omega_c >> \omega_m$. Then prove that

(*i*) At any time t_0

$$\frac{1}{RC} \geq \omega_m \left(\frac{m \sin \omega_m t_0}{1 + m \cos \omega_m t_0}\right) \text{ and}$$

(*ii*) $$\frac{1}{RC} \geq \frac{m\omega_m}{\sqrt{1 - m^2}}$$

Solution: Figure 3.78 shows envelope detector and envelope of $x_{AM}(t)$.

Voltage across capacitor C is given by

$$v_c(t) = E_0 e^{-t/RC}$$

Interval between two peaks is

$$= \frac{1}{f_c} = \frac{2\pi}{\omega_c}$$

and $R_C >> \frac{1}{\omega_c}$, then $v_c(t)$ discharges exponentially for a short time compared to its time constant. Hence, the exponential can be approximated by a straight line, *i.e.*,

$$v_c(t) \cong E\left(1 - \frac{t}{RC}\right)$$

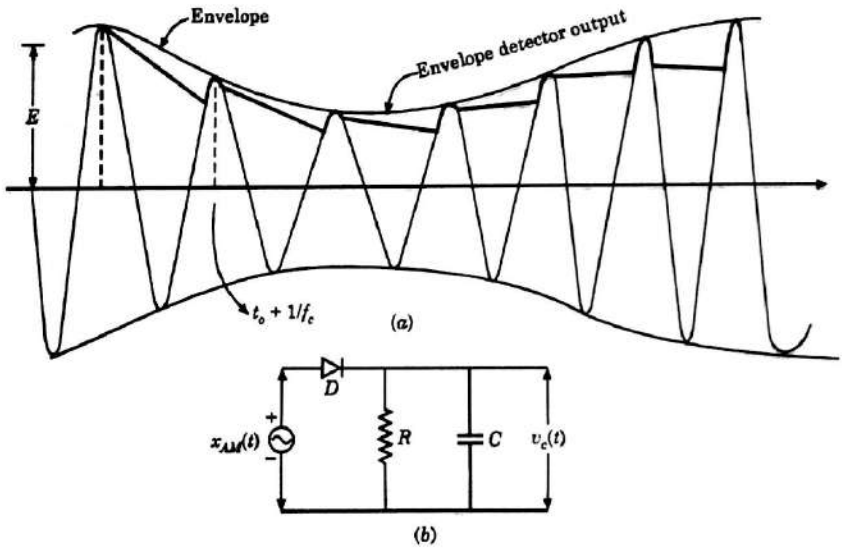


Fig. 3.78.

Because, in the envelope detector, the output of the detector follows the envelope of modulated signal, it is prerequisite at any time t_0,

$$(1 + m\cos\omega_m t_0)\left(1 - \frac{1}{RC f_c}\right) \le 1 + m\cos\omega_m\left(t_0 + \frac{1}{f_c}\right)$$

If $\omega_m << \omega_c$, we shall have

$$1 + m\cos\omega_m\left(t_0 + \frac{1}{f_c}\right) = 1 + m\cos\left(\omega_m t_0 + \frac{\omega_m}{f_c}\right)$$

$$\cong 1 + m\cos\omega_m t_0 - \frac{m\omega_m}{f_c}.\sin\omega_m t_0$$

or $$(1 + m\cos\omega_m t_0)\left(\frac{1}{RC f_c}\right) \ge \frac{m\omega_m}{f_c}\sin\omega_m t_0$$

or $$\frac{1}{RC} \ge \omega_m.\left(\frac{m\sin\omega_m t_0}{1 + m\cos\omega_m t_0}\right)$$

(*ii*) Now from equation (*i*), we get

$$\frac{1}{RC} + \frac{m}{RC}\cos\omega_m t_0 \ge m\,\omega_m \sin\omega_m t_0$$

or $$m\left(\omega_m \sin\omega_m t_0 - \frac{1}{RC}\cos\omega_m t_0\right) \le \frac{1}{RC}$$

or $$m\sqrt{\omega_m^2 + \frac{1}{(RC)^2}}.\sin\left(\omega_m t_0 - \tan^{-1}\frac{1}{\omega_m R_C}\right) \ge \frac{1}{RC}$$

Because $\sin\theta \le 2$

Therefore, we have $\frac{1}{RC} \ge m\sqrt{\omega_m^2 + \left(\frac{1}{RC}\right)^2}$

or $$m^2\left[\omega_m^2 + \left(\frac{1}{RC}\right)^2\right] \le \left(\frac{1}{RC}\right)^2$$

or $$RC \le \frac{1}{\omega_m}\frac{\sqrt{1-m^2}}{m}$$ **Hence Proved.**

Example 3.71. Evaluate the condition for distortionless demodulation of a VSB signal, initially generated by passing a PSB signal through a vestigial filter, using synchronous detector.

Solution:

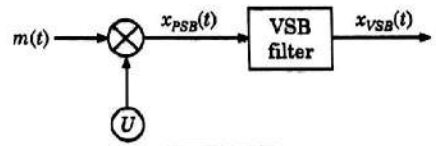


Fig. 3.79.

We have $x_{DSB}(t) = 2m(t)\cos\omega_c t$

Its spectrum is given by

$$X_{DSB}(\omega) = M(\omega - \omega_c) + M(\omega + \omega_c)$$

then, we have

$$X_{VSB}(\omega) = X_{DSB}(\omega).H(\omega)$$

where $H(\omega)$ is frequency response of the VSB filter

$$X_{VSB}(\omega) = [M(\omega - \omega_c) + M(\omega + \omega_c)]\,H(\omega) \quad ...(i)$$

Now, for demodulation using synchronous demodulator, we have

$$d(t) = x_{VSB}(t).\cos\omega_c t$$

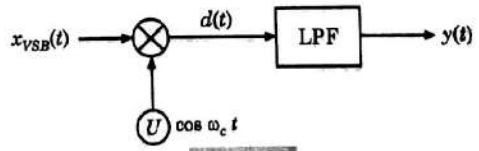


Fig. 3.80.

the Fourier transform is given by

$$D(\omega) = \frac{1}{2}[X_{VSB}(\omega - \omega_c) + X_{VSB}(\omega + \omega_c)] \quad ...(ii)$$

Using equations (*i*) and (*ii*), and eliminately the spectra at $\pm 2\omega_c$ by LPF, we get

$$y(t) \leftrightarrow \frac{1}{2} M(\omega)[H(\omega+\omega_c) + H(\omega-\omega_c)]$$

Now for distortionless detection, we must have

$$y(t) = km(t) \leftrightarrow kM(\omega)$$

where k is constant.

Hence, for distortionless demodulation, we have

$$H(\omega+\omega_c) + H(\omega-\omega_c) = \text{constant}, \ |\omega| \le \omega_M$$

Example 3.72. Prove that the output of the phase-shift modulator is an SSB and verify that upper-sideband (USB) or a lower sideband (LSB) signal results from subtraction or addition at the summer.

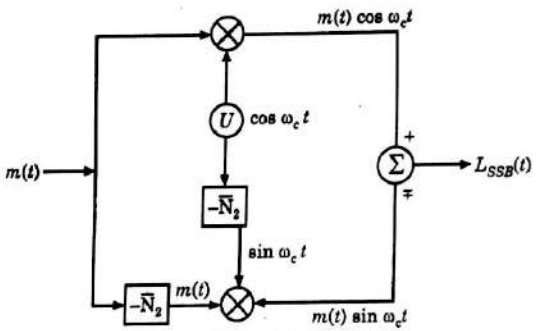


Fig. 3.81.

Solution:

Let $\quad m(t) = \cos \omega_m t$

Hilbert transform of $m(t)$ is given by

$$\hat{m}(t) = \cos|(\omega_m t - \pi/2)| = \sin \omega_m t$$

Then, $\quad x_{SSB}(t) = \cos \omega_m t \cos\omega_c t \mp \sin \omega_m t.\ \sin \omega_c t = \cos(\omega_c \pm \omega_m)t$

Now, with positive sign, we have

$$x_{SSB}(t) = \cos(\omega_c + \omega_m)t = m(t)\cos\omega_c t - \hat{m}(t).\sin\omega_c t \qquad ...(i)$$

and $\quad x_{SSB}(t) = \sin(\omega_c - \omega_m)t = m(t)\cos_c t + \hat{m}(t).\sin\omega_c t \qquad ...(ii)$

Now, tacking Fourier transform of equation (*i*), we get

$$X_{SSB}(\omega) = \frac{1}{2}M(\omega-\omega_c) + \frac{1}{2}M(\omega+\omega_c) - \frac{1}{2j}[\hat{M}(\omega-\omega_c) - \hat{M}(\omega+\omega_c)]$$

Using $\quad m(t).\cos\omega_c t \leftrightarrow \frac{1}{2}[M(\omega-\omega_c) + M(\omega+\omega_c)]$

and $\quad \hat{m}(t).\sin\omega_c t \leftrightarrow \frac{1}{2j}[\hat{M}(\omega-\omega_c) - \hat{M}(\omega+\omega_c)]$

Now, $\quad \hat{M}(\omega-\omega_c) = -j\,\text{sgn}(\omega-\omega_c).\ M(\omega-\omega_c)$

$\quad \hat{M}(\omega+\omega_c) = -j\text{sgn}(\omega+\omega_c).\ M(\omega+\omega_c)$

We have

$$X_{SSB}(\omega) = \frac{1}{2}M(\omega-\omega_c) + \frac{1}{2}M(\omega+\omega_c) - \left[-\frac{1}{2}\sin(\omega-\omega_c).\ M(\omega-\omega_c) + \frac{1}{2}sgn(\omega+\omega_c)\ M(\omega+\omega_c)\right]$$

Since, $\quad 1 + \text{sgn}(\omega-\omega_c) = \begin{cases} 2 & \text{for } \omega > \omega_c \\ 0 & \text{for } \omega < \omega_c \end{cases}$

and $\quad 1 - \text{sgn}(\omega-\omega_c) = \begin{cases} 2 & \text{for } \omega < -\omega_c \\ 0 & \text{for } \omega > -\omega_c \end{cases}$

Then, $\quad X_{SSB}(\omega) = \begin{cases} 0 & \text{for } |\omega| < \omega_c \\ M(\omega+\omega_c) & \text{for } \omega < -\omega_c \\ M(\omega-\omega_c) & \text{for } \omega > \omega_c \end{cases}$

Its sketch has been shown in figure 3.82.

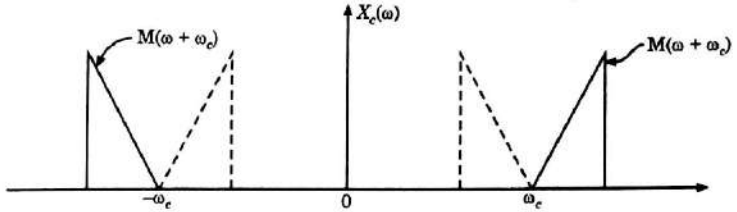


Fig. 3.82.

Its spectrum shows that X_{SSB} is an USB – SSB signal. In a similar fashion, the Fourier transform of equation (*ii*) yields

$$X_{SSB}(\omega) = \frac{1}{2}M(\omega-\omega_c)\ [1 - sgn(\omega-\omega_c)] + \frac{1}{2}M(\omega+\omega_c)\ [1 + sgn(\omega+\omega_c)]$$

As, $\quad 1 - \text{sgn}(\omega-\omega_c) = \begin{cases} 2 & \text{for } \omega < \omega_c \\ 0 & \text{for } \omega > \omega_c \end{cases}$

and $\quad 1 + \text{sgn}(\omega+\omega_c) = \begin{cases} 2 & \text{for } \omega > -\omega_c \\ 0 & \text{for } \omega < -\omega_c \end{cases}$

Example 3.73. The frequency response $H(\omega)$ of a VSB filter has been shown in figure 3.83.

(*a*) Find the VSB signal $x_{VSB}(t)$ when

$$m(t) = a_1 \cos \omega_1 t + a_2 \cos \omega_2 t$$

(*b*) Show that $x_{VSB}(t)$ can be demodulated by a synchronous demodulator.

(Vinayaka University, 2001)

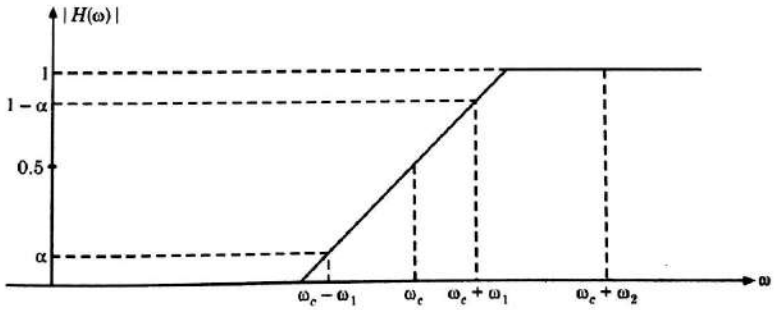


Fig. 3.83.

Solution: (*a*) We have

(*i*) $\quad x_{DSB}(t) = m(t)\cos\omega_c t = (a_1 \cos \omega_1 t + a_2 \cos \omega_2 t)\cos \omega_c t$

or $$x_{DSB}(t) = \frac{1}{2}a_1\cos(\omega-\omega_1)t + \frac{1}{2}a_1\cos(\omega_c+w_1)t + \frac{1}{2}a_2\cos(\omega_c-\omega_2)t + \frac{1}{2}a_2\cos(\omega_c-\omega_2)t$$

These sinusoids are transmitted through $H(\omega)$, shown in figure 3.85, which has gains of 0, α, $1-\alpha$, and 1 at $\omega_c-\omega_2$, $\omega_c-\omega_1$, $\omega_c+\omega_1$, and $\omega_c+\omega_2$, respectively. Thus, the VSB filter output $x_{VSB}(t)$ is

$$x_{VSB}(t) = \frac{1}{2}a_1\alpha\cos(\omega_c-\omega_1)t + \frac{1}{2}a_1(1-\alpha)\cos(\omega_c+\omega_1)t + \frac{1}{2}a_2\cos(\omega_c-\omega_2)t \quad ...(i)$$

(*b*) Referring to figure 3.85, we get

$$d(t) = x_{VSB}(t)\cos\omega_c t$$
$$= \frac{1}{4}\alpha_1\cos\omega_1 t + \alpha_2\cos\omega_2 t) + \frac{1}{4}[\alpha_1\alpha\cos(2\omega_c-\omega_1)t$$
$$+ a_1(1-\alpha)\cos(2\omega_c+\omega_1)t + a_2\cos(2\omega_c+\omega_2)t]$$

Using low-pass filtering to eliminate the double-frequency terms, we obtain

$$y(t) = \frac{1}{4}(a_1\cos\omega_1 t + a_2\cos\omega_2 t) = \frac{1}{4}m(t)$$

Example 3.74. A radio receiver used in the AM system is shown in figure 3.84. The mixer translates the frequency f_c to a fixed. If 455 kHz by using a local oscillator of frequency f_{LO}. The broadcase-band frequencies range from 540 to 1600 kHz.

(*a*) Determine the range of tuning that must be provided in the local oscillator (*i*) when f_{LO} is higher than f_c (superheterodyne receiver) and (*ii*) when f_{LO} is lower than f_c.

(*b*) Explain why the usual AM radio receiver uses a superheterodyne system.

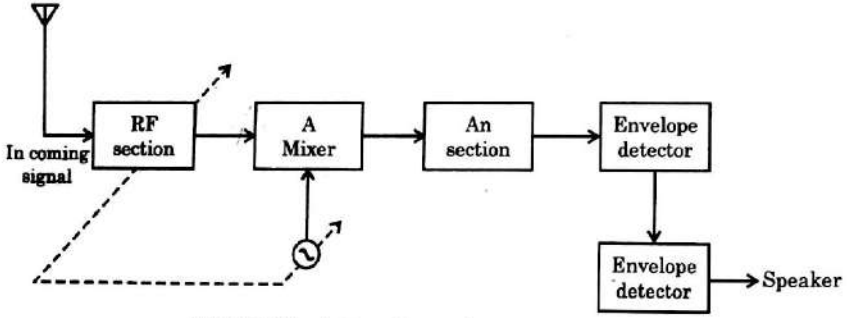


Fig. 3.84. *A Superheterodyne AM receiver.*

Solution: (*a*) (*i*) When $f_{LO} > f_c$,

$$540 < f_c < 1600$$
$$f_{LO} - f_c = 455$$

where both f_c and f_{LO} are expressed in kilohertz. Thus

$$f_{LO} = f_c + 455$$

When $f_c = 540$ kHz, we get $f_{LO} = 995$ kHz; and $f_c = 1600$ kHz, we get $f_{LO} = 2055$ kHz. Thus, the required tuning range of the local oscillator is 995 – 2055 kHz.

(*ii*) When $f_{LO} < f_c$,

$$f_{LO} = f_c - 455$$

When $f_c = 540$ kHz, we get $f_{LO} = 85$ kHz; and when $f_c = 1600$ kHz, we get $f_{LO} = 1145$ kHz. Thus, the required tuning range of the local oscillator for this case is 85 – 1145 kHz.

(*b*) The frequency ratio, that is, the ratio of the highest f_{LO} to the lowest f_{LO} is 2.07 for case (*i*) and 13.47 for case (*ii*). It is much easier to design an oscillator that is tunable over a smaller frequency ratio, that is the reason why the usual AM radio receiver uses the superheterodyne system.

Example 3.75. The spectrum of a message signal $m(t)$ is shown in figure 3.85(*a*). The ensure communication privacy, this signal is applied to a system (known as a scrambler) shown in figure 3.85(*b*). Analyze the system and sketch the spectrum of the output $x(t)$.

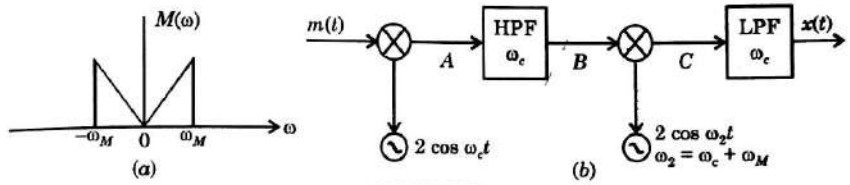


Fig. 3.85.

Solution: The spectrum of the signal at each point is shown in figure 3.86. We see that the spectrum of the output $x(t)$, $X(\omega)$, consists of the two reversed lobes of $M(\omega)$.

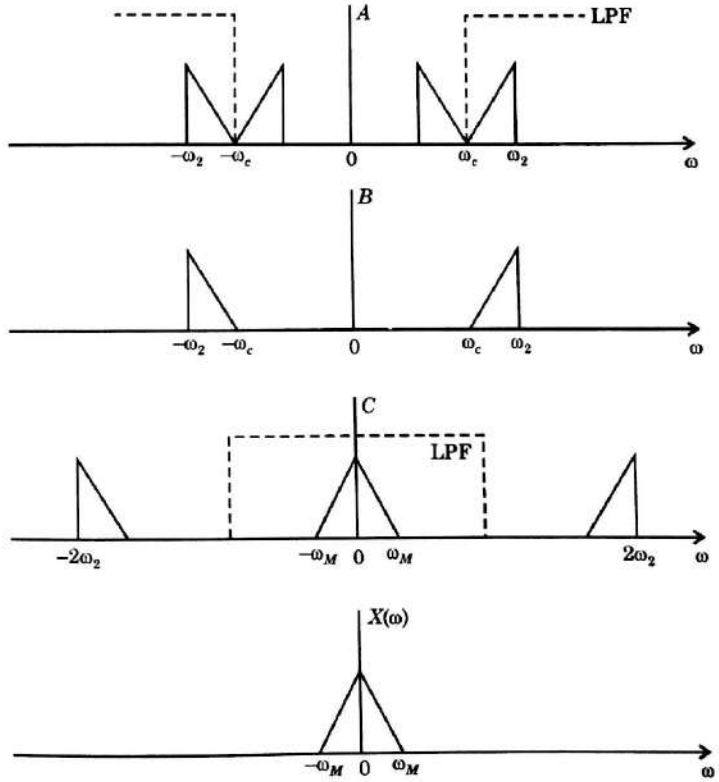


Fig. 3.86

Example 3.76. Using the orthogonality of sine and cosine makes it possible to transmit and receive two different signals simultaneously on the same carrier frequency. A scheme for doing this, known as *quadrature multiplexing*, or *quadrature, amplitude modulation* (QAM),

is shown in figure 3.87. Show that each signal can be recovered by synchronous detection of the received signal by using two local carriers of same frequency but in phase quadrature.
(U.P.S.C., I.E.S., Examination, 1998)

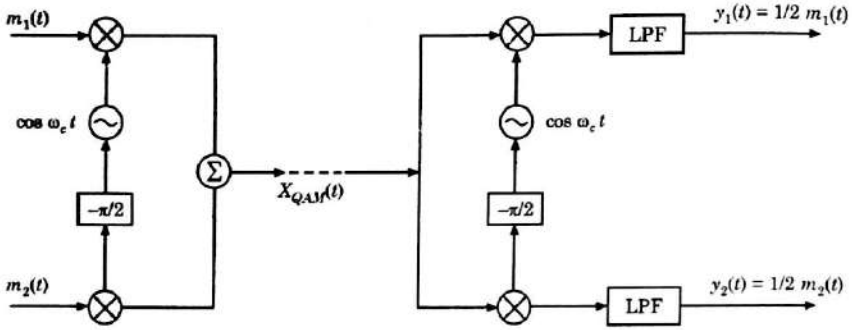


Fig. 3.87. *Quadrature multiplexing system.*

$$x_{QAM}(t) = m_1(t) \cos \omega_c t + m_2(t) \sin \omega_c t$$

$$x_{QAM}(t) \cos \omega_c t = m_1(t) \cos^2 \omega_c t + m_2(t) \sin \omega_c t \cos \omega_c t$$

$$= \frac{1}{2} m_1(t) + \frac{1}{2} m_1(t) \cos 2\omega_c t + \frac{1}{2} m_2(t) \sin 2\omega_c t$$

$$x_{QAM}(t) \sin \omega_c t = m_1(t) \cos \omega_c t \sin \omega_c t + m_2(t) \sin^2 \omega_c t$$

$$= \frac{1}{2} m_1(t) \sin 2\omega_c t + \frac{1}{2} m_2(t) - \frac{1}{2} m_2(t) \cos 2\omega_c t$$

All terms at $2\omega_c$ are filtered out by the low-pass filter, yielding

$$y_1(t) = \frac{1}{2} m_1(t) \quad \text{and} \quad y_2(t) = \frac{1}{2} m_2(t)$$

Note that quadrature multiplexing is an efficient method of transmitting two messages signals within the same bandwidth. It is used in the transmission of colour information signals in commercial television broadcasts.

SUMMARY

1. Modulation is a fundamental requirement of a communication system. Modulation may be defined as a process by which some characteristic of a signal known as carrier is varied according to the instantaneous value of another signal known as modulating signal. The signals containing intelligence or information to be transmitted are called modulating signals.
2. Also the carrier frequency is greater than the modulating frequencies and the signal which results from the process of modulation is known as modulated signal.
3. Modulation may be classified as continuous wave modulation and pulse modulation.
4. If the carrier waveform is continuous in nature then the modulation process is called as continuous-wave (*CW*) modulation. The examples of this type of modulation are amplitude modulation and angle modulation.
5. On the other hand, if the carrier waveform is a pulse-type waveform, then the modulation process is called as pulse modulation. The examples of this type of modulation are Pulse Amplitude Modulation (*PAM*), Pulse Width Modulation (*PWM*), Pulse Code Modulation (*PCM*) etc.
6. Amplitude Modulation and Angle Modulation are the two families of continuous-wave (*CW*) modulation systems.

7. In amplitude modulation, the amplitude of a sinusoidal carrier wave is varied in accordance with the baseband (modulating) signal. On the other hand, in angle modulation, the angle of the sinusoidal carrier wave is varied in accordance with baseband signal.
8. Amplitude Modulation may be defined as a system in which the maximum amplitude of the carrier wave is made proportional to the instantaneous value (amplitude) of the modulating or baseband signal.
9. Carrier signal [*i.e.*, $c(t) = A \cos \omega_c t$] is a fixed frequency signal having frequency ω_c. The modulating or baseband signal $x(t)$ contains the information or intelligence to be transmitted. In the process of amplitude modulation, this information is superimposed upon the carrier signal in the form of amplitude variations of the carrier signal. This means that the information to be transmitted, is now, contained in the amplitude variations of the carrier signal. In other words, in amplitude modulation, the information is transmitted in the form of amplitude variations of the carrier signal.
10. The resulting signal from the process of amplitude-modulation is called amplitude modulated signal or simply AM wave.
11. In the process of amplitude modulation, the frequency and phase of the carrier remain constant whereas the maximum amplitude varies according to the instantaneous value of the information signal.
12. The *AM* wave has a time-varying amplitude called as the envelope of the *AM* wave. Figure 3.1(*c*) shows that the envelope of AM wave consists of the modulating or baseband signal.
13. This means that the unique property of AM is that the envelope of the modulated carrier has the same shape as the message signal or baseband signal.
14. For positive frequencies, a portion of the spectrum of *AM* wave is lying above the carrier frequency w_c. This band of frequency which is lying above the carrier frequency w_c is known as the upper sideband (USB) whereas the symmetrical portion below carrier frequency w_c is known as the lower sideband (LSB). For negative frequencies, the upper sideband (USB) is represented by the portion of the spectrum below-w_c and the lower sideband by the symmetrical portion above $-w_c$.
15. We generally keep $w_c > w_m$ which ensures that the two sidebands do not overlap each other.
16. The negative frequencies appeared in spectrum analysis is due to the exponential representation of Fourier Transform. These negative frequencies are used for mathematical convenience. For a real function, these negative frequencies have no meaning. For general purpose, it is sufficient to consider only the positive frequency region while treating the negative frequency region as a replica of the positive frequencies. This means that to calculate the BW of a signal, we consider only positive side.
17. For positive side, the highest frequency component present in the spectrum of AM wave is $w_c + w_m$ and the lowest frequency component is $w_c - w_m$.
18. The difference between these two extreme frequencies is equal to the bandwidth of the AM wave. Therefore,
    Bandwidth $\quad B = (w_c + w_m) - (w_c - w_m) = 2w_m$
19. The bandwidth of the amplitude modulated wave is twice the highest frequency present in the baseband or modulating signal.
20. In *AM* system the modulation index is defined as the measure of extent of amplitude variation about an unmodulated maximum carrier. It is represented by m_a.
    Mathematically,
    Modulation Index $\quad m_a = \dfrac{|x(t)|_{max}}{\text{maximum carrier amplitude}}$
    or Modulation Index $\quad m_a = \dfrac{|x(t)|_{max}}{A}$.
21. The baseband or modulating signal will be preserved in the envelope of the AM signal only if we have
    $$|x(t)|_{max} \le A$$
    i.e. modulation index m_a is less than or equal to unity.

22. If $m_a > 1$ or the per centage modulation is greater than 100, the baseband signal is not preserved in the envelope. It means that in this case, the baseband signal recovered from the envelope will be distorted. This type of distortion is called envelope distortion and the AM signal with $m_a > 1$ or $m_a > 100\%$ is called overmodulated signal.
23. In which the modulating or baseband signal consists of only one (single) frequency *i.e.* modulation is done by a single frequency or tone. This type of amplitude modulation is known as single tone amplitude modulation.
24. The total power P of the AM wave is the sum of the carrier power P_c and sideband power P_s.
25. The total modulated power of an *AM* signal is expressed as

$$P_t = P_c + P_s = \frac{1}{2}[A^2 + \overline{x^2(t)}]$$

26. In AM wave, the amount of useful message power P_s may be expressed by a term known as transmission efficiency η.
27. Hence transmission efficiency of *AM* wave may be defined as the per centage of total power contributed by the sidebands.
28. Mathematically,

$$\text{Transmission Efficiency,} \quad \eta = \frac{P_s}{P_t} \times 100 = \frac{\frac{1}{2}\overline{x^2(t)}}{\frac{1}{2}A^2 + \frac{1}{2}\overline{x^2(t)}} \times 100 = \frac{100\,\overline{x^2(t)}}{A^2 + \overline{x^2(t)}}$$

29. The maximum transmission efficiency of the AM is only 33.33%. This implies that only one-third of the total power is carried by the sidebands and the rest two-third is wasted.
30. In *AM*, it is generally more convenient to measure the AM transmitter current than the power. In this case, the modulation index may be calculated from the values of unmodulated and modulated currents in the *AM* transmitter.
31. The device which is used to generate an Amplitude-modulated (AM) wave is known as Amplitude Modulator.

    The methods of AM generation may be broadly classified as follow:

    (*i*) Low-level AM Modulation

    (*ii*) High-level AM Modulation
32. Square-low diode modulation and switching modulation are examples of low-level modulation.
33. In a practical modulation circuit, the frequencies other than the carrier and two sideband frequencies are rejected with the help of a tuned circuit.
34. Square law diode modulation circuit make use of nonlinear current-voltage characteristics of diode. This method is suited at low voltage levels because of the fact that current-voltage characteristic of a diode is highly nonlinear particularly in the low voltage region.
35. The process of extracting a modulating or baseband signal from the modulated signal is called demodulation or detection. In other words, demodulation or detection is the process by which the message is recovered from the modulated signal at receiver. The devices used for demodulation or detection are called demodulators or detectors.
36. For amplitude modulation, detectors or demodulators are categorized as:

    (*i*) Square-Law detectors

    (*ii*) Envelope detectors
37. AM signal with large carrier are detected by using the envelope detector. The envelope detector uses the circuit which extracts the envelope of the AM wave. Infact, the envelope of the AM wave is the baseband or modulating signal. But a low-level amplitude modulated signal can only be detected by using square-law detectors in which a device operating in the non-linear region is used to detect the modulating signal.

38. The square law detector circuit is used for detecting modulated signal of small magnitude (*i.e.*, below 1 volt) so that the operating region may be restricted to the non-linear portion of the *V-I* characteristics of the device.
39. It is a known fact that a diode operating in a linear region of its *V-I* characteristics can extract the envelope of an AM wave. This type of detector is known as envelope detector or a linear detector. Envelope detector is most popular in commercial receiver circuits since it is very simple and is not expensive, also at the same time, it gives satisfactory performance for the reception of broadcast programmes.
40. The time constant RC cannot be chosen either too high or too low. If the time constant RC is quite low, the discharge curve during the non-conducting period is almost vertical which results in large fluctuations in a output voltage. Whereas, if the time constant RC is very large, the discharge curve is almost horizontal and it then misses several peaks of the rectified output voltage during negative peaks.
41. One source of distortion in linear diode detector has already been discussed, namely distortion due to improper selection of time constant RC. If RC is too low, removal of radio frequency components is inadequate and the output voltage waveform is spiky in nature. On the other hand, if RC is too large, distortion may result due to clipping during the negative peaks of the modulation wave.
42. The modulated signal, which contains no carrier but two sidebands is called Double-Sideband Suppressed Carrier (DSB-SC) modulation.
43. A balanced modulator may be defined as a circuit in which two non-linear devices are connected in a balanced mode to produce a DSB-SC signal.
44. Ring modulator is another product modulator, which is used to generate DSB-SC signal.
45. The DSB-SC signal may be demodulated by following two methods:

    (*i*) Synchronous detection method

    (*ii*) Using envelope detector after carrier re-insertion.
46. A small amount of carrier signal known as pilot carrier is transmitted alongwith the modulated signal from the transmitter. This small amount of carrier signal is called pilot carrier.
47. Following are the properties of Hilbert transform:

    (*i*) A signal $x(t)$ and its Hilbert transform $x_h(t)$ have the same energy density spectrum.

    (*ii*) A signal $x(t)$ and its Hilbert transform $x_h(t)$ have the same autocorrelation function.

    (*iii*) A signal $x(t)$ and its Hilbert transform $x_h(t)$ are mutually orthogonal.

    $$\text{Mathematically,} \quad \int_{-\infty}^{\infty} x(t)x_h(t)\,dt = 0$$

    (*iv*) If $x_h(t)$ is a Hilbert transform of $x(t)$, then the Hilbert transform of $x_h(t)$ is $-x(t)$, *i.e.*,

    If $\quad H[x(t)] = x_h(t) \quad$ then $\quad H[x_h(t)] = -x(t)$

    Here 'H' denotes the Hilbert transform.
48. SSB-SC signals may be generated by two methods as under:

    (*i*) Frequency discrimination method or filter method

    (*ii*) Phase discrimination method or phase-shift method
49. The phase-shift method avoids filter. This method makes use of two balanced modulators and two phase-shifting networks.
50. The demodulation or detection of AM signal is simpler than DSB-SC and SSB systems. The conventional AM can be demodulated by rectifier or envelope detector. Detection of DSB-SC and SSB is rather difficult and expensive alsop. Furthermore, it is quite easier to generate conventional AM signals at high power levels as compared to DSB-SC and SSB signals. For this reason, conventional AM systems are used for broadcasting purpose.

51. The advantage of DSB-SC and SSB systems over conventional AM system is that the former requires lesser power to transmit the same information.
52. SSB scheme needs only one-half of the bandwidth required in DSB-SC system and less than that required in VSB also. Thus, we can say that SSB modulation scheme is the most efficient scheme among DSB-SC and VSB schemes. SSB modulation scheme is used for long distance transmission of voice signals because it allows longer spacing between repeaters.
53. A device that performs the frequency translation of a modulated signal is known as a *frequency mixer*. The operation is often called *frequency mixing, frequency conversion*, or *heterodyning*.
54. Multiplexing is a technique in which several message signals are combined into a composite signal for transmission over a common channel. In order to transmit a number of these signals over the same channel, the signals must be kept apart so that they do not interfere with each other, and hence they can be separated easily at the receiver end.

REVIEW QUESTIONS

1. What is meant by the term amplitude modulation.
2. Define the term modulation index for AM.
3. Derive an expression for single-tone amplitude-modulated wave.
4. Draw the single-sided frequency spectrum for a single tone amplitude modulated signal.
5. Derive the power relations for single-tone amplitude-modulated wave.
6. Derive the current relations for single-tone amplitude modulated wave.
7. Derive an expression for the transmission efficiency of AM wave.
8. Explain the low-level and high-level AM modulation methods with help of figures.
9. Explain the square-law diode modulation method for AM generation.

SHORT QUESTIONS WITH ANSWERS

Q.1. Write the expression for AM wave?

Ans. The standard equation for amplitude modulated (AM) wave may be expressed as

$$s(t) = x(t)\cos\omega_c t + A\cos\omega_c t$$

or

$$s(t) = [A + x(t)]\cos\omega_c t$$

Q.2. What is the envelope of AM wave?

Ans. The expression for *AM* wave is

$$s(t) = [A + x(t)]\cos\omega_c t$$

or

$$s(t) = E(t)\cos\omega_c t$$

where

$$E(t) = A + x(t)$$

$E(t)$ is called the envelope of *AM* wave. This envelope consists of the baseband signal $x(t)$. Hence, the modulating or baseband signal may be recovered from an AM wave by detecting the envelope.

Q.3. What is the BW for AM wave?

Ans. The difference between these two extreme frequencies is equal to the bandwidth of the AM wave. Therefore,

Bandwidth

$$B = (w_c + w_m) - (w_c - w_m)$$

$$B = 2w_m$$

Q.4. Define modulation index for AM wave in AM system?

Ans. The modulation index is defined as the measure of extent of amplitude variation about an unmodulated maximum carrier. It is represented by m_a.

Mathematically,

Modulation Index $$m_a = \frac{|x(t)|_{max}}{\text{maximum carrier amplitude}}$$

or Modulation Index $$m_a = \frac{|x(t)|_{max}}{A}$$

where $|x(t)|_{max}$ represents the maximum amplitude of modulating signal and A represents the maximum amplitude of carrier signal.

The modulation index is also known as depth of modulation, degree of modulation or modulation factor.

Also, the absolute value of m_a multiplied by 100 is known as per centage modulation.

Q.5. What are the frequency components in an AM wave?

Ans. The AM signal has three frequency components as follow:

(*i*) carrier frequency ω_c having amplitude A

(*ii*) upper sideband $(\omega_c + \omega_m)$ having amplitude $\frac{m_a \cdot A}{2}$

(*iii*) lower sideband $(\omega_c - \omega_m)$ having amplitude $\frac{m_a \cdot A}{2}$

With the help of these frequency components, we can plot the frequency-spectrum of single-tone amplittude modulated (AM) wave. Figure 3.88 shows the one-sided frequency spectrum of single-tone AM wave.

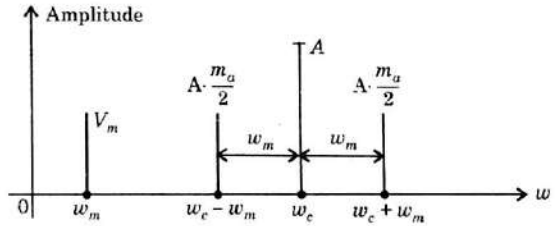


Fig. 3.88. *Single-sided frequency spectrum of single-tone AM wave.*

Q.6. Define transmission efficiency in AM wave?

Ans. In AM wave, the amount of useful message power P_s may be expressed by a term known as transmission efficiency η.

Q.7. What is the current relationship for AM wave?

Ans.

$$I_t = I_c\sqrt{1 + \frac{m_a^2}{2}}$$

Q.8. What is demodulation?

Ans. The process of extracting a modulating or baseband signal from the modulated signal is called demodulation or detection. In other words, demodulation or detection is the process by which the message is recovered from the modulated signal at receiver. The devices used for demodulation or detection are called demodulators or detectors. For amplitude modulation, detectors or demodulators are categorized as:

(*i*) Square-Law detectors, (*ii*) Envelope detectors.

Q.9. How can you obtain a DSB-SC signal?

Ans. A DSB-SC signal is obtained by simply multiplying modulating signal $x(t)$ with carrier signal cos $w_c t$. This is achieved by a product modulator. The block diagram of a product modulator is shown in figure 3.89.

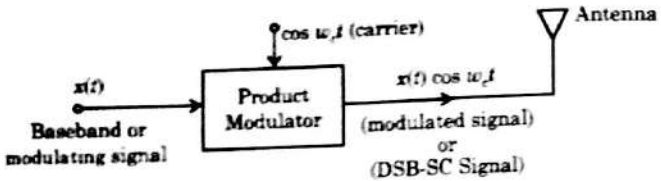


Fig. 3.89.

Q.10. What is the BW of DSB-SC signal

Ans. Bandwidth $B = (w_c + w_m) - (w_c - w_m)$

$= 2\,w_m$

It is obvious that the bandwidth of DSB-SC modulation is same as that of general AM wave.

Q.11. What are the demodulation methods for DSB-SC signal.

Ans. The DSB-SC signal may be demodulated by following two methods:

(*i*) Synchronous detection method

(*ii*) Using envelope detector after carrier reinsertion.

Q.12. What do you mean by Hilbert transfer and inverse Hilbert transform.

(Anna University, Chennai-1999)

Ans. It may be observed that the function $x_h(t)$ obtained by providing $(-\pi/2)$ phase shift to every frequency component present in $x(t)$, actually represents the Hilbert transform of $x(t)$. This means that $x_h(t)$ is the Hilbert transform of $x(t)$ defined as

$$x_h(t) = \frac{1}{\pi} x(t) \otimes \frac{1}{t} = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x(\tau)}{t-\tau}\, d\tau$$

Also, the inverse Hilbert transform is defined as

$$x(t) = -\frac{1}{\pi} \int_{-\infty}^{\infty} \frac{x_h(\tau)}{t-\tau}\, d\tau$$

Q.13. Write few applications of Hilbert transform.

Ans. (*i*) for generation of SSB signals,

(*ii*) for designing of minimum phase type filters,

(*iii*) for representation of bandpass signals.

Q.14. Write few properties of Hilbert transform.

Ans. Following are the properties of Hilbert transform:

(*i*) A signal $x(t)$ and its Hilbert transform $x_h(t)$ have the same energy density spectrum.

(*ii*) A signal $x(t)$ and its Hilbert transform $x_h(t)$ have the same autocorrelation function.

(*iii*) A signal $x(t)$ and its Hilbert transform $x_h(t)$ are mutually orthogonal.

Mathematically,

$$\int_{-\infty}^{\infty} x(t) x_h(t)\, dt = 0$$

(*iv*) If $x_h(t)$ is a Hilbert transform of $x(t)$, then the Hilbert transform of $x_h(t)$ is $-x(t)$, *i.e.*,

If $\quad H[x(t)] = x_h(t)$

then $\quad H[x_h(t)] = -x(t)$

Here 'H' denotes the Hilbert transform.

Q.15. What are the generating methods for SSB-SC signal.

Ans. SSB-SC signals may be generated by two methods as under:

(*i*) Frequency discrimination method or filter method

(*ii*) Phase discrimination method or phase-shift method

Q.16. What do you mean by multiplexing? *(Anna University, Chennai-2000)*

Ans. Multiplexing is a technique in which several message signals are combined into a composite signal for transmission over a common channel. In order to transmit a number of these signals over the same channel, the signals must be kept apart so that they do not interfere with each other, and hence they can be separated easily at the receiver end.

NUMERICAL PROBLEMS

1. An audio signal whose mathematical description is 25 sin (2π 1000t) modulates a carrier described as 75 sin (2π 150,000 t)
   (*a*) Sketch the audio signal.
   (*b*) Sketch the carrier.
   (*c*) Construct the modulated wave showing all amplitude magnitudes.
   (*d*) Calculate the modulation factor and per cent modulation.
   (*e*) What is the frequency of the audio signal of the carrier?
   (*f*) What frequencies would show up in a spectrum analysis of the modulated wave?
   (JNTU, Hyderabad-1998)
   [**Ans.** (*d*) 0.333, 33.33%, (*e*) 1000 Hz, 150,000 Hz, (*f*) 149,000 Hz, 150,000 Hz, 151,000 Hz]

2. An audio signal described as 30 sin (2π 2500 t) amplitude modulates a carrier which is described as 65 sin (2π 250, 000 t)
   (*a*) Sketch the audio signal
   (*b*) Sketch the carrier
   (*c*) Construct the modulated wave
   (*d*) What is the modulation factor and per cent modulation?
   (*e*) What is frequency of the audio signal of the signal?
   (*f*) What frequencies would show up in a spectrum analysis of the modulated wave?
   [**Ans.** (*d*) 0.4615, 46.15%, (*e*) 2500 Hz, 250,000 Hz, (*f*) 247, 500 Hz, 252,500 Hz]

3. A 2000 Hz audio signal having an amplitude of 15 V modulates a 100 kHz carrier which has a peak value of 25 V when not modulated.
   (*a*) Sketch the audio signal to scale.
   (*b*) Sketch the carrier to scale.
   (*c*) Construct the modulation factor and per cent modulation of the modulated wave.
   (*d*) Calculate the modulation factor and per cent modulation of the modulated wave.
   (*e*) What frequencies appear in a spectrum analysis of the modulated wave? *(Pune University-1999)*
   [**Ans.** (*d*) 0.60, 60% ; (*e*) 98 kHz, 100 kHz, 102 kHz]

4. An 1800 Hz signal which has an amplitude of 30 V amplitude modulates a 50 MHz carrier which when the unmodulated has an amplitude of 65 V.
   (*a*) Sketch the modulating signal.
   (*b*) Sketch the carrier.
   (*c*) Construct the modulated wave.
   (*d*) Calculate the modulation factor and per cent modulation.

(*e*) What frequencies would show up in a spectrum analysis of the AM wave?

(*f*) Write the trigonometric equations for the carrier and for the audio signal.

[**Ans.** (*d*) 0.4614, 46.15% ; (*e*) 49.9982 MHz, 50.0000 MHz, 50.0018 MHz ; (*f*) 65 sin [2π (50×10^6)t]

5. How many AM broadcast station can be accommodated in a 6 MHz bandwidth if each station transmits a signal which was modulated by an audio signal having a maximum frequency of 5 kHz? [**Ans.** 600 stations]

6. A bandwidth of 12 MHz becomes available for assignment. If assigned for TV broadcast service, only two channels could be accommodated. Determine the number of AM stations that could broadcast simultaneously if the maximum modulating frequency is limited to 5 kHz. [**Ans.** 1200 stations]

7. A 90 kHz bandwidth is to accommodate six AM broadcasts simultaneously. What maximum modulating frequency must each station be limited to? [**Ans.** 7500 Hz]

8. An antenna transmits an AM signal having a total power content of 15 kW. Determine the power being transmitted at the carrier frequency and at each of the sidebands when the per cent modulation if 85%. [**Ans.** 11,019 W, 1990 W]

9. Calculate the power content of the carrier and of each of the sidebands of an AM signal whose total broadcast power is 40 kW when the per cent modulation is 60%. [**Ans.** 33,898, 3050 W]

10. A 3500 Hz audio tone amplitude modulates a 200 kHz carrier resulting in a modulated signal having a per cent modulation of 85%. The total power being transmitted is 15 kW.

    (*a*) What frequencies would appear in a spectrum analysis of the modulated wave?

    (*b*) Determine the power content at each of the frequencies that appear in a spectrum analysis of the modulated wave. [**Ans.** (*a*) 196,500 Hz, 200,000 Hz, 203,500 Hz; (*b*) 11,019 W, 1990 W]

11. Determine the power contained at the carrier frequency and within each of the sidebands for an AM signal whose total power content is 15 kW when the modulation factor is 0.70.

[**Ans.** 12,048, 1475 W]

12. An amplitude-modulated signal contains a total of 6 kW. Calculate the power being transmitted at the carrier frequency and at each of the sidebands when the per cent modulation is 100%.

[**Ans.** 4000 W, 1000 W]

13. An AM wave has a power content of 1800 W at its carrier frequency. What is the power content of each of the sidebands when the carrier is modulated 85% ? [**Ans.** 325 W]

14. An AM signal contains 500 W at its carrier frequency and 100 W in each of its sidebands.

    (*a*) Determine the per cent modulation of the AM signal.

    (*b*) Find the allocation of power if the per cent modulation is changed to 60%.

    Note: The power content of the carrier of an AM wave does not vary with per cent modulation.

[**Ans.** (a) 89.44% (b) 500 W, 45 W, 45 W]

15. 1200 W is contained at the carrier frequency of an AM signal. Determine the power content of each of the sidebands for each of the following per cent modulations: (*a*) 40%, (*b*) 50%, (*c*) 75%, (*d*) 100%.

    Note: The power content of the carrier of an AM wave does not vary with per cent modulation.

[**Ans.** (*a*) 89.44%, (*b*) 75 W, (*c*) 168.75 W, (*d*) 300 W]

@sTc

CHAPTER 4

Angle Modulation

Inside this Chapter

4.1. Introduction 229
4.2. Concept of Angle Modulation Basic Definition 223
4.3. Frequency Deviation 232
4.4. Relationship Between Phase Modulation and Frequency Modulation (FM) 233
4.5. Single Tone Frequency Modulation 234
4.6. Phasor Representation of Angle Modulated (*i.e.*, FM and PM Waves) 239
4.7. Types of Frequency Modulation (FM) 239
4.8. Wideband FM 242
4.9. Effect of Variation of Modulation Index m_f on the Spectrum of FM Signal 246
4.10. Transmission Bandwidth of FM Signal 247
4.11. FM bandwidth for an Arbitrary Modulating Signal $x(t)$ 249
4.12. Narrowband FM Versus Wideband FM 249
4.13. Multiple Frequency Modulation 251
4.14. Square Wave Modulation 252
4.15. Linear and Nonliner Modulation 254
4.16. Phase modulation : An Analytic View 255
4.17. Comparison of Angle Modulated Waves and Amplitude Modulated Wave 257
4.18. Comparison of Frequency Modulation and Amplitude Modulation 258
4.19. FM generation 258
4.20. Drawbacks of Direct Method for FM Generation 260
4.21. The Indirect or Armstrong Method of FM Generation 260
4.22. Practical Armstrong Method for FM Generation 260
4.23. FM Demodulators or Detectors 262
4.24. Types of FM Demodulators 262
4.25. Phase Difference Detectors 264

4.1. Introduction

In last chapter, we discussed the Amplitude (linear) Modulation in which we varied the amplitude of the carrier according to the message signal. Also the transmission bandwidth of AM signal never exceeds twice the maximum frequency component present in AM signal. Linear or amplitude modulation has another important property that the signal-to-noise ratio (SNR) at the destination is no better than one which does not involve modulation (*i.e.*, baseband modulation). The SNR can

be improved only be increasing the transmitted power. There is another significant method of modulation used for message transmission known as *angle modulation*. In angle modulation, the frequency or phase of the carrier signal is varied according to the message signal. In this method of modulation, the amplitude of the carrier is kept constant.

4.2. Concept of Angle Modulation Basic Definition

Angle modulation may be defined as the process in which the total phase angle of a carrier wave is varied in accordance with the instantaneous value of the modulating or message signal while keeping the amplitude of the carrier constant.

Mathematical Representation

Let us consider that an unmodulated carrier signal is expressed as

$$c(t) = A\cos(\omega_c t + \theta_0) \quad ...(4.1)$$

where A = amplitude of carrier

ω_c = carrier frequency

θ_0 = some phase angle

Substituting $\omega_c t + \theta_0 = \phi$, we get

$$c(t) = A\cos\phi \quad ...(4.2)$$

where $$\phi = (\omega_c t + \theta_0) \quad ...(4.3)$$

is the total phase angle of the carrier signal.

We may consider equation (4.2) as the real part of rotating phasor $A\,e^{j\phi}$. Let us denote this rotating phasor by C, so that

$$C = A\,e^{j\phi} \quad ...(4.4)$$

Also $$c(t) = Re\,(A\,e^{j\phi}) = A \cdot Re\,(e^{j\phi}) = A\cos\phi \quad ...(4.5)$$

This phasor C rotates at a constant angular velocity ω_c. Clearly θ_0 is the phase angle of the unmodulated carrier at $t = 0$.

From equation (4.3), the constant angular velocity ω_c of the phasor C is related to its total phase angle 'ϕ' as

$$\phi = \omega_c t + \theta_0$$

Differentiating both sides of this equation with respect to t, we have

$$\omega_c = \frac{d\phi}{dt} \quad ...(4.6)$$

It may be noted at this point that for unmodulated carrier, the derivative $\frac{d\phi}{dt}$ is constant. However, this derivative $\frac{d\phi}{dt}$ may not be a constant with time, in general *i.e.* this may vary with time. Therefore, the angular velocity of the phasor C would also vary with time. Hence, this time-dependent angular velocity or angular frequency is known as **instantaneous angular velocity** or **instantaneous angular frequency**. This instantaneous angular frequency is denoted by w_i.

Hence, for this case, equation (4.6) becomes

$$\frac{d\phi}{dt} = \omega_i \quad ...(4.7)$$

where angular frequency ω_i is time-dependent.

From equation (4.7), we get

$$\phi = \int \omega_i\, dt \quad ...(4.8)$$

Again writing equation (4.1) of unmodulated carrier as

$$c(t) = A\cos(\omega_c t + \theta_0) = A\cos\phi$$

where ϕ is the total phase angle of the unmodulated carrier expressed as

$$\phi = \omega_c t + \theta_0$$

Now, if this angle ϕ is varied according to the instantaneous value of the message or modulating signal, the carrier signal is then said to be angle modulated.

Types of Angle Modulation

We can vary this phase angle ϕ in two ways and thus there are two types of angle modulation as under:

(*i*) Phase Modulation (PM)

(*ii*) Frequency Modulation (FM)

4.2.1. Phase Modulation (PM)

Definition

Phase modulation (PM) is that type of angle modulation in which the phase angle ϕ is varied linearly with a baseband or modulating signal $x(t)$ about an unmodulated phase angle $(\omega_c t + \theta_0)$. This means that in Phase Modulation, the instantaneous value of the phase angle is equal to the phase angle of the unmodulated carrier $(\omega_c t + \theta_0)$ plus a time-varying component which is proportional to modulating signal $x(t)$.

Mathematical Representation

We know that unmodulated carrier signal is expressed as

$$c(t) = A\cos(\omega_c t + \theta_0) = A\cos\phi$$

where $$\phi = \omega_c t + \theta_0$$

Neglecting θ_0, we get

total phase angle of unmodulated carrier is

$$\phi = \omega_c t \quad ...(4.9)$$

Now, according to Phase Modulation, this phase angle 'ϕ' is varied linearly with a baseband or modulating signal $x(t)$.

Let the instantaneous value of phase angle be denoted by ϕ_i.

Therefore,

$$\phi_i = \omega_c t + k_p \cdot x(t) \quad ...(4.10)$$

where k_p is the proportionality constant and is known as **phase sensitivity** of the modulator. This is expressed in radians/volts.

Since, the expression for unmodulated carrier wave is

$$c(t) = A\cos\phi$$

Therefore, the expression for phase modulated wave will be

$$s(t) = A\cos\phi_i \quad ...(4.11)$$

Putting the value of ϕ_i in equation (4.11) from equation (4.10), we get

$$s(t) = A\cos[\omega_c t + k_p \cdot x(t)] \quad ...(4.12)$$

which is the required mathematical expression for a phase modulated wave.

4.2.2. Frequency Modulation

Definition *(U.P. Tech., Semester, Exam., 2004-05) (4 marks)*

Frequency modulation is that type of angle modulation in which the instantaneous frequency ω_i is varied linearly with a message or baseband signal $x(t)$ about an unmodulated carrier frequency ω_c. This means that the instantaneous value of the angular frequency ω_i will be equal to the carrier frequency ω_c plus a time-varying component proportional to the baseband signal $x(t)$.

Mathematical, Representation

We know that the instantaneous frequency is given by

$$\omega_i = \omega_c + k_f \cdot x(t) \quad ...(4.13)$$

where k_f is proportionality constant and is known as the **frequency sensitivity** of the modulator. This is expressed in Hz/volt.

Now, let the expression for unmodulated carrier signal be

$$c(t) = A\cos(\omega_c t + \theta_0) \quad ...(4.14)$$

or

$$c(t) = A\cos\phi \quad ...(4.15)$$

where

$$\phi = \omega_c t + \theta_0 \quad ...(4.16)$$

'ϕ' is the total phase angle of the unmodulated carrier.

Let ϕ_i be the instantaneous phase angle of the modulated signal.

From equation (4.15), the equation for unmodulated carrier is

$$c(t) = A\cos\phi$$

On frequency modulation amplitude A must remain constant and only angle 'ϕ' will change.

Hence, the expression for frequency modulated wave will be

$$s(t) = A\cos\phi_i \quad ...(4.17)$$

where ϕ_i = instantaneous phase angle

From equation (4.16)

$$\phi = \omega_c t + \theta_0$$

On differentiation, we get

$$\frac{d\phi}{dt} = \omega_c$$

or

$$\phi = \int \omega_c dt \quad ...(4.18)$$

Based on equation (4.18), we may write the expression for instantaneous phase angle ϕ_i as

$$\phi_i = \int \omega_i\, dt \quad ...(4.19)$$

where ω_i = instantaneous frequency of frequency modulated wave.

Putting the value of w_i in equation (4.19) form equation (4.13), we get

$$\phi_i = \int [\omega_c + k_f x(t)]\, dt = \omega_c t + k_f \int x(t)\, dt \quad ...(4.20)$$

Putting this value of ϕ_i in equation (4.17), we get the expression for frequency modulated wave will be

$$s(t) = A\cos\left[\omega_c t + k_f \int x(t)\, dt\right] \quad ...(4.21)$$

Now, if the phase angle of the unmodulated carrier is taken at $t = 0$, then the limit of integration in equation (4.21) will be 0 to t.

In this case the expression for FM wave will be

$$s(t) = A\cos\left[\omega_c t + k_f \int_0^t x(t)\, dt\right]$$

which is the required **general expression for FM wave.**

4.3. Frequency Deviation

(U.P. Tech., Semester, Exam., 2003-04) (05 marks)

We know that the instantaneous frequency of FM wave is given as

$$\omega_i = \omega_c t + k_f \cdot x(t) \quad ...(4.22)$$

The instantaneous frequency of FM signal varies with time around the carrier frequency ω_c. This means that the instantaneous frequency of FM signal varies according to the modulating signal.

The maximum change in instantaneous frequency form the average frequency ω_c is called **frequency deviation.***

* The frequency deviation is a useful parameter for determining the bandwidth of FM signals.

This maximum change in instantaneous frequency ω_i from the average or carrier frequency ω_c depends upon the magnitude and sign of $k_f \cdot x(t)$. This means that the frequency deviation would be either positive or negative depending upon the sign of $k_f \cdot x(t)$. However, the amount of frequency deviation in both these cases is given by the maximum magnitude *i.e.* $|k_f \cdot x(t)|_{max}$.

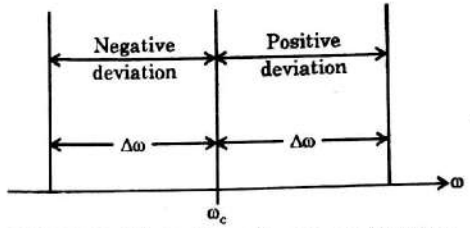


Fig. 4.1. *Illustration of frequency deviation*

Maximum frequency deviation is generally denoted by $\Delta\omega$.

Therefore,

Frequency deviation $\Delta\omega = |k_f \cdot x(t)|_{max}$

Figure 4.1 further illustrates the concept of frequency deviation.

4.4. Relationship between Phase Modulation (PM) and Frequency Modulation (FM)

(U.P. Tech., Semester, Exam., 2004-05) (06 marks)

We know that an angle modulated wave is given as

$$s(t) = A\cos\phi_i \quad ...(4.23)$$

where A = Amplitude

ϕ_i = Instantaneous total phase angle of the angle modulated wave.

The expression for phase modulated (PM) wave is

$$s(t) = A\cos[\omega_c t + k_p x(t)] \quad ...(4.24)$$

Similarly, the expression for frequency modulated (FM) wave is

$$s(t) = A\left[\cos\omega_c t + k_f \int_0^t x(t)\, dt\right] \quad ...(4.25)$$

It may be observed from above equations that Phase Modulation (PM) and Frequency Modulation (FM) are closely related to each other because in both the cases there is a variation in the total phase angle. In Phase Modulation (PM), the phase angle varies linearly with baseband signal $x(t)$ whereas in case of Frequency Modulation (FM), the phase angle varies linearly with the integral of baseband signal $x(t)$. This means that FM wave may be obtained by using PM. Conversely, PM wave may be obtained by using FM.

To get FM by using PM, we first integrate the baseband signal and then apply to the phase modulator. This process is illustrated with the help of a block diagram shown in figure 4.2.

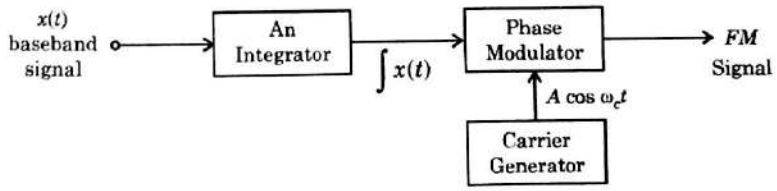


Fig. 4.2. *Generation of FM using phase modulator.*

Similarly, PM wave may be generated by using frequency modulator by first differentiating modulating or baseband signal $x(t)$ and then applying to the frequency modulator. This process is illustrated with the help of a block diagram shown in figure 4.3.

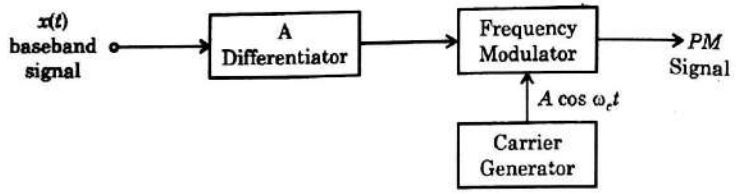


Fig. 4.3. *Generation of PM using frequency modulator.*

4.5. Single Tone Frequency Modulation

Definition

We know that the general expression for FM wave is given as

$$s(t) = A\cos\left[\omega_c t + k_f \int_0^t x(t)\,dt\right] \quad \text{...(4.26)}$$

From this equation, it is clear that the FM wave $s(t)$ is a nonlinear function of the baseband or modulating signal $x(t)$. This means that frequency modulation is a **nonlinear modulation process**. Due to this reason unlike amplitude modulation, the spectrum of an FM wave is related in a complex manner with the modulating or baseband signal.

Therefore, to study the spectrum or spectral properties of FM signal, we shall consider the simplest case of FM wave known as **single tone frequency modulation**.

In general expression of FM wave equation (4.26), we consider a modulating or baseband signal $x(t)$ which may consist of any number of frequency components.

On the other hand, a single tone frequency modulation is that type of frequency modulation (*FM*) in which the modulating or baseband signal contains a single frequency.

Let us consider a carrier signal as

$$c(t) = A\cos\omega_c t$$

Let the modulating signal be

$$x(t) = V_m \cos\omega_m t^*$$

where V_m = maximum amplitude of modulating signal

w_m = frequency of modulating signal

Then according to the definition of *FM*, the frequency of carrier signal is varied according to the instantaneous value of the modulating or baseband signal, $x(t)$. Figure 4.4 illustrates the process of single tone frequency modulation.

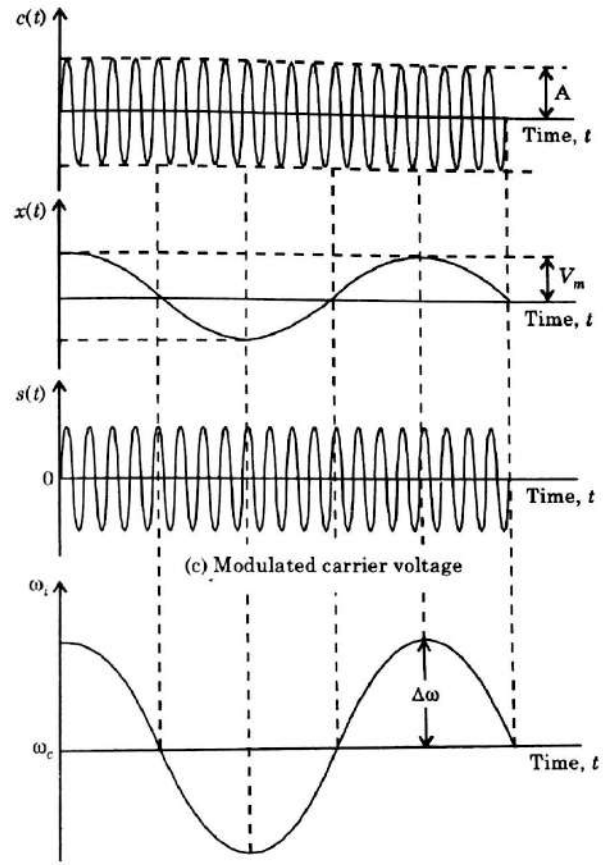
Few Points

(*i*) Figure 4.4 shows the *a* single-tone FM wave. It is clear that the frequency of carrier is varied according to the modulating signal keeping the maximum amplitude of the carrier as constant.

(*ii*) Figure 4.4 shows the variation of instantaneous carrier frequency ω_i with time around the unmodulated carrier frequency or centre frequency ω_c. This variation of instantaneous carrier frequency is proportional to the modulating signal. This means that the instantaneous carrier frequency varies according to the shape of modulating or baseband signal.

(*iii*) From figure 4.4 it may also be observed that when the modulating or message signal is applied, the instantaneous carrier frequency deviates up and down from its centre frequency ω_c. This change or shift either above or below the resting or centre frequency is called frequency deviation denoted by $\Delta\omega$.

* Note that here we have considered the modulating signal $x(t)$ as a voltage signal $v(t)$.



(*d*) Variation of carrier frequency

Fig. 4.4. *Illustration of FM (a) Unmodulated carrier (b) Modulating signal (c) FM signal (d) Instantaneous carrier frequency.*

(*iv*) The total variation in frequency from the lowest to the highest point is called carrier swing. Obviously

The carrier swing = 2 × frequency deviation = $2 \times \Delta\omega$

(*v*) The amount of frequency deviation or variation depends upon the amplitude (loudness) of the modulating (audio) signal. This means that louder the sound, greater the frequency deviation and vice versa.

However, in FM broadcast, it has been internationally agreed to restrict maximum deviation to 75 kHz on each side of the centre frequency for sounds of maximum loudness. This means that the sounds of lesser loudness are permitted proportionally less frequency deviation. A maximum frequency deviation of 25 kHz is permitted for the sound portion of television broadcasts.

(*vi*) As stated earlier, the frequency deviation is useful in determining the FM signal bandwidth. Since, a maximum frequency deviation of 75 kHz is allowed for commercial FM broadcast stations using a band of 88 MHz to 108 MHz, therefore approximate FM channel width is $2 \times 75 = 150$ kHz. Allowing a 25 kHz guard band on either side, the channel width becomes

2(75 + 25) = 200 kHz. This guardband is meant to prevent interference between adjacent channels.

In FM broadcast, the highest audio frequency transmitted is 15 kHz.

(*vii*) For FM, the modulation index is defined as the ratio of frequency deviation to the modulating frequency.

Mathematically,

$$\text{Modulation index, } m_f = \frac{\text{Frequency deviation}}{\text{Modulation frequency}} = \frac{\Delta\omega}{\omega_m} \quad ...(4.27)$$

This modulation index may be greater than unity.

(*vii*) The term "per cent modulation" as it is used in reference to FM refers to the ratio of actual frequency deviation to the maximum allowable frequency deviation. Thus 100% modulation corresponds to 75 kHz for the commercial FM broadcast band and 25 kHz for the commercial FM broadcast band and 25 kHz for television.

$$\text{Per cent modulation } M = \frac{\Delta f_{actual}}{\Delta f_{max}} \times 100$$

4.5.1. Mathematical Expression for Single-Tone Frequency Modulation

Let the expression for a carrier signal be

$$c(t) = A \cos \omega_c t$$

Let the modulating signal be

$$x(t) = V_m \cos \omega_m t$$

Let the expression for FM wave be

$$s(t) = A \cos \phi_i \quad ...(4.28)$$

where ϕ_i is the instantaneous phase angle of the modulated wave.

We know that the instantaneous frequency of the modulated signal is given as

$$\omega_i = \omega_c + k_f \cdot x(t) \quad ...(4.29)$$

Putting the value of $x(t)$, we get

$$\omega_i = \omega_c + k_f \cdot V_m \cos \omega_m t$$

But we know that frequency deviation is given as

$$\Delta\omega = |k_f \cdot x(t)|_{max} = k_f |x(t)|_{max} = k_f \cdot V_m$$

Therefore,

$$\omega_i = \omega_c + \Delta\omega . \cos \omega_m t \quad ...(4.30)$$

The total phase angle ϕ_i of the modulated wave is given as

$$\phi_i = \int \omega_i \, dt$$

Putting the value of ω_i from equation (4.30), we get

$$\phi_i = \int [\omega_c + \Delta\omega \cos \omega_m t] \, dt = \omega_c t + \frac{\Delta\omega}{\omega_m} \sin \omega_m t \quad ...(4.31)$$

But, modulation index m_f is given as

$$m_f = \frac{\text{Frequency deviation}}{\text{Modulating frequency}} = \frac{\Delta\omega}{\omega_m}$$

Therefore, putting the value of m_f in equation (4.31), we obtain

$$\phi_i = \omega_c t + m_f \sin \omega_m t$$

Substituting this value of ϕ_i in equation (4.28), we get the expression for single-tone FM wave

$$s(t) = A \cos \phi_i = A \cos [\omega_c t + m_f \sin \omega_m t \quad ...(4.32)$$

which is the required mathematical expression for single tone FM wave.

Example 4.1. A single-tone FM is represented by the voltage equation as:

$$v(t) = 12 \cos (6 \times 10^8 t + 5 \sin 1250 t)$$

Determine the following:

(*i*) **carrier frequency**

(*ii*) **modulating frequency**

(*iii*) **the modulation index**

(*iv*) **maximum deviation**

(*v*) **what power will this FM wave dissipate in 10 Ω resistors.**

Solution: We know that the standard expression for a single-tone FM wave is given as

$$v(t) = A \cos (w_c t + m_f \sin w_m t) \quad ...(i)$$

The given expression is

$$v(t) = 12 \cos (6 \times 10^8 t + 5 \sin 1250 t) \quad ...(ii)$$

Comparing equations (*i*) and (*ii*), we get

(*i*) carrier frequency

$$w_c = 6 \times 10^8 \text{ rad/sec.}$$

or

$$f_c = \frac{6 \times 10^8}{2\pi} = 95.5 \text{ MHz} \quad \textbf{Ans.}$$

(*ii*) modulating frequency

$$w_m = 1250 \text{ rad/sec.}$$

or

$$f_m = \frac{1250}{2\pi} = 199 \text{ Hz} \quad \textbf{Ans.}$$

(*iii*)

$$m_f = 5 \quad \textbf{Ans.}$$

(*iv*) maximum frequency deviation is given as

$$m_f = \frac{\Delta w}{w_m} = \frac{\Delta_f}{f_m}$$

or

$$\Delta_f = m_f \cdot f_m$$

$$= 5 \times 199 = 995 \text{ Hz} \quad \textbf{Ans.}$$

(*v*) The power dissipated is

$$P = \frac{V_{rms}^2}{R} = \frac{(12/\sqrt{2})^2}{10}$$

or

$$P = \frac{72}{10} = 7.2 \text{ watts} \quad \textbf{Ans.}$$

Example 4.2. A 107. 6 MHz carrier signal is frequency modulated by a 7 kHz since wave. The resultant FM signal has a frequency deviation of 50 kHz. Determine the following:

(*i*) **the carrier swing of the FM signal.**

(*ii*) **the highest and the lowest frequencies attained by the modulated signal.**

(*iii*) **the modulation index of the FM wave.**

Solution: Given that

$$f_c = 107.6 \text{ MHz}$$

$$f_m = 7 \text{ kHz}$$

$$\Delta_f = 50 \text{ kHz}$$

(*i*) Carrier swing = 2 × frequency deviation

$$= 2 \times 50 = 100 \text{ kHz} \quad \textbf{Ans.}$$

(*ii*) The highest frequency attained by the modulated signal is equal to the resting or carrier frequency plus the frequency deviation, *i.e.*

$$f_H = f_c + \Delta_f = 107.6 \times 10^6 + 50 \times 10^3$$
$$= (107600 \times 10^3) + (50 \times 10^3)$$

or $$f_H = 107650 \times 10^3 \text{ Hz} = 107.65 \text{ MHz} \quad \textbf{Ans.}$$

Similarly the lowest frequency attained by the modulated signal is equal to the resting or carrier frequency minus the frequency deviation i.e.

$$f_L = f_c - \Delta_f = 107.6 \times 10^6 - 50 \times 10^3)$$
$$= (107600 \times 10^3) - (50 \times 10^3)$$

or $$f_L = 107550 \times 10^3 \text{ Hz} = 107.55 \text{ MHz} \quad \textbf{Ans.}$$

(iii) The modulation index is given as

$$m_f = \frac{\text{frequency deviation}}{\text{modulating frequency}} = \frac{\Delta_f}{f_m}$$

or $$m_f = \frac{50 \times 10^3}{7 \times 10^3} = 7.143 \quad \textbf{Ans.}$$

Example 4.3. Determine the frequency deviation and carrier swing for a frequency-modulated (FM) signal which has a resting frequency of 105.00 MHz and whose upper frequency is 105.007 MHz when modulated by a particular wave. Find the lowest frequency reached by the FM wave.

Solution: Given that

Carrier frequency $f_c = 105.000$ MHz

$$f_H = 105.007 \text{ MHz}$$

We know that frequency deviation Δ_f is defined as the maximum change in frequency of the modulated signal away from the resting or carrier frequency *i.e.*,

$$\Delta_f = f_H - f_c = (105.007 - 105.000)\ 10^6$$

or $$\Delta_f = 0.007 \times 10^6 \text{ Hz}$$

or $$\Delta_f = 7000 \text{ Hz} = 7 \text{ kHz} \quad \textbf{Ans.}$$

Carrier swing is given by

carrier swing = 2 × frequency deviation

$$= 2 \times \Delta_f = 2 \times 7 = 14 \text{ kHz.} \quad \textbf{Ans.}$$

Lowest frequency reached by the modulated wave is given as

$$f_L = f_c - \Delta_f = (105.000 - 0.007) \text{ MHz.}$$
$$= 104.993 \text{ MHz.} \quad \textbf{Ans.}$$

Example 4.4. What is the modulation index of an FM signal having a carrier swing of 100 kHz when the modulating signal has a frequency of 8 kHz?

Solution: Given that

carrier swing = 100 kHz

modulating frequency $f_m = 8$ kHz

Modulation index is given as

$$m_f = \frac{\text{frequency deviation}}{\text{modulating frequency}} = \frac{\Delta_f}{f_m} \quad \text{...(i)}$$

But we know that

carrier swing = $2 \times \Delta_f$

$$\Delta_f = \frac{\text{carrier swing}}{2} = \frac{100}{2} = 50 \text{ kHz}$$

Using equation *(i)* we get

$$m_f = \frac{50}{8} = 6.25 \quad \textbf{Ans.}$$

Example 4.5. An FM transmission has a frequency deviation of 20 kHz.

(i) determine the per cent modulation of this signal if it is broadcasted in the 88–108 MHz band.

(ii) Calculate the per cent modulation if this signal is broadcasted as the audio portion of a television broadcast.

Solution: Given that

$$\Delta_f = 20 \text{ kHz}$$

(i) Per cent modulation for an FM wave is defined as

$$M = \frac{\Delta f_{actual}}{\Delta f_{max}} \times 100$$

Δf_{actual} is given = 20 kHz

The maximum frequency deviation Δf_{max} permitted in the FM broadcast band is 75 kHz.

Thus, $$M = \frac{20 \times 10^3}{75 \times 10^3} \times 100 = 26.67\% \quad \textbf{Ans.}$$

(ii) $$M = \frac{\Delta f_{actual}}{\Delta f_{max}} \times 100$$

$$\Delta f_{actual} = 20 \text{ kHz}$$

The maximum frequency deviation Δf_{max} permitted for the FM audio portion of a TV broadcast is 25 kHz.

Thus $$M = \frac{20 \times 10^3}{25 \times 10^3} \times 100 = 80\% \quad \textbf{Ans.}$$

4.6. Phasor Representation of Angle Modulated (*i.e.* FM and PM Waves)

The expression for FM wave is

$$s(t)_{FM} = A \cos\left[w_c t + k_f \int_0^t x(t)\, dt\right] \quad \text{...(4.33)}$$

Similarly, the expression for PM wave is

$$s(t)_{PM} = A \cos[w_c t + k_p \cdot x(t)] \quad \text{...(4.34)}$$

We may consider the signa ls $s(t)_{FM}$ and $s(t)_{PM}$ as real parts of phasors C_{FM} and C_{PM} respectively expressed as

$$C_{FM} = Ae^{j[(w_c t + k_f \int x(t)\, dt)]}$$

and $$C_{PM} = Ae^{j[w_c t + k_p x(t)]}$$

Assuming $$\int x(t)\, dt = y(t)$$

So that $$C_{FM} = Ae^{j[w_c t + k_f y(t)]} \quad \text{...(4.35)}$$

4.7. Types of Frequency Modulation (FM)

(U.P. Tech., Semester, Exam., 2003-04) (04 marks)

In article 4.3, we discussed that the bandwidth of an *FM* signal depends upon the frequency deviation $\Delta w = k_f\, x(t)$.

Thus, as a matter of fact, if the frequency deviation is high, the bandwidth will be large. Similarly, if the frequency deviation is low, the bandwidth will be small.

Since deviation is given by $\Delta w = k_f\, x(t)$, it means that for any given $x(t)$, the frequency deviation and therefore the bandwidth will depend upon frequency sensitivity k_f. Hence, when k_f is quite small then the bandwidth will be narrow and when k_f is large then the bandwidth will be large.

Thus, depending upon the value of frequency sensitivity k_f, *FM* may be divided as under –

(i) Narrowband FM: In this case, k_f is small and hence the bandwidth of *FM* is narrow.*

(ii) Wideband FM: In this case, k_f is large and hence the *FM* signal has a wide bandwidth.**

4.7.1. Narrowband FM

We know that the general expression for FM wave in the phasor form is given by

$$C_{FM} = Ae^{j[w_c t + k_f\, y(t)]} \qquad ...(4.36)$$

However, for a narrowband FM, we have

$k_f \times y(t) << 1$ for all the values of t

Thus $\quad e^{jk_f \cdot y(t)} \cong 1 + jk_f \cdot y(t)$

Hence, the FM phasor expression, now, becomes

$$C_{FM} = Ae^{jw_c t} \cdot e^{jk_f\, y(t)} = A[1 + jk_f \cdot y(t)]\; e^{jw_c t}$$

Also, since the *FM* signal is the real part of its phasor representation, therefore, we write

$$s(t)_{FM} = Re[C_{FM}] = A \cos w_c t - Ak_f y(t) \sin w_c t \qquad ...(4.37)$$

In the same way, on the other hand, the expression for narrowband PM is given as

$$s(t)_{PM} = A \cos w_c t - Ak_p\, x(t) \sin w_c t \qquad ...(4.38)$$

Thus, the expressions for narrowband FM and PM are quite similar to the expression for AM signal with a little modification. Therefore, the bandwidth of a narrowband FM is almost same as that of the AM signal.

Now, we can also show that the signals $y(t)$ and $x(t)$ have the same bandwidth.

Since

$$y(t) = \int_0^t x(t)\, dt$$

Therefore, $$F[y(t)] = F\left[\int_0^t x(t)\, dt\right] = \frac{1}{jw} \cdot X(w)$$

or $$Y(w) = \frac{1}{jw} \cdot X(w) \qquad ...(4.39)$$

Conclusion

Hence, from above expression it is clear that if $X(w)$ is band limited to w_m, then $Y(w)$ is also band-limited to w_m. Both FM and PM wave expressions have carrier and sideband terms which are identical to the AM expression except with differing phase relations between carrier and sideband terms.***

4.7.2. Generation of Narrowband FM

We know that the expressions for narrowband FM and narrowband PM are given by

$$s(t)_{FM} = A \cos w_c t - Ak_f \cdot y(t) \sin w_c t \qquad ...(4.40)$$

and $$s(t)_{PM} = A \cos w_c t - Ak_p\, x(t) \sin w_c t \qquad ...(4.41)$$

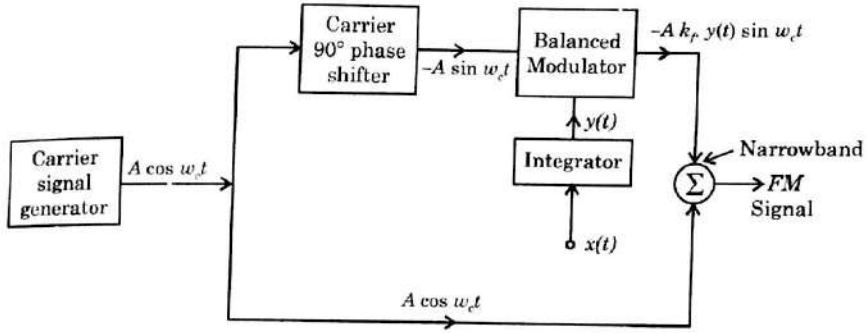
* We shall observe, later on, that the BW of a narrowband FM is almost same as that of AM.

** We shall study, later on, that unlike AM, the bandwidth of wideband FM is too large.

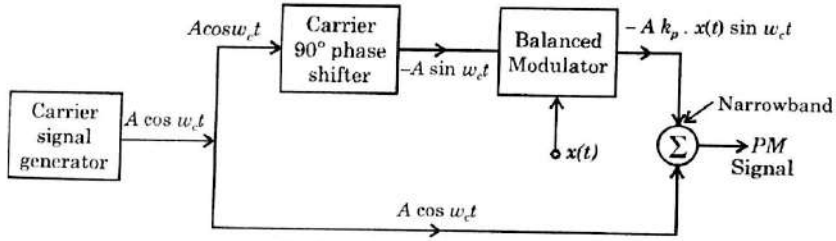
*** In fact, this is the difference that makes the amplitude of FM constant.

The equations (4.40) and (4.41) provide an idea to generate narrowband FM and narrowband PM respectively. This means that the sideband terms are generated by a balanced modulator, as in a DSB-SC system, and then the carrier term is added to sideband terms.

Based upon equations (4.40) and (4.41), figure 4.5 shows the methods to generate narrowband FM and narrowband *PM*. It may be observed that the block diagrams shown in figure 4.5 satisfy the corresponding equations for narrowband *FM* and narrowband *PM*.



(a) Narrowband FM Generator



(b) Narrowband PM Generator.

Fig. 4.5. *Generation schemes for narrowband FM and PM*

4.7.3. Expression for Single-tone Narrowband FM

We know that the general expression for narrowband FM is given as

$$s(t)_{FM} = A \cos w_c t - Ak_f\, y(t) \sin w_c t \qquad ...(4.42)$$

Here $$y(t) = \int x(t)\, dt$$

where $x(t)$ is any modulating signal

and $A \cos w_c t$ is a carrier signal.

For a single-tone narrowband *FM*, we take modulating signal as consisting of a single frequency

i.e.

$$x(t) = V_m \cos w_m t^* \qquad ...(4.43)$$

where w_m is the single modulating frequency.

Thus $$y(t) = \int x(t)\, dt$$

$$y(t) = \int V_m \cos w_m t = \frac{V_m}{w_m} \cdot \sin w_m t$$

Putting this value of $y(t)$ in the equation (4.42), we get

* Here, we have taken the signal $x(t)$ as a voltage signal.

$$s(t)_{FM} = A\cos w_c t - Ak_f \cdot \frac{V_m}{w_m} \sin w_m t . \sin w_c t$$

But we know that

$$\frac{k_f V_m}{w_m} = m_f = \text{modulation index}$$

Hence $$s(t)_{FM} = A\cos w_c t - Am_f . \sin w_m t \sin w_c t \quad ...(4.44)$$

which is the required expression for a single-tone narrowband FM.

4.8. Wideband FM

Definition

When the value of modulation index m_f is quite large, then in FM, a large number of sidebands are produced and hence the bandwidth of FM is sufficiently large. This type of FM system is known as wideband FM.

Now, we shall analyse wideband FM by considering the case of a single-tone FM system.

The expression for a single-tone FM wave is given as

$$s(t) = A\cos(w_c t + m_f \sin w_m t) \quad ...(4.45)$$

This expression may be considered as a real part of the exponential phasor given by

$$C_{FM}(t) = Ae^{j(w_c t + m_f \sin w_m t)}$$

$$C_{FM}(t) = Ae^{jw_c t} e^{jm_f \sin w_m t} \quad ...(4.46)$$

In above expression, the second exponential is a periodic function of period $\frac{1}{f_m}$ and may be expanded in the form of a complex Fourier series as under:

$$e^{jm_f \sin w_m t} = \sum_{n=-\infty}^{\infty} C_n e^{jnw_m t} \quad \text{for } -\frac{1}{2f_m} \le t \le \frac{1}{2f_m} \quad ...(4.47)$$

Here, the coefficient C_n is given by

$$C_n = f_m \int_{\frac{-\pi}{w_m}}^{\frac{\pi}{w_m}} e^{j(m_f \sin w_m t)} e^{-jn w_m t}\, dt$$

Substituting $x = w_m t$, we have

$$C_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{j(m_f \sin x - nx)} dx \quad ...(4.48)$$

In above expression, the integral on the right hand side is the **n^{th} order Bessel function of the first kind and argument m_f.** This function is represented by $J_n(m_f)$.

Thus

$$C_n = J_n(m_f) \quad ...(4.49)$$

Putting the value of C_n in equation (4.47), we get

$$e^{jm_f \sin w_m t} = \sum_{n=-\infty}^{\infty} J_n(m_f) e^{jnw_m t}$$

Substituting above expression in equation (4.46), we have

$$C_{FM}(t) = Ae^{jw_c t} \sum_{n=-\infty}^{\infty} J_n(m_f) e^{jnw_m t}$$

$$C_{FM}(t) = A \sum_{n=-\infty}^{\infty} J_n(m_f) e^{j(w_c + nw_m)t} \quad ...(4.50)$$

In above expression, the real part of the right hand side provides the expression for FM signal *i.e.*,

$$s(t) = A \sum_{n=-\infty}^{\infty} J_n(m_f) \cos(w_c + nw_m t) \quad ...(4.51)$$

Thus, we have converted original single-tone FM expression into modified form which consists of Bessel function.

Now, the Bessel function $J_n(m_f)$ may be expanded in a power series given by

$$J_n(m_f) = \sum_{m=0}^{\infty} \frac{(-1)^m \left(\frac{1}{2} m_f\right)^{n+2m}}{\angle m \, \angle n+m} \quad ...(4.52)$$

The above expression *i.e.*, the Bessel function $J_n(m_f)$ with respect to m_f has been plotted in figure 4.6 for various positive integer values of n.

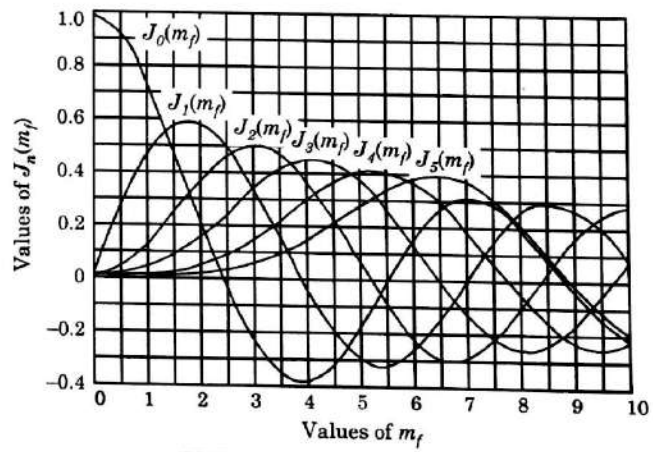


Fig. 4.6. *Bessel functions.*

Few important properties of the Bessel function may be summarized as follows:

(*i*) $J_n(m_f) = J_{-n}(m_f)$ for even n

$J_n(m_f) = -J_{-n}(m_f)$ for odd n

(*ii*) For a small value of m_f, we have

$$J_o(m_f) \cong 1$$

$$J_1(m_f) \cong \frac{m_f}{2}$$

$$J_n(m_f) \cong 0 \text{ for } n > 1$$

(*iii*) $$\sum_{n=-\infty}^{\infty} J_n^2(m_f) = 1$$

Making use of first property, the equation (4.52) may be expressed as

$$s(t) = A J_o(m_f) \cos w_c t + AJ_1(m_f) [\cos (w_c + w_m)t - \cos (w_c - w_m)t]$$
$$+ AJ_2(m_f) [\cos (w_c + 2w_m) t + \cos (w_c - 2w_m)t]$$
$$+ AJ_3(m_f) [\cos (w_c + 3w_m)t - \cos (w_c - 3 w_m)t] + \ldots \quad \ldots(4.53)$$

In order to evaluate the value of a given pair of sidebands or the value of the carrier, it is necessary to know the value of the corresponding Bessel function. Table 4.1 provides this type of information.

We may observe the following points from above expression:

(*i*) The expression contains a carrier term $\cos w_c t$ having magnitude $AJ_o(m_f)$. This means that the magnitude of the carrier term is reduced by a factor $J_o(m_f)$. Thus in FM, unlike in AM, the amplitude of the carrier component does not remain constant.

(*ii*) From expression in equation (4.53) it may be noted that in FM, theoretically an infinite number of sidebands are produced and the amplitude of each sideband is determined by the corresponding Bessel function $J_n(m_f)$. Thus, the presence of an infinite number of sideband components makes the ideal bandwidth for FM signal infinite. But practically, the distant sidebands with small or insignificant amplitudes are ignored and sidebands with significant amplitudes are considered in calculating the bandwidth for FM signal. Thus the practical bandwidth for FM is finite.

The number of significant sidebands generated in FM depends upon the value of modulation index m_f *i.e.* modulation index determines how many sideband components have significant amplitudes. The modulation index m_f in turn, depends upon the maximum frequency deviation (Δw) and modulating frequency (w_m). This means that practical bandwidth of FM system depends upon the value of modulation index.

(*iv*) From Table 4.1, it is clear that for a small value of modulation index m_f (*i.e.*, less than 0.6), only $J_o(m_f)$ and $J_1(m_f)$ have significant values and the higher terms are negligible. Thus, for small values of m_f, FM has a carrier term and only one pair of sidebands. This is equivalent to narrowband FM.

(*v*) Since the amplitude of FM signal remains unchanged, the power of the FM signal will be same as that of the unmodulated carrier.

(*vi*) From point (*v*), we know that the total power of FM remains same as that of the unmodulated carrier which is unlike AM where the total power depends upon the modulation index m_f

Actually in FM, out of the total power $\frac{A^2}{2}$, the power carried by the carrier term depends upon the value of coefficient $J_o(m_f)$ and the power carried by a sideband depends upon the value of corresponding $J_o(m_f)$. Now if m_f is adjusted to have a value of 2.4 or 5.52 so that $J_o(m_f) = 0$, then the power carried by the carrier term in FM would be zero. Thus, all the power is carried by the sidebands and thus giving a 100 per cent transmission efficiency. However, for other values of m_f, some amount of power is carried by the carrier also and hence the transmission efficiency is less than 100 per cent. This means that by choosing the value of m_f, one can get FM efficiency much more than AM which is 33% and may approach equal to the efficiency of DSB-SC (100%). Hence the efficiency of FM is between AM and DSB-SC.

(*viii*) There are a number of international regulations for frequency modulation prescribed by CCIR (*i.e.* Consultative Committee for International Radio). These regulations must be followed by all the commercial FM broadcast stations to avoid interference problem. They are as follows:

Table 4.1. Bessel Functions of the First Kind

J_0	J_1	J_2	J_3	J_4	J_5	J_6	J_7	J_8	J_9	J_{10}	J_{11}	J_{12}	J_{13}	J_{14}	J_{15}	J_{16}	J_{16}
0.00	1.00	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
0.25	0.98	0.12	—	—	—	—	—	—	—	—	—	—	—	—	—	—	—
0.5	0.94	0.24	0.03	—	—	—	—	—	—	—	—	—	—	—	—	—	—
1.0	0.77	0.44	0.11	0.02	—	—	—	—	—	—	—	—	—	—	—	—	—
1.5	0.51	0.56	0.23	0.06	0.01	—	—	—	—	—	—	—	—	—	—	—	—
2.0	0.22	0.58	0.35	0.13	0.03	—	—	—	—	—	—	—	—	—	—	—	—
2.5	–0.05	0.50	0.45	0.22	0.07	0.02	—	—	—	—	—	—	—	—	—	—	—
3.0	–0.26	0.34	0.49	0.31	0.13	0.04	0.01	—	—	—	—	—	—	—	—	—	—
4.0	–0.40	–0.07	0.36	0.43	0.28	0.13	0.05	0.02	—	—	—	—	—	—	—	—	—
5.0	–0.18	–0.33	0.05	0.36	0.39	0.26	0.13	0.05	0.02	—	—	—	—	—	—	—	—
6.0	0.15	–0.28	–0.24	0.11	0.36	0.36	0.25	0.13	0.06	0.02	—	—	—	—	—	—	—
7.0	0.30	0.00	–0.30	–0.17	0.16	0.35	0.34	0.23	0.13	0.06	0.02	—	—	—	—	—	—
8.0	0.17	0.23	–0.11	–0.29	–0.10	0.19	0.34	0.32	0.22	0.13	0.06	0.03	—	—	—	—	—
9.0	–0.09	0.24	0.14	–0.18	–0.27	–0.06	0.20	0.33	0.30	0.21	0.12	0.06	0.03	0.01	—	—	—
10.0	–0.25	0.04	0.25	0.06	–0.22	–0.23	–0.01	0.22	0.31	0.29	0.20	0.12	0.06	0.03	0.01	—	—
12.0	0.05	–0.22	–0.08	0.20	0.18	–0.07	–0.24	–0.17	0.05	0.23	0.30	0.27	0.20	0 12	0.07	0.03	0.01
15.0	–0.01	0.21	0.04	–0.19	–0.12	–0.13	0.21	0.03	–0.17	–0.22	–0.09	0.10	0.24	0 28	0.25	0.18	0.12

(a) Maximum modulating frequency is 15 kHz
(b) Maximum frequency deviation is ± 75 kHz
(c) Frequency stability of the carrier is ± 2 kHz
(d) Allowable bandwidth per channel = 200 kHz.

4.9. Effect of Variation of Modulation Index m_f on the Spectrum of FM Signal

As discussed in previous section, the number of sidebands produced in FM increases with increase in the value of modulation index m_f.

The modulation index m_f is given as

$$m_f = \frac{\text{frequency deviation}}{\text{modulating frequency}}$$

or

$$m_f = \frac{\Delta w}{w_m} = \frac{k_f V_m}{w_m}$$

Thus according to above expression m_f may be increased

(i) by reducing w_m and keeping amplitude of modulating signal V_m constant.
(ii) by increasing amplitude of modulating signal V_m and keeping modulating frequency w_m constant.

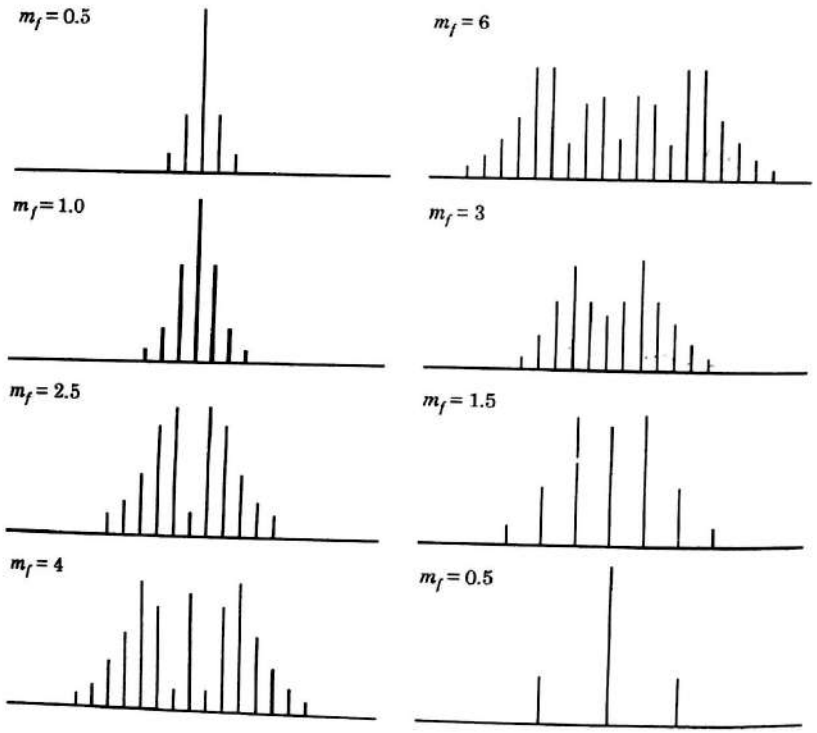


Fig. 4.7.

Infact, it is possible to evaluate the size of the carrier and each sideband for each specific value of the modulation index. When this is done, the frequency spectrum of the FM wave for that particular

value of modulation index m_f may be plotted. This has been done in figure 4.7 which shows these frequency spectrums first for increasing deviation and f_m constant and then for decreasing modulating frequency f_m and frequency deviation Δf constant. From figure 4.7, it may be observed that as modulation depth increases, so does the bandwidth and also that reduction in modulating frequency f_m increases the number of sidebands, though not necessarily the bandwidth. Another point which may be observed is that although the number of sideband components is theoretically infinite, but in practice a lot of the higher sidebands have insignificant relative amplitudes and this is why they are not shown in the spectrums.

4.10. Transmission Bandwidth of FM Signal

As discussed in previous article, the bandwidth of FM signal depends upon the value of modulation index m_f. With the increase of modulation index m_f, more and more number of sidebands acquire significant amplitudes and thus bandwidth is increased.

For a particular value of modulation index m_f, we can note the number of significant sidebands from Table 4.1.

Basically the effective bandwidth is the separation between the two extreme significant side frequencies on either side of the carrier.

Fourier analysis reveals that the number of side frequencies which contain a significant amount of power and thus the effective bandwidth of the FM signal is dependent on the modulation index of the modulated wave.

For practical purpose, a side frequency is considered to be significant if its amplitude is at least one per cent of the unmodulated carrier amplitude.

Let us consider that in the spectrum of FM wave the number of significant sidebands is n. Since the upper sidebands are separated by w_m, therefore, they form a frequency band of nw_m.

A similar frequency band of nw_m is formed by the lower sidebands.

Thus, for n sidebands the bandwidth of FM wave is given by

$$BW = 2\,nw_m \text{ radians/sec.} = 2nf_m \text{ Hz.} \qquad ...(4.55)$$

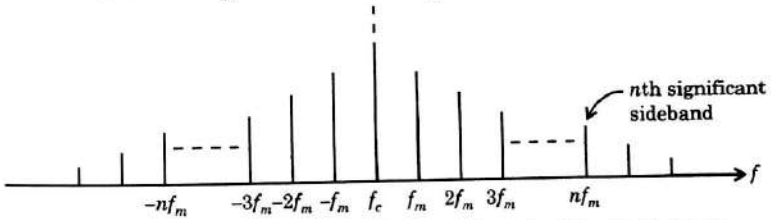


Fig. 4.8. *Frequency spectrum of FM wave for n significant sidebands.*

Universal Curve

Schwartz developed a graph for determining the bandwidth of an FM signal if the modulation index is known. This graph has been shown in figure 4.9. Schwartz uses as his criterion the rule of thumb that any frequency component with a signal strength (voltage) less than 1% of that of the unmodulated carrier will be considered too small to be significant. This graph is also called **universal curve** which shows the variation of the Bandwidth B normalised with respect to Δf against m_f.

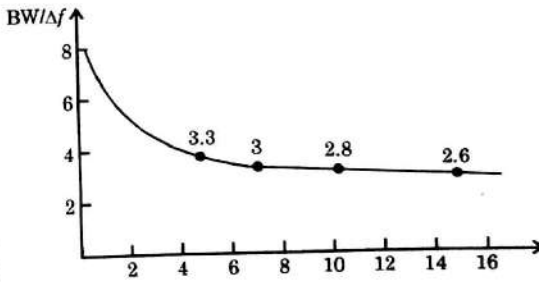


Fig. 4.9. *Schwartz's universal curve.*

4.10.1. Carson's Rule

Carson's rule provides a thumb formula to calculate the bandwidth of a single-tone wideband FM. According to this rule, the FM bandwidth is given as twice the sum of the frequency deviation and the highest modulating frequency. However, it must be remembered that this rule is just an approximation.

Mathematically,

$$BW = 2\,(\Delta w + w_m)$$

But $$m_f = \frac{\Delta w}{w_m}$$

Therefore, $$BW = 2\,(m_f\, w_m + w_m) = 2(1 + m_f)w_m \quad ...(4.56)$$

Regarding Carson's rule, we can consider two special cases:

(*i*) $\because$ $$BW = 2(1 + m_f)\, w_m$$

If $\Delta w << w_m$ *i.e.* $m_f << 1$ as is the case for narrowband FM

then $$BW = 2\, w_m{}^{*} \quad ...(4.57)$$

(*ii*) $\because$ $$BW = 2(1 + m_f)\, w_m$$

If $\Delta w >> w_m$ *i.e.* $m_f >> 1$ as is the case for wideband FM

then $$BW = 2(m_f)w_m$$

But $$m_f\, w_m = \Delta w$$

Thus $$BW = 2\Delta w^{**} \quad ...(4.58)$$

Example 4.6. Find the bandwidth of a commercial FM transmission if frequency deviation Δf = 75 kHz and modulating frequency f_m = 15 kHz.

Solution: According to Carson's rule, the bandwidth is given by

$$BW = 2(\Delta_f + f_m) = 2(75 + 15)$$

$$BW = 2 \times 90 = 180 \text{ kHz} \quad \textbf{Ans.}$$

Example 4.7. Determine the bandwidth of a narrowband FM signal which is generated by a 4 kHz audio signal modulating a 125 MHz carrier.

Solution: Given that

$$f_m = 4 \text{ kHz}$$

and $$f_c = 125 \text{ MHz}$$

Since this is a narrowband FM signal the bandwidth is found merely by doubling modulating frequency as

$$BW = 2f_m = 2 \times 4 \times 10^3$$

$$BW = 8 \text{ kHz} \quad \textbf{Ans.}$$

Example 4.8. A 2 kHz audio signal modulates a 50 MHz carrier causing a frequency deviation of 2.5 kHz. Determine the bandwidth of the FM signal.

Solution: Given that

$$\Delta f = 2 \text{ kHz}$$

$$f_c = 50 \text{ MHz}$$

$$\Delta f = 2.5 \text{ kHz}$$

Modulation index m_f is given by

* This case is equivalent to AM.

** For large values of m_f, this bandwidth relation has a very small error and hence can be assumed to be true bandwidth for all practical purposes.

$$m_f = \frac{\Delta_f}{f_m} = \frac{2.5}{2} = 1.25$$

Since the value of m_f is less than $\pi/2$, therefore we are dealing with a narrowband signal.

Thus,

$$BW = 2\, f_m = 2 \times 2 \times 10^3 = 4 \text{ kHz} \quad \textbf{Ans.}$$

Example 4.9. The maximum deviation allowed in an FM broadcast system is 75 kHz. If the modulating signal is a single-tone sinusoid of 8 kHz, determine the bandwidth of FM signal. What will be the bandwidth when modulating signal amplitude is doubled?

Solution: Given that

$$\Delta f = 75 \text{ kHz}$$

$$f_m = 8 \text{ kHz}$$

Bandwidth is given by

$$BW = 2(\Delta f + f_m) = 2\,(75 + 8) = 166 \text{ kHz}$$

Now when the modulating signal amplitude is doubled, the frequency deviation Δf becomes $2 \times 75 = 150$ kHz.

Therefore the bandwidth becomes

$$BW = 2\,(\Delta_f + f_m) = 2(150 + 8) = 316 \text{ kHz} \quad \textbf{Ans.}$$

4.11. FM Bandwidth for an Arbitrary Modulating Signal *x(t)*

Till now, we have discussed the bandwidth for a single-tone FM. Now let us consider that the modulating signal $x(t)$ is an arbitrary signal *i.e.* it may consists of large number of frequency components. In this case to calculate the bandwidth, we first find a parameter known as deviation ratio D defined as

$$D = \frac{\text{Peak frequency deviation corresponding to the maximum possible amplitude of } x(t)}{\text{Maximum frequency component present in the modulating signal } x(t)} \quad ...(4.59)$$

Note: *The deviation ratio D has same meaning to estimate the bandwidth as m_f does for the single-tone FM. These for this case assuming deviation ratio D as equivalent to m_f, we can calculate the transmission bandwidth of FM using Carson's rule.*

4.12. Narrowband FM Versus Wideband FM

An examination of the Schwartz bandwidth curve of figure 4.10 shows that at high values of m_f the curve tends towards a horizontal asymptote and at low values of m_f it tends toward the vertical. Detailed mathematical study shows that the bandwidth of an FM signal for which m_f is less than $\pi/2$ is dependent mainly upon the frequency of the modulating signal and is quite independent of frequency deviation. Further analysis shows that the bandwidth of an FM signal for which m_f is less than $\pi/2$ is equal to twice the modulating frequency.

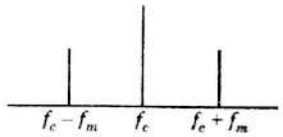


Fig. 4.10.

$$\text{Bandwidth} = 2f_m \quad \text{for } m_f < \pi/2$$

Just as with AM, and unlike the situation in which $m_f > \pi/2$, two side frequencies show up for each modulating frequency, one above and one below the frequency of the carrier, each spaced f_m away from the carrier frequency. Because of the limited bandwidth of FM signals with $m_f < \pi/2$, such modulations are known as narrowband FM, and FM signals with $m_f > \pi/2$ are known as *wideband FM*.

Though the spectrum for an AM signal and a narrowband FM signal appear to be the same, a Fourier analysis shows that the magnitude and phase relationships for AM and FM are quite different. See Figure 4.10 for the frequency spectrum of a narrowband FM signal.

Many of the advantages obtained with wideband FM, such as noise reduction, are not available with narrowband FM. Why, then, would one want to use narrowband FM rather than AM? One reason is that with narrowband FM (as well as with wideband FM) the power content at the carrier frequency decreases as the modulation increases so that we have desirable situation of putting the power where the information is.

Example 4.10. A 5 kHz audio tone is used to modulate a 50 MHz carrier causing a frequency deviation of 2 kHz. Determine (*i*) the modulation index and (*ii*) the bandwidth of the FM signal.

Solution : Given that $f_m = 5$ kHz,

$f_c = 50.0$ MHz,

$\Delta f = 20$ kHz

(*i*) Modulation index is defined as

$$m_f = \frac{\Delta f}{f_m} = \frac{20 \times 10^3}{5 \times 10^3} = 4 \quad \textbf{Ans.}$$

(*ii*) Referring to the Schwartz bandwidth curve, Figure 4.11, and entering on the horizontal axis with $m_f = 4$, it is found that

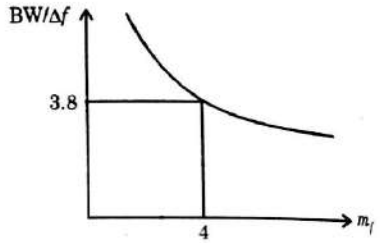


Fig. 4.11.

$$\frac{BW}{\Delta f} = 3.8$$

This is shown in figure 4.11. Substituting 20×10^3 for Δf as given,

$$\frac{BW}{20 \times 10^3} = 3.8$$

Solving for BW, we have

$$BW = 3.8 \times 20 \times 10^3 = 76 \times 10^3 = 76 \text{ kHz} \quad \textbf{Ans.}$$

Example 4.11. Determine the frequency of the modulating signal which is producing an FM signal having a bandwidth of 50 kHz when the frequency deviation of the FM signal is 10 kHz.

Solution: Given that BW = 50 kHz;

$\Delta f = 1$ kHZ

In order to find f_m, reference must be made to the Schwartz bandwidth curve, *i.e.*, figure 4.12. In order to enter this curve, we have to determine BW/Δf as

$$\frac{BW}{\Delta f} = \frac{50 \times 10^3}{10 \times 10^3} = 5$$

From figure 4.12, we have

$$m_f = 2 = \frac{\Delta f}{f_m}$$

Therefore, $$2 = \frac{10 \times 10^3}{f_m}$$

$$f_m = \frac{10 \times 10^3}{2} = 5 \text{ kHz} \quad \textbf{Ans.}$$

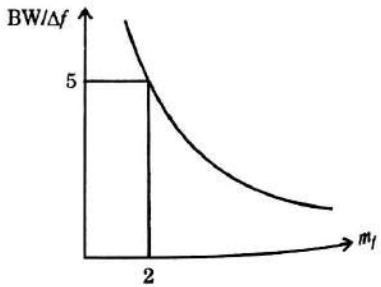


Fig. 4.12.

4.13. Multiple Frequency Modulation

In article 4.5, we discussed a specific case of a single frequency-modulating signal. Now, let us extend these results to the case of multiple frequencies. First only two frequencies will be considered. It can then be generalized to any number of frequencies.

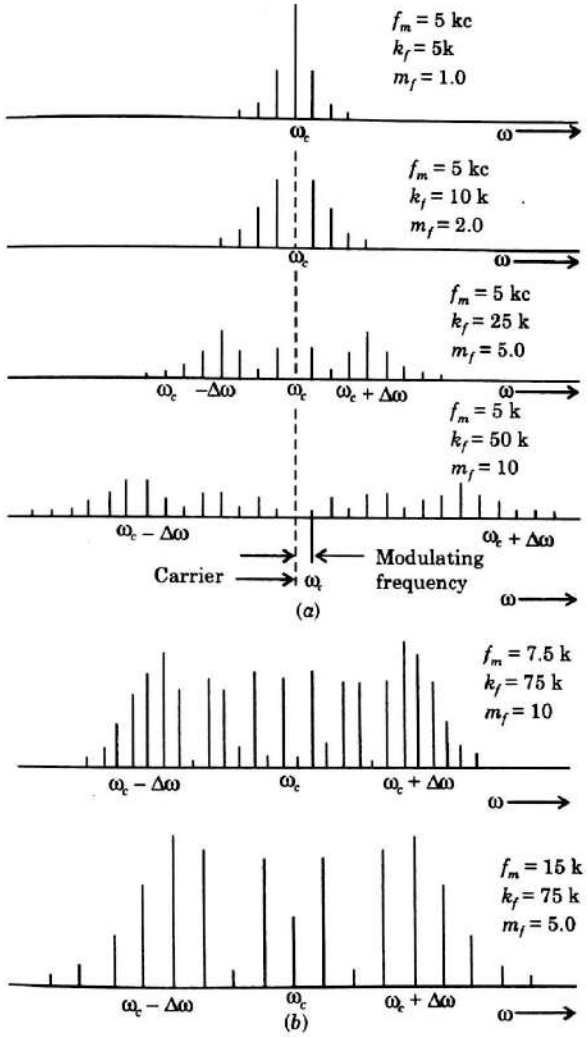


Fig. 4.13.

Let us consider

$$x(t) = a_1 \cos \omega_1 t + a_2 \cos \omega_2 t$$

$$\omega_i = \omega_c + k_f x(t) = \omega_c + k_f (a_1 \cos \omega_1 t + a_2 \cos \omega_2 t)$$

The maximum frequency deviation will be

$$\Delta\omega = (a_1 + a_2) k_f$$

and $$\phi = \int \omega_i \, dt = \omega_c t + \frac{a_1 k_f}{\omega_1} \sin \omega_1 t + \frac{a_2 k_f}{\omega_2} \sin \omega_2 t$$

or $$\phi = \omega_c t + m_1 \sin \omega_1 t + m_2 \sin \omega_2 t$$

where $$m_1 = \frac{a_1 k_f}{\omega_1} \quad \text{and} \quad m_2 = \frac{a_2 k_f}{\omega_2}$$

Also, $$\hat{s}(t) = Ae^{j\theta(t)} = Ae^{j(\omega_c t + m_1 \sin \omega_1 t + m_2 \sin \omega_2 t)} \,*$$

or $$\hat{s}(t) = Ae^{j\omega_c t} \left(e^{jm_1 \sin \omega_1 t}\right)\left(e^{jm_2 \sin \omega_2 t}\right) \qquad ...(4.60)$$

The exponentials in the brackets are obviously periodic functions with periods $2\pi/\omega_1$ and $2\pi/\omega_2$, respectively. These exponentials can be represented by Fourier series using Bessel functions.

Hence, we have

$$s(t) = Ae^{j\omega_c t}\left[\sum_{n=-\infty}^{\infty} J_n(m_1)\, e^{jn\omega_1 t}\right]\left[\sum_{k=-\infty}^{\infty} J_k(m_2) e^{jk\omega_1 t}\right]$$

or $$s(t) = A \sum_{n=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} J_n(m_1)\, J_k(m_2)\, e^{j(\omega_c + n\omega_1 + k\omega_2)t} \qquad ...(4.61)$$

and $$s(t) = A \sum_{n=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} J_n(m_1)\, J_k(m_2) \cos\left[(\omega_c + n\omega_1 + k\omega_2)t\right] \qquad ...(4.62)$$

It is evident from this result that when $x(t)$ is composed of two frequencies ω_1 and ω_2, the spectrum of the FM signal contains sidebands $(\omega_c \pm n\omega_1)$ and $(\omega_c \pm k\omega_2)$, corresponding to frequencies ω_1 and ω_2. In addition, there are cross modulation terms $(\omega_c \pm n\omega_1 \pm k\omega_2)$.

It may be noted that this behaviour contrasts to that observed in AM. In AM, each new frequency added to the modulating signal gives rise to its own side bands only. There are no cross modulation terms. For this reason, AM is called linear modulation, whereas FM is a nonlinear type of modulation.

The results derived in this section for two frequencies can easily be extended to any number of frequencies.

4.14. Square Wave Modulation

Now, let us consider one more special case of FM modulation where the modulating signal $x(t)$ is a square wave as shown in figure 4.14 (*a*). The method discussed here is very general and is applicable to any periodic modulating signal. For a square wave (figure 4.14 (*a*)), the instantaneous frequency of the FM carrier is given by

$$\omega_i = \omega_c + k_f x(t)$$

and the phase ϕ is given by

$$\phi = \int \omega_i \, dt = \omega_c t + k_f \int x(t)\, dt$$

* Here, we are using the notation $\hat{s}(t)$ for the exponential representation of $s(t)$.
For example, if

$$s(t) = A \cos \theta(t)$$

then, $$\hat{s}(t) = Ae^{j\theta(t)}$$

and $$s(t) = Re\ \hat{s}(t)$$

or $$\phi = \omega_c t + \psi(t)$$

where $$\psi(t) = k_f \int x(t)\, dt$$

is shown in figure 4.14. This is a triangular periodic function with period T.

It may be noted that the maximum carrier frequency deviation $\Delta\omega$ in this case is k_f since $|x(t)|_{max} = 1$. Hence,

$$k_f = \Delta\omega$$

and

$$\psi(t) = \begin{cases} (\Delta\omega)t & \text{for } -\frac{T}{4} < t < \frac{T}{4} \\ (\Delta\omega)\left[\frac{T}{2} - t\right] & \text{for } \frac{T}{4} < t < \frac{3T}{4} \end{cases} \qquad ...(4.63)$$

and $$\psi(t) = \psi(t \pm nT)$$

Also $$\hat{s}(t) = Ae^{j\theta(t)} = Ae^{j\psi(t)} e^{j\omega_c t}$$

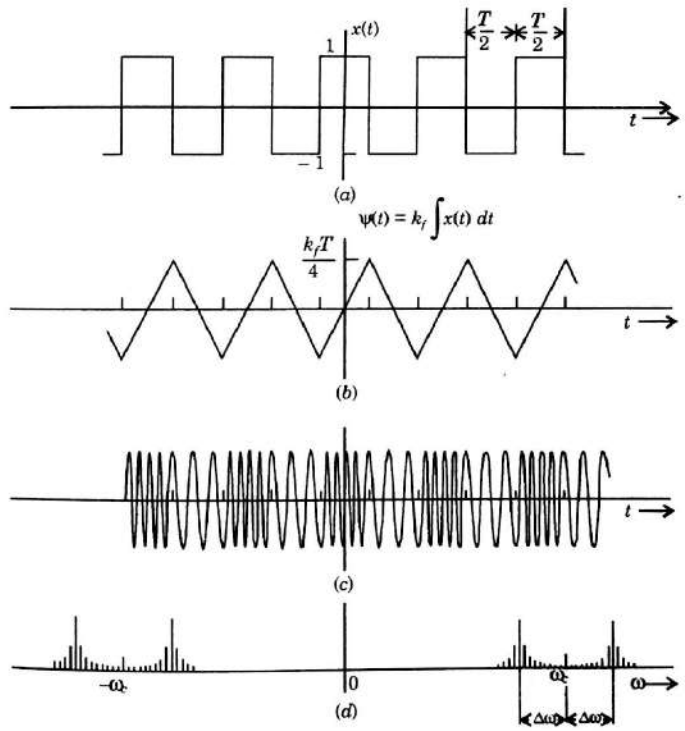


Fig. 4.14.

The function $e^{j\psi(t)}$ is itself a periodic function of period T, and it can be expressed by Fourier series as

$$e^{j\psi(t)} = \sum_{n=-\infty}^{\infty} \alpha_n e^{jn\omega_s t}$$

$$\omega_s = \frac{2\pi}{T}$$

where $$\alpha_n = \frac{1}{T}\int_{-T/4}^{3T/4} e^{j\psi(t)} e^{-jn\omega_s t} dt$$

Substitution of equation (4.63) in above equation and the subsequent integration will yield

$$\alpha_n = \frac{1}{2}\left\{Sa\left[\frac{\pi}{2}(m_f - n)\right] + (-1)^n Sa\left[\frac{\pi}{2}(m_f + n)\right]\right\} \quad ...(4.64)$$

where $$\beta = \frac{\Delta\omega}{\omega_s}$$

Hence, $$\hat{s}(t) = Ae^{j\psi(t)} e^{j\omega_c t} = A\sum_{n=-\infty}^{\infty} \alpha_n e^{j(\omega_c + n\omega_s)t}$$

and $$s(t) = A\sum_{n=-\infty}^{\infty} \alpha_n \cos(\omega_c + n\omega_s)t$$

The frequency spectrum of $s(t)$ is shown in figure 4.14(*d*). The method discussed here is general and can be applied to any modulating signal $x(t)$ that is periodic and has a zero mean.

4.15. Linear and Nonlinear Modulation

In the case of AM signals, the sidebands follow the principle of superposition. Thus, if signals $x_1(t)$ and $x_2(t)$ give rise to sidebands φ_1 and φ_2, respectively, the sidebands created by the composite signal $x_1(t) + x_2(t)$ will be $\varphi_1 + \varphi_2$. There is no intermodulation or cross-product sidebands as we observed in FM. For this reason, AM is called the linear type of modulation.*

The linear modulation lends itself easily to mathematical manipulations and generalizations. The spectrum of a modulated signal due to sum of two modulating signals can be found by calculating the spectrum due to each individual signal and then adding them together. This proves very useful in noise calculations in communication systems. For linear modulation systems, the effect of additive noise over the channel can be calculated by assuming the signal to be zero. This is not true of a nonlinear modulation where cross-modulation terms arise. For these reasons, we are interested in approximating a nonlinear modulation by a linear model. This is analogous to the case of system analysis where one can approximate a nonlinear system by a linear one over a limited range of signal amplitudes. Now, let us show that FM closely approximates a linear behaviour for a small modulation index.

4.15.1. Linearization of FM Signal for Small Modulation Index

For a small modulating index, FM closely exhibits linear behaviour. Let us consider again the case of modulating signal $x(t)$ containing two frequencies ω_1 and ω_2. The FM signal in this case is given by equation (4.60).

If m_1 and $m_2 < 1$, then for $x(t)$, we have

$$x(t) = a_1 \cos \omega_1 t + a_2 \cos \omega_2 t$$

$$\hat{s}(t) \cong A(1 + jm_1 \sin \omega_1 t)(1 + jm_2 \sin \omega_2 t)\, e^{j\omega_c t} \cong A(1 + jm_1 \sin \omega_1 t + jm_2 \sin \omega_2 t)e^{j\omega_c t}$$

* The modulated signal is a function of the modulating signal $x(t)$. Let $\varphi[x(t)]$ be the modulated signal. Then, the modulation is linear if $(d/d[x(t)]\}\ \{\phi[x(t)]\}$ is independent of $x(t)$. Otherwise it is a nonlinear modulation. The reader can easily verify, according to this definition that AM is linear, whereas FM is not.

Note that if

$$x_1(t) = a_1 \cos \omega_1 t \text{ and } m_1 << 1$$

then $$\hat{s}(t) = Ae^{jm_1 \sin \omega_1 t} e^{j\omega_c t}$$

or $$\hat{s}(t) \cong A[1 + jm_1 \sin \omega_1 t]e^{j\omega_c t}$$

If $$x_2(t) = a_2 \cos \omega_2 t \quad \text{and } m_2 << 1$$

then $$\hat{s}(t) \cong A(1 + jm_2 \sin \omega_2 t)e^{j\omega_c t}$$

It is easy to see that under the conditions $m_1, m_2 << 1$, the sidebands due to modulating signal $x_1(t) + x_2(t)$ are the sum of sidebands due to individually. Hence, FM can be assumed to be a linear modulation for a small modulation index. The cross-modulation terms under this assumption can be ignored.

4.16. Phase Modulation: An Analytic View

As discussed earlier, phase modulation may be analysed in a similar way as the FM signal. The PM has some distinct features as compared to FM. The main distinct feature of phase-modulation is that the deviation in the carrier frequency w_c in PM is linearly proportional to the baseband or modulating frequency w_m. However in FM, the deviation is independent of the baseband or modulating frequency.

Let us find an expression for frequency deviation for a PM wave.

We know that the total phase angle of a PM wave is expressed as

$$\phi_i = w_c t + k_p . x(t) \quad ...(4.66)$$

For a single-tone modulating signal, we have

$$x(t) = V_m \cos w_m t \quad ...(4.67)$$

Thus $$\phi_i = w_c t + k_p . V_m \cos w_m t \quad ...(4.68)$$

In above expression, the maximum departure in the phase is $k_p . V_m$. This is called as phase deviation denoted by θ_d.

Hence phase deviation, $\theta_d = k_p . V_m$ ...(4.69)

Now, the expression for the PM wave becomes

$$s(t)_{PM} = A \cos \phi_i = A \cos[w_c t + \theta_d \cos w_m t] \quad ...(4.70)$$

Also, the instantaneous frequency related to ϕ_i is expressed as

$$w_i = \frac{d\phi_i}{dt} = \frac{d}{dt}[w_c t + k_p V_m \cos w_m t]$$

or $$w_i = w_c - k_p V_m w_m \sin w_m t \quad ...(4.71)$$

Thus, the maximum departure in the frequency from w_c is $k_p . V_m . w_m$.

Hence, for PM, the frequency deviation is given by

$$\Delta w_{PM} = k_p . V_m . w_m \quad ...(4.72)$$

which depends upon the modulating frequency w_m. On the other hand, the frequency deviation for FM is

$$\Delta w_{FM} = k_f . V_m \quad ...(4.73)$$

Thus we conclude that for an equal bandwidth in FM and PM, we have

$$k_f = k_p . w_m \quad ...(4.74)$$

4.16.1. Transmission Bandwidth of PM

We may estimate the transmission bandwidth for PM using Carson's rule as

$$BW_{PM} \cong 2(\Delta w) \cong 2k_p . V_m . w_m \quad ...(4.75)$$

Hence, the bandwidth of the PM wave varies fastly with the variation in the modulating frequency w_m. On the other hand, the FM bandwidth varies slowly with modulating frequency w_m.

For PM, the modulation index m_p will be same as the deviation θ_d and will be expressed as

$$m_p = k_p . V_m = \theta_d \quad \text{...(4.76)}$$

Example 4.12. A baseband or modulating signal $x(t) = 5 \cos 2\pi 15 \times 10^3 t$ angle modulates a carrier signal $A \cos w_c t$.

(i) Determine the modulation index and bandwidth for

(a) FM system (b) PM system

(ii) Find the change in the bandwidth and modulation index for both FM and PM if modulating frequency f_m is reduced to 5 kHz.

Assume $k_f = k_p = 15$ **kHz/volt**

Solution: (i) Given that

$$V_m = 5\ V$$
$$f_m = 15 \text{ kHz}$$

(a) For FM system

Frequency deviation $\Delta f = k_f . V_m$

$$m_f = 15 \times 10^3 \times 5 = 75 \text{ kHz}$$

Hence, modulation index

$$m_f = \frac{\Delta f}{m_f} = \frac{75}{15} = 5$$

Thus, according to Carson's rule the bandwidth is

$$BW = 2(\Delta f + f_m) = 2(75 + 15) = 2 \times 90$$

or $BW = 180$ kHz

(b) For PM system

Frequency deviation $\Delta f = k_p . V_m . f_m = 15 \times 10^3 \times 5 \times 15 \times 10^3$

$$\Delta f = 1125 \text{ MHz}$$

Then the bandwidth will be

$$BW = 2(\Delta f + f_m) \cong 2\Delta f = 2 \times (1125 \text{ MHz})$$
$$BW = 2250 \text{ MHz}$$

It is clear that the bandwidth is quite large as compared to FM. This is due to the fact that the modulation index in PM is quite large as

$$m_p = k_p . V_m = 15 \times 10^3 \times 5 = 75{,}000$$

This very large value of modulation index m_p produces a large number of sidebands and thus this is the reason for very large bandwidth

(ii) In this case $f_m = 5$ kHz

(a) For FM

Since the frequency deviation Δf is independent of f_m and hence remains 75 kHz.

Thus $$m_f = \frac{\Delta f}{f_m} = \frac{75}{5} = 15$$

In this case the bandwidth will be

$$BW = 2(\Delta f + f_m) = 2(75 + 5) = 2 \times 80 = 160 \text{ kHz}$$

This means that in FM, the modulation index changes sufficiently with a change in f_m, however, the bandwidth changes only slightly.

(b) For PM

In this case the frequency-deviation Δf is dependent on f_m and is given as

$$\Delta f = k_p . V_m . f_m = 15 \times 10^3 \times 5 \times 5 \times 10^3$$
$$\Delta f = 375 \text{ MHz}$$

Thus $$BW = 2(\Delta f + f_m) \cong 2\Delta f = 2 \times 375 = 750 \text{ MHz}$$

Hence, the bandwidth has changed considerably as compared to previous one *i.e.* 2250 MHz.

The modulation index is independent of f_m and is given by

$$m_p = k_p . V_m = 15 \times 10^3 \times 5$$
$$m_p = 75 \text{ kHz}$$

which is same as the previous one *i.e.* 75 kHz.

4.17. Comparison of Angle Modulated Waves and Amplitude Modulated Wave

Figure 4.15 shows a single-tone modulating signal, a carrier signal, amplitude-modulated (AM) wave and angle-modulated (*i.e.* FM and PM) waves. On comparison, we note the following differences between two techniques:

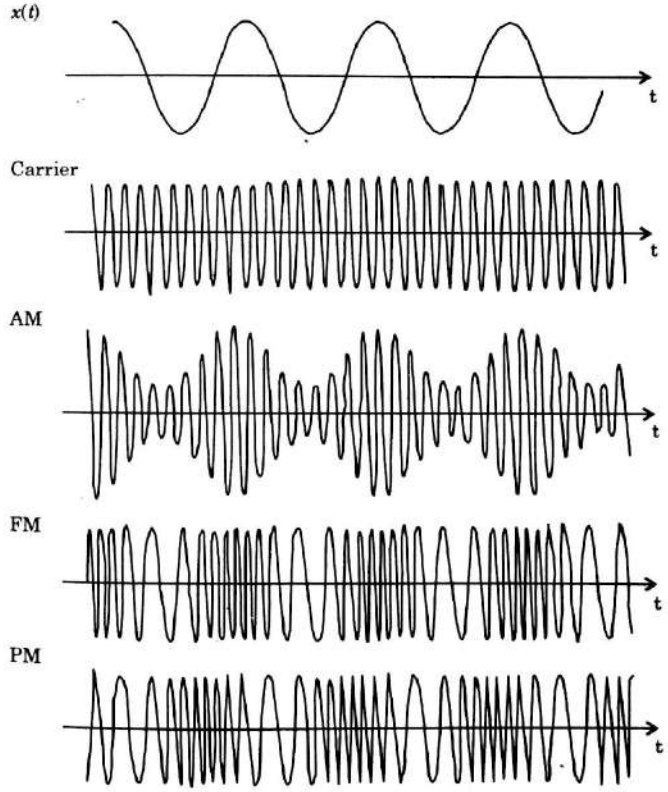


Fig. 4.15. *AM FM, and PM waveforms.*

(i) The envelope of FM wave or PM wave is constant and is equal to the unmodulated carrier amplitude. On the other hand, the envelope of AM wave is dependent on the modulating signal $x(t)$.

(ii) The zero crossings (*i.e.* the instants of time at which the waveform changes from negative to a positive value or vice-versa) of a FM wave or a PM wave no longer exhibit a perfect regularity in their spacing like AM wave. Thus this makes the instantaneous frequency of the angle modulated wave depend upon time.

4.18. Comparison of Frequency Modulation and Amplitude Modulation

(A) Advantages of FM over AM

The frequency modulation (FM) has the following advantages over AM:

(i) FM receivers may be fitted with amplitude limiters to remove the amplitude variations caused by noise.

This makes FM reception a good deal more immune to noise than AM reception.

(ii) It is possible to reduce noise still further by increasing the frequency-deviation. This is a feature which AM does not have because it is not possible to exceed 100 per cent modulation without causing severe distortion.

(iii) Standard Frequency Allocations provide a guard band between commercial FM stations. Due to this, there is less adjacent-channel interference than in AM.

(iv) FM broadcasts operate in the upper VHF and UHF frequency ranges at which there happens to be less noise than in the MF and HF ranges occupied by AM broadcasts.

(v) The amplitude of the FM wave is constant. It is thus independent of the modulation depth, whereas in AM, modulation depth governs the transmitted power. This permits the use of low-level modulation in FM transmitter and use of efficient class *C* amplifiers in all stages following the modulator. Further since all amplifiers handle constant power, the average power handled equals the peak power. In AM transmitter the maximum power is four times the average power. Finally in FM, all the transmitted power is useful whereas in AM, most of the power is carrier power which does not contain any information.

(B) Disadvantages of FM over AM

Following are the disadvantages of FM over AM:

(i) A much wider channel typically 200 kHz is required in FM as against only 10 kHz in AM broadcast. This forms serious limitation of FM.

(ii) FM transmitting and receiving equipments particularly used for modulation and demodulation tend to be more complex and hence more costly.

4.19. FM Generation

The FM modulator circuits used for generating FM signals may be put into two categories as under.

(i) The direct method or parameter variation method.

(ii) The indirect method or the Armstrong method.

4.19.1. The Direct Method or Parameter Variation Method

In direct method or parameter variation method, the baseband or modulating signal directly modulates the carrier. The carrier signal is generated with the help of an oscillator circuit. This oscillator circuit uses a parallel tuned L-C circuit. Thus the frequency of oscillation of the carrier generation is governed by the expression

$$\omega_c = \frac{1}{\sqrt{LC}}$$

Now, we can make the carrier frequency w_c to vary in accordance with the baseband or modulating signal $x(t)$ if L or C is varied according to $x(t)$. An oscillator circuit whose frequency is controlled by a modulating voltage is called voltage controlled oscillator (VCO). The frequency of VCO is varied according to the modulating signal simply by putting a shunt voltage variable capacitor with its tuned circuit. This voltage variable capacitor is called varactor or varicap. This type of property is

exhibited by reverse biased semiconductor diodes. Also the capacitance of bipolar transistors (BJT) and field-effect transistors (FET) is varied by the Miller-effect. This miller capacitance may be utilized for frequency modulation.

In addition to this, the electron tubes may also provide variable reactance (either it is inductive or capacitive) which is proportional to modulating or baseband signal. This type of tubes are called reactance tubes and may be used for FM generation.

The inductance *L* of the tuned circuit may also be varied in accordance with the baseband or modulating signal $x(t)$. The FM circuit using such inductors is called saturable reactor modulator. Frequency modulation can also be achieved from voltage controlled devices such as PIN diode, Klystron oscillators and multivibrators.

4.19.2. Varactor Diode Method for FM Generation

The varactor diode is a semiconductor diode whose junction capacitance changes with d.c. bias voltage. This varactor diode is connected in shunt with the tuned circuit of the carrier oscillator. This arrangement is shown in figure 4.16.

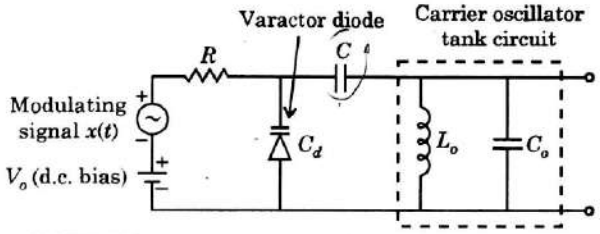


Fig. 4.16. *Varactor diode method of FM generation.*

In varactor diode FM generation arrangement, the capacitor *C* is made much smaller than the varactor diode capacitance C_d so that the radio frequency (RF) voltage from oscillator across the diode is small as compared to reverse bias d.c. voltage across the varactor diode. In addition to this, the reactance of the capacitor *C* at the highest modulating frequency is made large enough compared to resistor *R* so that the shunting of the baseband or modulating signal through the tuned circuit may be checked.

Mathematical Analysis

The capacitance C_d of the varactor diode is expressed as

$$C_d = \frac{k}{\sqrt{v_D}} = k(v_D)^{-1/2} \quad ...(4.77)$$

Here, v_D is the total instantaneous voltage across the varactor diode and is given by

$$v_D = V_o + x(t) \quad ...(4.78)$$

Also, *k* is a constant of proportionality.

The oscillation frequency is given as

$$\omega_c = \frac{1}{\sqrt{LC}} \quad ...(4.79)$$

Now, the total capacitance of the oscillator tank circuit will be $C_o + C_d$ and thus the instantaneous frequency of oscillation w_i is expressed as

$$\omega_i = \frac{1}{\sqrt{L_o(C_o + C_d)}} \quad ...(4.80)$$

In above equation, substituting the value of C_d from equation, we have

$$\omega_i = \frac{1}{\sqrt{L_o(C_o + kv_D^{-1/2})}}$$

We conclude that the instantaneous frequency w_i of FM signal depends upon v_D which in turn depends upon the value of the modulating signal $x(t)$. Thus, the instantaneous oscillator frequency w_i also depends upon the baseband or modulating signal $x(t)$ and hence frequency modulation is generated.

4.20. Drawbacks of Direct Method for FM Generation

Following are the drawbacks of the direct method:

(i) In direct method of FM generation, it is not easy to get a high order stability in carrier frequency. This is due to the fact that generation of carrier signal is directly affected by the baseband or modulating signal. The baseband signal directly controls the tank circuit of the carrier generator and thus a stable oscillator circuit (*i.e.* crystal oscillator) cannot be used. This means that the carrier generation cannot be of high stability which is a necessary requirement.

A solution to this problem is the indirect method *i.e.* Armstrong method of FM generation. In fact, in this method, the carrier oscillator does not respond to the modulating signal directly rather the generation of the carrier is made isolated from other parts of the circuit. Thus stable crystal oscillators may be used to generate carrier signal.

(ii) The non-linearity of the varactor diode produces a frequency variation due to harmonics of the modulating or baseband signal and therefore the FM signal is distorted. We will have to take the proper cares to keep this type of distortion minimum.

Despite all above drawbacks, the direct method is utilized for high power FM generation in several applications.

4.21. The Indirect or Armstrong Method of FM Generation

(U.P. Tech. Semester, Exam. 2003-04)

In Armstrong method of FM generation, we can get very high frequency stability since in this case the crystal oscillator may be used as a carrier frequency generator. The working principle of Armstrong method is to generate a narrowband FM (NBFM) indirectly by utilizing the phase-modulation technique and then changing this narrowband FM into a wideband FM as shown in figure 4.17. Since in narrowband FM the modulation index is small, therefore the distortion is very low in narrowband FM. Here we prefer phase modulation technique because its generation is easy. The multiplier circuit apart from multiplying the carrier frequency also increases the frequency deviation and hence the narrowband FM (with small frequency deviation) is converted into wideband FM (with large frequency deviation).

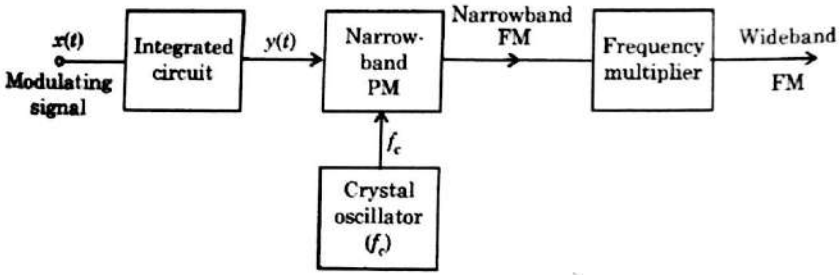


Fig. 4.17. *Armstrong method for FM generation.*

4.22. Practical Armstrong Method for FM Generation

In this section, we shall discuss a practical method of generating FM signal. Figure 4.18 shows a simplified block diagram of a commercial FM generation system using Armstrong Method.

We know that in commercial use we require to transmit audio signal consisting frequencies in the range of 50 Hz to 15 kHz and the value of Δf = 75 kHz. Let us assume that the final carrier frequency of the FM required is f_c = 100 MHz.

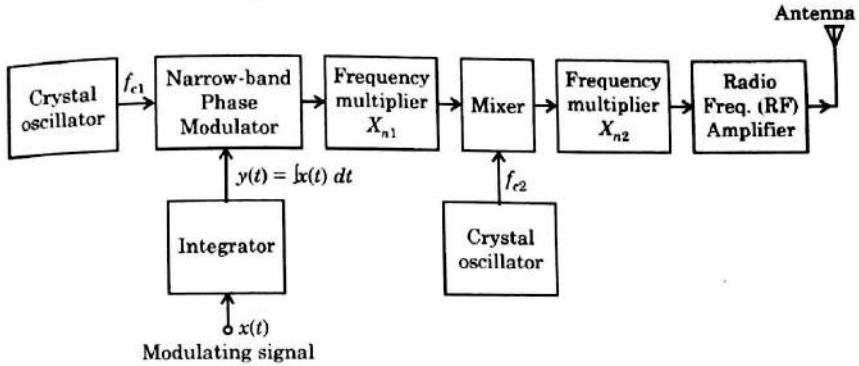


Fig. 4.18.

Let us begin with narrowband FM having a carrier frequency f_{c1} = 100 kHz generated by a crystal oscillator. To check the harmonic distortion produced by narrowband phase modulator, the modulation index m_f to a maximum value of 0.3 radians.

Let us assume m_{f1} = 0.2 radian. The lowest modulation frequency of 100 Hz produces a frequency deviation of $\Delta_{f1} = 0.2 \times 50 = 10$Hz at the narrowband phase modulator output whereas the largest modulating frequency 15 kHz produces a frequency deviation of

$$\Delta_{f2} = 0.2 \times 15 \text{ kHz} = 3 \text{ kHz} \quad ...(4.81)$$

Thus, the lowest modulating frequency is of immediate concern. We take the value of Δ_{f1} = 10 Hz, so that at the highest modulating frequency, m_f becomes even less.

To produce a frequency deviation of Δf = 75 kHz at the output, a frequency multiplication is required. As an example Δf_1 = 10 Hz and the required deviation is Δf = 75 kHz. We thus require a total frequency multiplication by a factor

$$n = \frac{75000}{10} = 7500 \quad ...(4.82)$$

However, a straight forward frequency multiplication equal to this value leads to a very high value of carrier frequency than the required 100 MHz. To get the required frequency-deviation and carrier frequency we use a two-stage frequency multiplier as shown in figure 4.18. This arrangement makes use of two multipliers and a mixer. With the help of a mixer, the carrier frequency is translated suitably without altering frequency-deviation Δf. The final stage multiplier provides the required carrier frequency and deviation.

Let us assume that n_1 and n_2 are the frequency multiplication factors for the two multipliers, so that

$$n = n_1 \cdot n_2 \quad ...(4.83)$$

$$= \frac{\Delta f}{\Delta f_1} = \frac{75000}{10} = 7500 \quad ...(4.84)$$

The carrier frequency at the output of first multiplier is translated downwards to frequency $(f_2 - nf_1)$ by mixing it with a carrier wave of frequency f_2 which is produced by another oscillator.

The carrier frequency at the input of the second multiplier is $\frac{f_c}{n_2}$.

Hence, $f_2 - n_1 f_1 = \frac{f_C}{n_2}$

Thus with $f_1 = 0.1$ MHz

and $f_2 = 8.5$ MHz, we have

$$8.5 - 0.1\, n_1 = \frac{100}{n_2}$$

Using equation (4.78), we have

$$n_1 = 100 \text{ and } n_2 = 75$$

4.23. FM Demodulators or Detectors

The process of getting a modulating or baseband signal from a frequency modulated (FM) signal is called demodulation or detection. The electronic circuits which perform the demodulation process are known as FM demodulators or detectors. The FM demodulator or detector performs the extraction of modulating signal in two steps as follows:

(*i*) It converts the frequency-modulated (FM) signal into a corresponding amplitude modulated (AM) signal with the help of frequency dependent circuits, *i.e.*, the circuits whose output voltage depends upon the input frequency. These circuits are generally known as **frequency discriminators.**

(*ii*) The original modulating or baseband signal is recovered from this AM signal with the help of the linear diode detector or envelope detector.

A simpel R-L circuit may be used as a discriminator, but this circuit has a very poor sensitivity as compared to a tuned L-C circuit. Therefore, L-C circuits are generally used as frequency-discriminators.

4.24. Types of FM Demodulators

FM demodulators are of following types:

1. Slope detector

The principle of operation of slope detectors depends upon the slope of the frequency response characteristics of a frequency selective network. The two main FM detectors, which use detuned resonant circuits come under this category, are as under:

(*i*) Single-tuned detector circuit or simple slope detector.

(*ii*) Stagger-tuned detector circuit or balanced slope detector.

2. Phase Difference Detectors

The following two circuits come under this category:

(*i*) Foster-Seeley detector

(*ii*) Ratio detector

4.24.1. Slope Detector

Figure 4.19 shows the circuit diagram of a slope detector.

The circuit consists of a tuned circuit which is slightly detuned from the carrier frequency w_c. In other words, the circuits uses two tuned circuits which are tuned to two different freqencies. First one is tuned to the incoming FM carrier frequency w_c whereas the second one is tuned to a frequency slightly different from the carrier frequency w_c. Therefore, this portion of the circuit which contains two tuned circuits tuned to different frequencies, is called discriminator.

This circuit converts the FM singal into an AM signal as shown in the slope detector characteristic curve. The another portion of the circuit is envelope detector. The AM signal from the output of the discriminator is applied at the input of envelope detector.

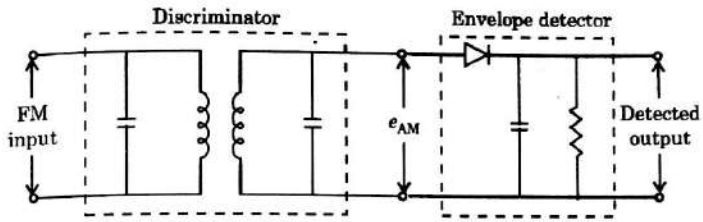


Fig. 4.19. *Circuit diagram of a slope detector.*

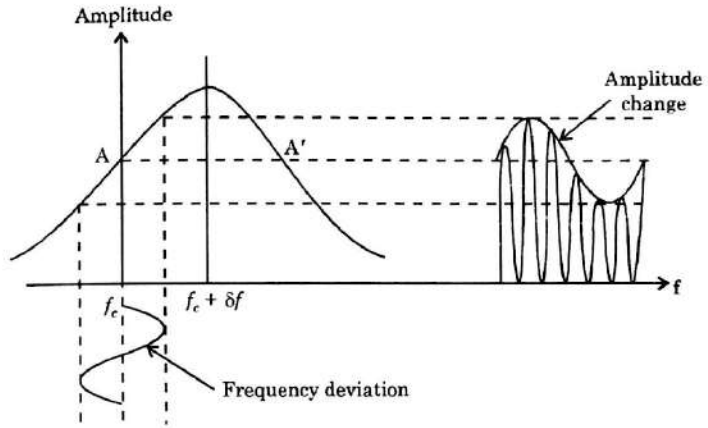


Fig. 4.20. *Slope detector characteristic curve.*

At the output of envelope detector, the original modulating or baseband signal is obtained. Although this circuit is simple and inexpensive, it has following drawbacks.

(*i*) The circuit's non-linear characteristic produces a harmonic distortion. The non-linearity is obvious from the fact that the slope is not the same at each point of the characteristics.

(*ii*) The circuit does not eliminate the amplitude variations and the output is sensitive to any amplitude variations in the input FM signal which is obviously not a desirable feature. A good discriminator circuit must respond only to frequency variations and not to amplitude variations.

4.24.2. Balanced Slope Detector *(U.P. Tech., Semester, Exam., 2004-05) (05 marks)*

The balanced slope detector is an improvement over the simple slope detector. This circuit of balanced slope detector consists of two L-C circuits as shown in the figure 4.21. The frequency response for the circuit has been shown in figure 4.22.

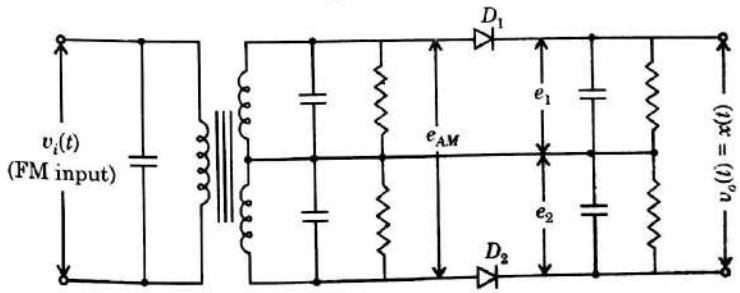


Fig. 4.21. *Circuit diagram of a balanced slope detector.*

third phasor diagram. Hence the amplitude of the voltage V_{a1} and V_{a2} will vary with the instantaneous frequency f in the manner shown in figure 4.25.

The RF voltage V_{a1} and V_{a2} are separately rectified by the diodes D_1 and D_2 respectively to produce voltage V_{01} and V_{02}. The RF components are by passed by the capacitors leaving only modulating frequency component and a d.c. term. The voltage V_{01} and V_{02} then represent the amplitude variations of V_{a1} and V_{a2} respectively. The diodes are so arranged that the output voltage V_0 is equal to the arithmetic difference $|V_{02}| - |V_{01}|$.

i.e., $V_0 = |V_{02}| - |V_{01}|$

Hence, the voltage V_0 will vary with instantaneous frequency in accordance with the difference $|V_{a2}| - |V_{a1}|$, as shown by the dotted curve in figure 4.25(c). This dotted curve is known as discriminator characteristic in figure 4.25.

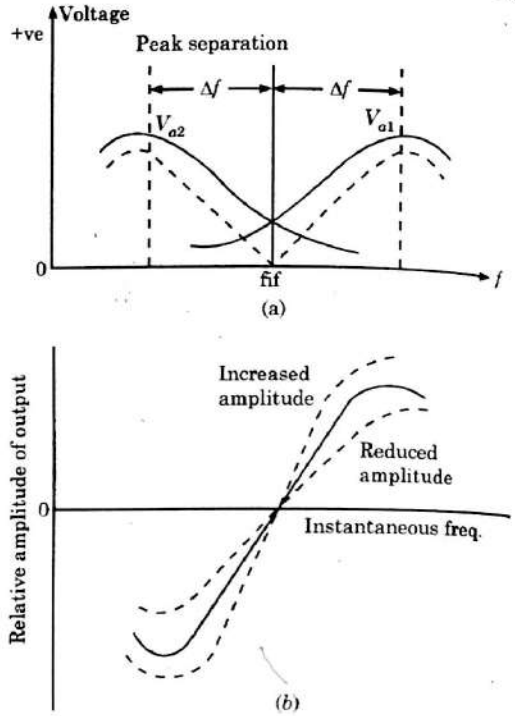


Fig. 4.25. *Discriminator characteristics.*

4.25.2. Ratio Detector

Ratio detector is an improvement over the Foster-Seeley discriminator and is most widely used. Since it does not respond to amplitude variations, therefore limiter is not needed. The circuit diagram of a ratio detector has been shown in figure 4.26. The circuit is similar to the circuit of Foster-Seeley discriminator (figure 4.26), except the following:

(i) The polarity of diode D_2 has been reversed.

(ii) The output V_o is taken from the center tap of a resistor R which shunts the load impedance of the two diodes. The output voltage varies with the input signal frequency (FM) exactly in the same way as it does in the Foster-Seeley discriminator, however its magnitude is reduced to half. This can be explained as under:

For instant, we shall ignore the two resistances (both denoted by R) and the capacitance C. The voltage V_{o1} and V_{o2} have the same magnitude as in the case of a Foster-Seeley discriminator, but V_{o2} is now reversed in polarity. Therefore, the voltage V_R is now sum of V_{o1} and V_{o2} unlike in the Foster-Seeley case, where it was the difference of V_{o2} and V_{o1}. Hence,

$$V_R = |V_{o1}| + |V_{o2}| \qquad ...(4.85)$$

The output voltage V_0 is taken across the terminal t_1 and t_2

From the circuit diagram, we have

$$V_o = V_{t1,\,t2} = |V_{t,\,t1}| - |V_{t,\,t2}| = |V_{o2}| - \left|\frac{V_R}{2}\right|$$

But $$V_R = |V_{o1}| + |V_{o2}|$$

Thus, $$V_o = V_{o2} - \frac{V_{o1} - V_{o2}}{2} = \frac{V_{o2} - V_{o1}}{2}$$

Thus, the ratio detector has exactly same behaviour except that its output is reduced.

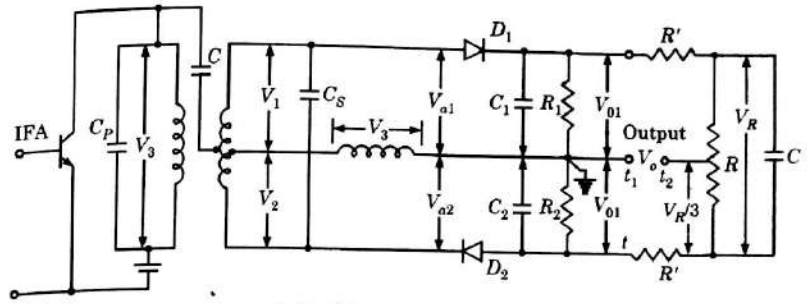


Fig. 4.26. *Ratio detector.*

4.25.3. PLL-FM Demodulator

A phase-locked loop (PLL) is primarily used in tracking the phase and frequency of the carrier component of an incoming FM signal. PLL is also useful for synchronous demodulation of AM-SC (*i.e.*, Amplitude Modulation with Suppressed Carrier) signals or signals with few cycles of pilot carrier. Further, PLL is also useful for demodulating FM signals in presence of large noise and low signal power.

This means that, PLL is most suitable for use in space vehicle-to-earth data links or where the loss along the transmission line or path is quite large. Recently, it has found application in commercial FM receivers.

A Phase-Locked Loop (PLL) is basically a negative feedback system. It consists of three major components. These components are multiplier, a loop filter and a voltage controlled oscillator (VCO) connected together in the from of a feedback loop. A VCO is a sine wave generator whose frequency is determined by the voltage applied to it from an external source.

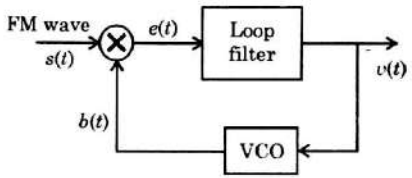


Fig. 4.27. *The block diagram of a Phase-Locked Loop (PLL).*

It means that any frequency modulator can work as a VCO.

Operation

The operation of a PLL is similar to any other feedback system. In any feedback system, the feedback signal tends to follow the input signal. If the signal fedback is not equal to the input signal, the error signal will change the value of the fedback signal until it is equal to the input signal. The difference signal between $s(t)$ and $b(t)$ is called an **error signal**. A PLL operates on a similar principle except for the fact that the quantity fedback is not the amplitude, but a generalised phase $\phi(t)$.

The error signal or difference signal $e(t)$ is utilised to adjust the VCO frequency in such a way that the instantaneous phase angle comes close to the angle of the incoming signal $s(t)$. At this point, the two signals $s(t)$ and $b(t)$ are in synchronism and the PLL is locked to the incoming signal $s(t)$.

The following mathematical steps would help us understand how FM demodulation of detection can be performed by using PLL.

Here, we have assumed that the VCO is adjusted initially so that when the control voltage comes to zero, the following two conditions are satisfied:

(i) The frequency of the VCO is precisely set at the unmodulated carrier frequency f_c and

(ii) The VCO output has a 90° phase-shift w.r.t. the unmodulated carrier wave.

Let the input signal applied to the PLL be an FM wave. It is defined as

$$s(t) = A \sin [\omega_c t + \phi_1(t)] \quad ...(4.86)$$

where A is the unmodulated carrier amplitude and $\omega_c = 2\pi f_c$ = Angular carrier frequency

and
$$\phi_1(t) = 2\pi k_f \int_0^t x(t)\, dt \quad ...(4.87)$$

where $x(t)$ is the message or baseband signal or modulating signal and

k_f = frequency sensitivity of frequency modulator.

Let the VCO output be defined by

$$b(t) = A_v \cos [\omega_c t + \phi_2(t)] \quad ...(4.88)$$

where A_v = Amplitude of VCO output when the control voltage applied to the VCO is denoted by $v(t)$, then, we have

$$\phi_2(t) = 2\pi k_v \int_0^t v(t)\, dt \quad ...(4.89)$$

Here, k_v is the frequency sensitivity of VCO, measured in Hertz/volt.

It may be observed from equations (4.86) and (4.88) that the VCO output and the incoming signals are 90° out of phase, while the VCO frequency in absence of $v(t)$ is precisely equal to the unmodulated frequency of the FM signal. The incoming FM have $s(t)$ and the VCO output $b(t)$ are applied to a multiplier.

The output of the multiplier has the following components:

(*i*) A high frequency component represented by

$$k_m A A_v \sin [2\omega_c t + \phi_1(t) + \phi_2(t)] \quad ...(4.90)$$

(*ii*) A low frequency component represented by

$$k_m A A_v \sin [\phi_1(t) - \phi_2(t)] \quad ...(4.91)$$

where k_m = Multiplier Gain measured in per volt.

The high frequency component can be eliminated by using a filter. Hence, discarding the high frequency component, the effective input to the low pass filter (LPF) will be given by

$$e(t) = k_m A A_v \sin[\phi_1(t) - \phi_2(t)]$$

or
$$e(t) = k_m A A_v \sin[\phi_e(t)] \quad ...(4.92)$$

where $\phi_e(t)$ is the phase error and is expressed as

$$\phi_e(t) = \phi_1(t) - \phi_2(t)$$

This means that

$$\phi_e(t) = \phi_1(t) - 2\pi k_v \int_0^t v(t)\, dt \quad ...(4.93)$$

The loop filter operates on error signal $e(t)$ to produce the output $v(t)$. It is given by

$$v(t) = \int_{-\infty}^{\infty} e(\tau)\, h(t-\tau)\, d\tau \quad ...(4.94)$$

where $h(t)$ = Impulse response of the low-pass filter (LPF).

Using equations (4.92), (4.93) and (4.94), we get

$$\phi_e(t) = \phi_1(t) - 2\pi k_m k_v A A_v \int_0^t \int_{-\infty}^{\infty} \sin[\phi_e(\tau)]\, h(t-\tau)\, d\tau dt$$

or
$$\phi_e(t) = \phi_1(t) - 2\pi k_o \int_0^t \int_{-\infty}^{\infty} \sin[\phi_e(\tau)]\, h(t-\tau)\, d\tau dt \quad ...(4.95)$$

where
$$k_o = k_m k_v A A_v \quad ...(4.96)$$

Now, differentiating both sides of equation (4.95), we obtain

$$\frac{d\phi_e(t)}{dt} = \frac{d\phi_1(t)}{dt} - 2\pi k_o \int_{-\infty}^{\infty} \sin[\phi_e(\tau)]\, h(t-\tau)\, d\tau \quad ...(4.97)$$

Here k_o has the dimension of frequency. On the basis of equation (4.97), we can construct and equivalent model of PLL. It has been shown in figure 4.28.

In this model, $v(t)$ and $e(t)$ are also included utilising the relationship between them as given in equations $e(t) = k_m A A_v \sin[\phi_e(t)]$ and

$$v(t) = \int_{-\infty}^{\infty} e(\tau)]\, h(t-\tau)\, d\tau.$$

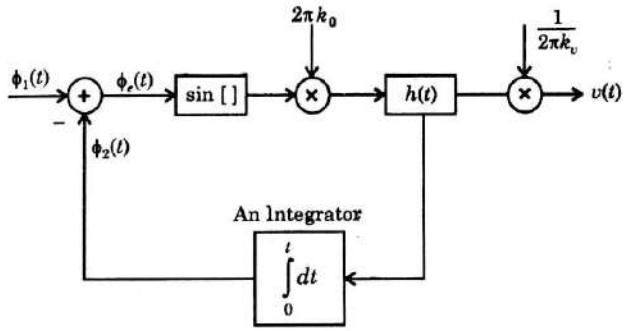


Fig. 4.28. *A non-linear equivalent model of PLL.*

A comparison of figures 4.27 and 4.28 reveals the fact that they are similar except for the fact that the multiplier in the equivalent model has been replaced by a subtractor and a sinusoidal non-linearity and the VCO by an integrator.

When the phase error $\phi_e(t)$ is zero, then PLL is said to be phase-locked. When the phase error $\phi_e(t)$ at all times is small compared to 1 radian, then we can approximate $\sin[\phi_e(t)]$ as $\phi_e(t)$, *i.e.*,

$$\sin[\phi_e(t)] \cong \phi_e(t) \quad ...(4.98)$$

It is fairly accurate as long as $\phi_e(t)$ is less than 0.5 radian. In this case, PLL is said to be Near-Lock Condition and the sinusoidal non-linearity can be disregarded. The linearised model of PLL is valid under above-mentioned condition as shown in figure 4.29*a*. In this model, phase error $\phi_e(t)$ is related to the input phase $\phi_1(t)$ by the Integro-differential equation. It is expressed as

$$\frac{d\phi_e(t)}{dt} + 2\pi k_o \int_{-\infty}^{\infty} \phi_e(\tau)\, h(t-\tau)\, d\tau = \frac{d\phi_1(t)}{dt} \quad ...(4.99)$$

Taking the Fourier transform of both sides of equation (4.99), we obtain

$$\Phi_e(f) = \frac{1}{1 + k_o \dfrac{H(f)}{jf}} \Phi_1(f) \quad ...(4.100)$$

where $\Phi_e(f)$ and $\Phi_1(f)$ are the Fourier transform of $\phi_e(t)$ and $\phi_1(t)$, respectively and $H(f)$ is the Fourier transform of impulse response $h(t)$ and is known as transfer function of the loop filter.

The quantity $k_0 H(f)/jf$ is called the open loop transfer function of the PLL.

$$L(f) = \frac{k_0 H(f)}{jf} \qquad \text{...(4.100a)}$$

Substituting equation (4.100a) in equation (4.100), we obtain

$$\Phi_e(f) = \frac{1}{1 + L(f)} \Phi_1(f) \qquad \text{...(4.10}$$

Now, let us consider that for all values of frequency f inside the baseband signal, we make the magnitude of $L(f)$ very large compared to unity.

Thus, from equation (4.101), we obtain

$$\Phi_e(f) \to 0 \text{ as } L(f) >> 1 \qquad \text{...(4.102)}$$

Under above-mentioned condition, the phase of the VCO becomes asymptotically equal to the phase of the incoming wave and the phase lock is thereby established.

From figure 4.29(a), we observe that $V(f)$ is related to $\Phi_e(f)$ by

$$V(f) = \frac{k_0}{k_v} H(f) \Phi_e(f) \qquad \text{...(4.10}$$

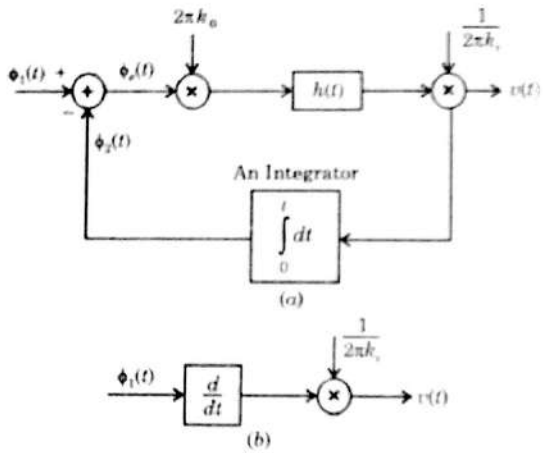


Fig. 4.29. *Illustration of an equivalent model of PLL.*

Since, we have

$$L(f) = k_0 \frac{H(f)}{jf} \qquad \text{...(4.10}$$

Above equation (4.104), may be obtained by taking Fourier transform of both sides of equation given below:

$$v(t) = \int_{-\infty}^{\infty} e(\tau)\, h(t - \tau)\, d\tau$$

and using the approximate of equation, $\sin[\phi_e(t)] \cong \phi_e(t)$.

Now, substituting the values of $H(f)$ from equation (4.104) in equation (4.103), we obtain

$$V(f) = \frac{k_0}{k_v}\left\{\frac{jf}{k_0} L(f)\right\} \Phi_e(f) = \frac{jf}{k_v} L(f) \Phi_e(f) \qquad \text{...(4.10}$$

Substituting the values of $\Phi_e(f)$ from equation (4.101) in equation (4.105), we have

$$V(f) = \frac{jf}{k_v} L(f) \Phi_e(f) = \frac{jf}{k_v} L(f) \frac{1}{1 + L(f)} \Phi_1(f)$$

or

$$V(f) = \left(\frac{jf}{k_v}\right)\left(\frac{L(f)}{1 + L(f)}\right) \Phi_1(f) \qquad \text{...(4.106)}$$

If $|L(f)| >> 1|$ then equation (4.106) may be approximated as

$$V(f) = \left(\frac{jf}{k_v}\right)[1]\, \Phi_1(f) = \left(\frac{jf}{k_v}\right) \Phi_1(f) \qquad \text{...(4.107)}$$

The corresponding time-domain representation of equation (4.107) can be obtained by taking inverse Fourier transform of both sides of equation (4.107). Hence, we have

$$\text{Inverse FT } [V(f)] = \text{Inverse } FT\left[\left(\frac{jf}{k_v}\right) \Phi_1(f)\right]$$

or

$$v(t) = \frac{1}{2\pi k_v} \frac{d\phi_1(t)}{dt} \qquad \text{...(4.108)}$$

Conclusion

Here, it can be concluded that if the magnitude of $L(f)$ is very large for all frequencies of interest, then PLL may be modeled as a differentiator with its output scaled by a factor $1/2\pi K_v$. It has been illustrated in figure 4.29(b). The simplified model of PLL shown in figure 4.29(b) may be used as an FM demodulator. This can be easily verified by substituting the value of $\phi_1(t)$ from equation $\phi_1(t) = 2\pi k_f \int_0^t x(t)\, dt$ into equation (4.108).

Therefore, we have

$$v(t) = \frac{1}{2\pi k_v} \frac{d\phi_1 t}{dt} = \frac{1}{2\pi k_v} \frac{d}{dt}\left\{2\pi k_f \int_0^t x(t)\, dt\right\}$$

or

$$v(t) = \frac{2\pi k_f}{2\pi k_v} \frac{d}{dt}\left\{\int_0^t x(t)\, dt\right\} = \frac{k_f}{k_v} x(t) \qquad \text{...(4.109)}$$

Hence, we can say that output $v(t)$ of PLL is approximately same except for a scaling factor k_f/k_v as the original baseband or modulating signal $x(t)$ and the frequency demodulation is performed.

Note: *It may be noted that the incoming FM wave can have much wider bandwidth than that of the loop filter with transfer function H(f). Here H(f) is restricted to baseband. The control signal of the VCO has a bandwidth of the baseband signal whereas the output of the VCO is a wide-band FM wave. The instantaneous frequency of the WBFM wave tracks the incoming FM. The complexity of a PLL is obtained by the transfer function H(f) of the loop filter. If H(f) = 1 then PLL is called simplest PLL. This means that there is no loop filter and the PLL is referred to as first-order PLL. The order of the PLL is determined by the order of the denominator polynomial of the closed loop transfer function. The transfer function H(f) determines the output Fourier transform V(f) in terms of input Fourier transform $\Phi_1(f)$. It is given by equation (4.107).*

MISCELLANEOUS SOLVED EXAMPLES

Example 4.13. Determine the permissible range in maximum modulation index for

(i) Commerical FM which has 30 Hz to 15 kHz modulating frequencies.

(ii) narrowband FM system which allows maximum deviation of 10 kHz and 100 Hz to 3 kHz modulating frequencies.

Solution: (i) The maximum deviation in commerical FM is given as

$$\Delta f = 75 \text{ kHz}$$

Modulation index in FM is

$$m_f = \frac{\Delta f}{f_m} \quad \text{...(i)}$$

Modulation index for commerical FM at $f_m = 30$ Hz is

$$m_f = \frac{\Delta f}{f_m} = \frac{75 \times 10^3}{30} = 2500$$

Modulation index for commerical FM at $f_m = 15$ kHz is

$$m_f = \frac{\Delta f}{f_m} = \frac{75 \times 10^3}{15 \times 10^3} = 5$$

Hence, the modulation index for commerical FM varies between 2500 and 5.

(ii) For a given Narrowband FM system, the maximum frequency deviation is given as $\Delta f = 10$ kHz.

Hence, modulation index for a given NBFM system varies between

$$m_f = \frac{\Delta f}{f_m} = \frac{10 \times 10^3}{100} = 100$$

and

$$m_f = \frac{\Delta f}{f_m} = \frac{10 \times 10^3}{3 \times 10^3} = 3.33 \quad \textbf{Ans.}$$

Example 4.14. A 100 MHz carrier wave has a peak voltage of 5 volts. The carrier is frequency modulated (FM) by a sinusoidal modulating signal or waveform of frequency 2 kHz such that the frequency deviation Δf is 75 kHz. The modulated waveform passes through zero and is increasing at $t = 0$. Determine the expression for the modulated carrier waveform.

Solution: Because the frequency modulated carrier waveform passes through zero and is increasing at $t = 0$, therefore, the FM signal must be sine wave (signal). Thus

$$s(t) = A\sin\left[2\pi f_c t + m_f \sin(2\pi f_m t)\right] \quad \text{...(i)}$$

where m_f = modulation index of FM wave $= \dfrac{\Delta f}{f_m}$

Given that f_c = Carrier wave frequency $= 100 \times 10^6 = 10^8$ hertz

Δf = frequency deviation $= 75$ kHz $= 75 \times 10^3$ hertz

f_m = modulating frequency $= 2$ kHz $= 2 \times 10^3$ hertz

A = peak voltage of carrier wave = 5 volt.

Now, $$m_f = \frac{\Delta f}{f_m} = \frac{75 \times 10^3}{2 \times 10^3} = 37.5$$

Substituting all the above values in equation (i), we get

$$s(t) = 5\sin\left[2\pi \times 10^8 t + 37.5\sin(2\pi \times 2 \times 10^3 t)\right]$$

or

$$s(t) = 5\sin\left[2\pi \times 10^8 t + 37.5\sin(4\pi \times 10^3 t)\right] \quad \textbf{Ans.}$$

Example 4.15. A carrier wave of frequency 1 GHz and amplitude 3 volts is frequency modulated (FM) by a sinusoidal modulating signal frequency of 500 Hz and of peak amplitude 1 volt. The frequency deviation Δf is 1 kHz. The level of the modulating waveform (signal) is changed to 5 volt peak and the modulating frequency is changed to 2 kHz. Obtain the expression for the new modulated waveform (FM).

Solution: We know that the FM wave is given by the expression

$$s(t) = A\cos\left[2\pi f_c t + m_f \sin(2\pi f_m t)\right]$$

where m_f = Modulation index of FM wave $= \dfrac{\Delta f}{f_m}$

and Δf = frequency deviation $= k_f A_m$

k_f = Sensitivity of frequency modulator

A_m = Amplitude of the modulating signal

Given that $f_c = 1$ MHz $= 1 \times 10^6$ Hz

$A = 3$ volt

$A_m = 1$ volt

and $\Delta f = 1$ kHz

Therefore, k_f can be found as

$$k_f = \frac{\Delta f}{f_m} = \frac{1 \times 10^3}{1} = 10^3 \text{ Hz/volt}$$

Now, for the second case, we have

when, $A_m = 5$ volt and $f_m = 2$ kHz

Modulation Index will be

$$m_f = \frac{\Delta f}{f_m} = \frac{10^3 \times 5}{2 \times 10^3} = 2.5$$

The desired FM signal can be expressed by

$$s(t) = A\cos\left[2\pi f_c t + m_f \sin(2\pi f_m t)\right]$$

Substituting all the values, we get

$$s(t) = 3\cos\left[2\pi 10^6 t + 2.5\sin(2\pi \times 2 \times 10^3 t)\right]$$

or

$$s(t) = 3\cos\left[2\pi 10^6 t + 2.5\sin(4\pi \times 10^3 t)\right] \quad \textbf{Ans.}$$

Example 4.16. Given a signal

$$s(t) = \cos(2\pi f_c t) + 0.2\cos(2\pi f_m t)\sin(2\pi f_c t)$$

(i) Prove that $s(t)$ is a combination of AM-FM signal

(ii) Draw the phasor diagram at $t = 0$. *(U.P. Tech., Sem., Exam., 2004-05) (05 marks)*

Solution: (i) The given signal $s(t)$ can be modified in the following form:

$$s(t) = \cos(2\pi f_c t) + 0.2\cos(2\pi f_m t)\sin(2\pi f_c t)$$

$$s(t) = \left[1 + \{0.2\cos(2\pi f_m t)\}^2\right]^{1/2} \times \cos\left[\left\{2\pi f_c t - \tan^{-1}(0.2\cos(2\pi f_m t))\right\}\right]$$

$$s(t) \cong \left[1 + \frac{0.04}{2}\cos(2\pi f_m t)\right]\cos[(2\pi f_c t) - 0.2\cos(2\pi f_m t)]$$

or $$s(t) = [1 + 0.01\cos(4\pi f_m t)]\cos[(2\pi f_c t) - 0.2\cos(2\pi f_m t)]$$

From equation (i), we know that the amplitude as well as the instantaneous phase angle of the given signal changes in accordance with the modulating or message signal. Hence, the given signal s(t) is a combination of AM-FM signal.

(ii) For the construction of phasor diagram, we can express the signal s(t) as under:

$$s(t) = \cos(2\pi f_c t) + 0.2\cos(2\pi f_m t)\cos(2\pi f_c t)$$

or $$s(t) = \cos(2\pi f_c t) + \frac{0.2}{2}\cos[2\pi(f_c + f_m)t] + \frac{0.2}{2}\cos[2\pi(f_c - f_m)t]$$

Phasor diagram for equation (ii) has been shown in figure 4.30.

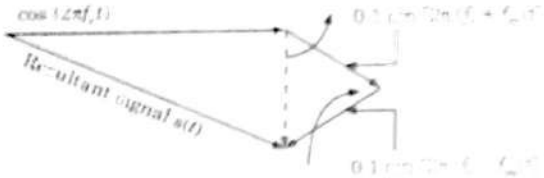


Fig. 4.30. *Illustration of Phasor diagram for example 4.19*

Example 4.17. (i) Show with a phasor diagram that the signal s(t) expressed by

$$s(t) = \cos(2\pi \times 10^6 t) + 0.02\cos\left[2\pi(10^6 + 10^3)t\right]$$

represents a carrier wave which is modulated both in amplitude and frequency.

(ii) Represent s(t) in the form

$$s(t) = \left[1 + m\cos(2\pi \times 10^6 t)\right]\cos\left[2\pi \times 10^6 t + m_f \sin(2\pi \times 10^3 t)\right]$$

Find the value of m and m_f. Also, write an expression for instantaneous frequency as a function of time t. Verify that both amplitude and frequency vary approximately sinusoidally with frequency 1 kHz.

Solution: (i) The phasor diagram of the signal

$$s(t) = \cos(2\pi \times 10^6 t) + 0.02\cos\left[2\pi(10^6 + 10^3)t\right]$$

has been shown in figure 4.31.

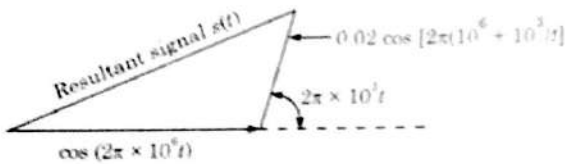


Fig. 4.31. *Phasor diagram.*

Observation

It may be easily observed from this figure that the resultant amplitude differs from the carrier ampltude A and the resultant is not in phase with the carrier. Hence, the given signal is modulated both in amplitude and phase. Hence Proved.

(ii) The given signal s(t) can be expressed as

$$s(t) = \cos(2\pi \times 10^6 t) + 0.02\cos\left[2\pi(10^6 + 10^3)t\right]$$

or $$s(t) = \cos(2\pi \times 10^6 t) + 0.02\cos(2\pi \times 10^6 t)\cos(2\pi \times 10^3 t) - 0.02\sin(2\pi \times 10^6 t)\sin(2\pi \times 10^3 t)$$

or $$s(t) = \left[1 + 0.02\cos(2\pi \times 10^3 t)\right]\cos(2\pi \times 10^6 t) - 0.02\sin(2\pi \times 10^6 t)\sin(2\pi \times 10^3 t)$$

$$s(t) = \left\{\left[1 + 0.02\cos(2\pi \times 10^3 t)\right]^2 + \left[0.02\sin(2\pi \times 10^3 t)\right]^2\right\}^{1/2} \times \cos\left\{2\pi \times 10^6 t + \tan^{-1}\frac{0.02\sin(2\pi \times 10^3 t)}{1 + 0.02\cos(2\pi \times 10^3 t)}\right\}$$

or $$s(t) \cong \left[1 + 0.02\cos(2\pi \times 10^3 t)\right]\cos\left[2\pi \times 10^6 t + 0.02\sin(2\pi \times 10^3 t)\right]$$

or $$s(t) \cong \left[\left[1 + m\cos(2\pi \times 10^3 t)\right]\cos\left[2\pi \times 10^6 t + m_f \sin(2\pi \times 10^3 t)\right]\right]$$

Hence, Modulation index of AM part, $m = 0.02$ and Modulation Index of FM part, $m_f = 0.02$

The instantaneous frequency is found as

$$f_i = \frac{1}{2\pi}\frac{d}{dt}\left[\theta_i(t)\right] = \frac{1}{2\pi}\frac{d}{dt}\left[2\pi \times 10^6 t + 0.02\sin(2\pi \times 10^3 t)\right]$$

or $$f_i = 10^6 + 20\cos(2\pi \times 10^3 t)$$

It can be observed that the amplitude and instantaneous frequency of modulated signal vary with $\cos(2\pi \times 10^3 t)$. **Ans.**

Example 4.18. Given a Narrowband Frequency Modulated (NBFM) signal as

$$s(t) = \cos[\omega_c t + \phi(t)]$$

where ω_c **= angular carrier frequency =** $2\pi f_c$

and $\phi(t) = m_1 \sin(2\pi f_1 t) + m_2 \sin(2\pi f_2 t)$

so that $|\phi(t)| << \frac{\pi}{2}$

Show that the superposition principle is satisfied by NBFM.

Solution: The Narrowband FM (NBFM) is given by

$$s(t) = \cos[\omega_c t + \phi(t)] \quad \text{...(i)}$$

where $\omega_c = 2\pi f_c$

This signal can be considered to be simultaneously modulated by two modulating signals of frequency f_1 and f_2 and thus can be written as under:

$$s(t) = \cos(\omega_c t)\cos[\phi(t)] - \sin(\omega_c t)\sin[\phi(t)] \quad \text{...(ii)}$$

For $|\phi(t)| << \frac{\pi}{2}$, we have

$$\cos[\phi(t)] \cong 1 \quad \text{and}$$

$$\sin[\phi(t)] \cong \phi(t) \quad \text{...(iii)}$$

Substituting equation (iii) in equation (ii), we get

$$s(t) = \cos(\omega_c t)1 - \sin(w_c t)\,\phi(t)$$

or $$s(t) = \cos(\omega_c t) - \sin(w_c t)\,\phi(t)$$

or $$s(t) = \cos(2\pi f_c t) - \sin(2\pi f_c t)\,\phi(t) \quad \text{...(iv)}$$

Now, substituting equation (*i*) in equation (*ii*), we get

$$s(t) = \cos(2\pi f_c t) - \sin(2\pi f_c t)\,\phi(t)$$

or

$$s(t) = \cos(2\pi f_c t) - \sin(2\pi f_c t)\left[m_1 \sin(2\pi f_1 t) + m_2 \sin(2\pi f_2 t)\right] \quad ...(iii)$$

Therefore, superposition principle holds for this NBFM. In other words, NBFM can be considered approximately, as a linear modulation scheme.

Example 4.19. A sinusoidal carrier wave of frequency of f_c = 100 MHz is exponentially modulated by a modulating signal $x(t)$. It has been shown in figure 4.32. Find the bandwidth when the resulting signal is

(*i*) FM with sensitivity $k_f = 10^5$ Hertz/volt.

(*ii*) PM with sensitivity $k_p = 10\pi$ radians/volt.

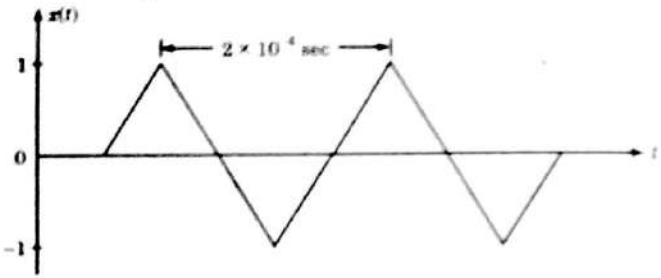


Fig. 4.32. *For example 4.19.*

Solution: We know that for FM, the instantaneous frequency is given by

$$f_i = f_c + k_f\,x(t) = f_c + 10^5\,x(t) \quad ...(i)$$

or

$$[f_i]_{max} = f_c + 10^5[x(t)]_{max} \quad ...(ii)$$

From figure 4.30, we have $x(t)_{max} = 1$.

Substituting $[x(t)]_{max} = 1$ in equation (*ii*), we get

$$[f_i]_{max} = f_c + 10^5\,[x(t)]_{max} = f_c + 10^5 \times 1 = f_c + 10^5 = 100 \times 10^6 + 10^5$$

or

$$[f_i]_{max} = 100.1 \times 10^6 \text{ Hertz} = 100.1 \text{ MHz}$$

Again, from figure 4.30, $[x(t)]_{min} = -1$

Substituting the $[x(t)]_{min} = -1$ in equation (*i*), we get

$$[f_i]_{min} = 100 \times 10^6 + 10^5(-1) = 100 \times 10^6 - 0.1 \times 10^6$$

$$= 99.9 \times 10^6 \text{ Hertz} = 99.9 \text{ MHz}$$

Frequency deviation Δf can be found as

$$\Delta f = [f_i]_{max} - f_c = (101.1 - 100) \text{ MHz}$$

$$= 0.1 \text{ MHz} = 10^5 \text{ Hertz}$$

The bandwidth of the resulting signal may be calculated using Carson's rule as under:

$$BW = 2\Delta f\left(1 + \frac{f_m}{\Delta f}\right) = 2\Delta f\left(1 + \frac{1}{D}\right) \quad ...(iii)$$

where $D = \dfrac{\Delta f}{f_m}$

f_m = Essential bandwidth of the modulating signal $x(t)$. This can be determined by expanding $x(t)$ shown in figure 4.30 in Fourier series.

The modulating signal $x(t)$ can be expanded as

$$x(t) = \sum_{n=-\infty}^{\infty} C_n \cos(2\pi n f_o t) \quad ...(iv)$$

where,

$$f_o = \frac{1}{T_o} = \frac{1}{2 \times 10^{-4}} = 5000 \text{ Hz} = 5 \text{ kHz}$$

and Fourier series coefficient

$$C_n = \begin{cases} \dfrac{8}{\pi^2 n^2}; & \text{for odd values of } n \\ 0 \quad ; & \text{for even values of } n \end{cases} \quad ...(v)$$

Now, it may be observed that the amplitudes of the various harmonic frequency components present in $x(t)$ decreases rapidly with increasing n. The powers of third harmonic and fifth harmonic are 1.21% and 0.16% of the power of the fundamental frequency. It is thus reasonable to assume that the essential bandwidth of modulating signal $x(t)$ is $3f_0 = 3 \times 5 \times 10^3 = 1.5 \times 10^4$ Hertz.

Hence, the bandwidth of frequency modulated (FM) wave is given by

$$BW = 2\Delta f\left(1 + \frac{f_m}{\Delta f}\right) = 2 \times 10^5\left(1 + \frac{1.5 \times 10^4}{2 \times 10^5}\right)$$

or

$$BW = 2 \times 10^5\left(1 + \frac{1.5}{20}\right) = 230 \times 10^3 \text{ Hz} = 230 \text{ kHz}$$

For the phase modulated case, we have

$$\theta_i(t) = 2\pi f_c t + k_p x(t)$$

Instantaneous frequency is found as

$$f_i = \frac{1}{2\pi}\frac{d\theta_i(t)}{dt} = \frac{1}{2\pi}\frac{d}{dt}\left[2\pi f_c t + k_p x(t)\right]$$

or

$$f_i = f_c + \frac{k_p}{2\pi}\frac{d}{dt}[x(t)] \quad ...(vi)$$

Maximum and minimum values of f_i will be

$$[f_i]_{max} = 100 \times 10^6 + \frac{10\pi}{2\pi}\frac{d}{dt}[x(t)]\Big|_{max}$$

or

$$[f_i]_{max} = 100 \times 10^6 + 5 \times 20{,}000 = 100 \times 10^6 + 0.1 \times 10^6$$

or

$$[f_i]_{max} = 100.1 \times 10^6 \text{ Hertz} = 100.1 \text{ MHz}$$

Further, $[f_i]_{max} = 99.9$ MHz

Hence, frequency deviation will be

$$\Delta f = [f_i]_{max} - f_i = [100.1. - 100] \text{ MHz} = 0.1 \text{ MHz} = 10^5 \text{ Hertz.}$$

Note: *The values of* $\left.\dfrac{d[m(t)]}{dt}\right|_{max}$ *has been obtained by noting that* $\dfrac{d[m(t)]}{dt}$ *switches back and forth between −20,000 and 20,000.*

The bandwidth of the PM signal is given by

$$BW_{PM} = 2\Delta f\left(1 + \frac{f_m}{\Delta f}\right) = 2 \times 10^5\left(1 + \frac{1.5 \times 10^4}{2 \times 10^5}\right) = 230 \text{ kHz} \quad \textbf{Ans.}$$

Example 4.20. Find the bandwidth of commercial FM transmission assuming frequency deviation Δf = 75 kHz and bandwidth of modulating signal $x(t)$, f_m = 15 kHz.

Solution: The deviation ratio for commerical FM transmission is given by

$$D = \frac{\Delta f}{f_m} = \frac{75 \text{ kHz}}{15 \text{ kHz}} = 5$$

Bandwidth of FM signal can be calculated by using Carson's rule as

$$BW = 2\Delta f\left(1 + \frac{f_m}{\Delta f}\right) = 2\Delta f\left(1 + \frac{1}{D}\right)$$

or $$BW = 2 \times 75 \times 10^3\left(1 + \frac{1}{5}\right) = 150 \times 10^3\left(\frac{6}{5}\right)$$

or $$BW = 180 \times 10^3 \text{ Hertz} = 180 \text{ kHz}$$

Also, using universal curve, replacing m_f by D, we get

$$BW = 3.2 \times \Delta f \text{ for } D = 5 = 3.2 \times 75 \times 10^3$$
$$= 240 \times 10^3 \text{ Hertz} = 240 \text{ kHz}$$

Now, we can find per centage of underestimation of bandwidth by using Carson's rule as

% under estimation of bandwidth will be

$$= \frac{240-180}{240} \times 100 = \frac{60}{240} \times 100 = 25\%$$

It means that Carson's rule underestimates the bandwidth by 25% as compared with the result obtained from the universal curve. **Ans.**

Example 4.21. A rule of bandwidth for FM signal is sometimes used as

$$BW = (2m_f + 1)f_m$$

Find the fraction of the signal power that is included in that frequency band. Assume that $m_f = 1$ and 10.

Solution: The bandwidth for FM signal can be calculated on the basis of 98% power requirement given by Carson's rule as

$$BW = 2\Delta f\left(1 + \frac{1}{m_f}\right) = 2(m_f + 1)\, f_m \quad \text{...(i)}$$

The fraction of signal power included in the frequency band B is

$$P = \frac{B}{BW} \times \left(\frac{98}{100}\right) \quad \text{...(ii)}$$

For $m_f = 1$ $$P = \frac{B}{BW} \times \left(\frac{98}{100}\right) = \frac{B}{BW} \times 0.98$$

or $$P = \frac{(2m_f + 1)f_m}{2(m_f + 1)f_m} \times 0.98 = \frac{2 \times 1 + 1}{2(1+1)} \times 0.98$$

or $$P = \frac{2 \times 1 + 1}{2(1+1)} \times 0.98 = \frac{3}{4} \times 0.98 = 73.5\%$$

For $m_f = 10$ $$P = \frac{B}{BW} \times 0.98 = \frac{(2m_f + 1)f_m}{2(m_f + 1)f_m} \times 0.98$$

or $$P = \frac{(2 \times 10 + 1)}{2(10+1)} \times 0.98$$

or $$P = \frac{21}{22} \times 0.98 = 97.6\% \quad \textbf{Ans.}$$

Example 4.22. A carrier is frequency modulated (FM) by a sinusoidal modulating signal $x(t)$ of frequency 2 kHz, it results in a frequency deviation Δf of 5 kHz. Find the bandwidth occupied by the FM waveform. The amplitude of the modulating sinusoid is increased by a factor of 3 and its frequency lowered by 1 kHz. Find the new bandwidth.

Solution: Given that

$$f_m = 2 \text{ kHz and}$$
$$\Delta f = 5 \text{ kHz}$$

Hence, the bandwidth of the FM signal will be

$$BW = 2(\Delta f + f_m) = 2(5 + 2)$$

or $$BW = 14 \text{ kHz}$$

When the amplitude of the modulating signal is trippled, then the frequency deviation increases three times.

Therefore, $$\Delta f = 3 \times 5 \text{ kHz} = 15 \text{ kHz}$$

Also, $$f_m = 1 \text{ kHz}$$

The new bandwidth will be

$$BW' = 2(\Delta f + f_m) = 2(15 + 1)$$

or $$BW' = 32 \text{ kHz} \quad \textbf{Ans.}$$

Example 4.23. Determine the relative power of the carrier wave and side frequencies when modulation index $m_f = 0.20$ for 10 kW FM transmitter.

Solution: For $$m_f = 0.20$$
$$J_0(m_f) = 0.99$$

(See Table 4.1 of Bessel Functions)

Thus, only significant side frequency pair is $f_c \pm f_m$ with relative amplitude,

$$J_1(m_f) = 0.099$$

Hence, the carrier power will be

$$P_c = J_0^2(m_f) \times \text{Power of FM transmitter}$$
$$P_c = (0.99)^2 \times 10 \times 10^3 = 9.8 \times 10^3 \text{ watt}$$
$$P_c = 9.8\, K \text{ watt} \quad \textbf{Ans.}$$

and the power in each side frequency is expressed as

$$P_{s1} = P_{s2}$$
$$= J_1^2(m_f) \times \text{Power of FM transmitter}$$
$$= (0.099)^2 \times 10 \times 10^3 = 98 \text{ watt.} \quad \textbf{Ans.}$$

Example 4.24. In an FM, system a 7 kHz modulating (or baseband) signal modulates 107.6 MHz carrier wave so that the frequency deviation is 50 kHz. Find

(i) carrier swing in the FM signal and modulating index m_f

(ii) the highest and lowest frequencies attained by the FM signal.

Solution: (i) Given that frequency deviation

$$\Delta f = 50 \text{ kHz} = 50 \times 10^3 \text{ Hertz}$$

Carrier swing in FM signal will be

$$= 2\Delta f = 2 \times 50 \times 10^3 \text{ Hz} = 100 \text{ kHz}$$

Modulation Index of FM wave,

$$m_f = \frac{\Delta f}{f_m} = \frac{50 \times 10^3 \text{ Hertz}}{7 \times 10^3 \text{ Hertz}} = 7.143$$

(*ii*) The upper or highest frequency attained by FM signal will be

$$= f_c + \Delta f$$
$$= 107.6 \times 10^6 + 50 \times 10^3 = 107.65 \text{ MHz}$$

The lower or lowest frequency attained by FM signal will be

$$= f_c - \Delta f = 107.6 \times 10^6 - 50 \times 10^3$$
$$= 107.55 \text{ MHz} \quad \textbf{Ans.}$$

Example 4.25. Determine the frequency deviation Δf and carrier swing for an FM signal which has a carrier frequency of 100 MHz and whose upper frequency is 100.007 MHz when modulated by a particular modulating signal or wave. Also find the lowest frequency reached by the FM wave.

Solution: We know that frequency deviation Δf is defined as the maximum change in frequency of the modulated signal away from the carrier frequency f_c.

This means that

i.e. $$\Delta f = f_v - f_c = [100.007 - 100.000] \text{ MHz}$$
$$= 0.007 \text{ MHz} = 7 \times 10^3 \text{ Hz} = 7 \text{ kHz}$$

Carrier swing $= 2\Delta f = 2 \times 7 = 14$ kHz

The lowest frequency f_L reached by the modulated FM wave is equal to the difference of the frequency deviation from the carrier frequency.

$$f_L = f_c - \Delta f = (100.00 - 0.007) \text{ MHz}$$

or $$f_L = 99.993 \text{ MHz} \quad \textbf{Ans.}$$

Example 4.26. Determine the modulation index m_f of an FM signal which is being broadcast in the 88 – 108 MHz band. This FM wave has a carrier swing of 125 kHz.

Solution: We know that the frequency deviation is given by

$$\Delta f = \frac{\text{Carrier swing}}{2} = \frac{125 \times 10^3}{2} \text{ Hz}$$

or $$\Delta f = 62.5 \times 10^3 \text{ Hz} = 62.5 \text{ kHz}$$

Since, maximum frequency deviation for the FM broadcast band is 75 kHz, therefore

$$m_f = \text{modulation index}$$

$$m_f = \frac{\Delta f}{f_m} = \frac{62.5}{75} = 83.3\% \quad \textbf{Ans.}$$

Example 4.27. The modulating signal in an FM wave is 500 Hz with amplitude 3.2 volt and frequency deviation is 6.4 kHz. If the audio frequency voltage is now increased to 8.4 volt, determine the new frequency deviation and modulation index. If the audio frequency voltage is raised to 20 volts while the audio frequency is dropped to 200 Hz, determine the frequency deviation and modulation index.

Solution: Given that $\Delta f = 64$ kHz

$$f_m = 500 \text{ Hz}$$
$$V_m = 3.2 \text{ volt}$$

We know that,

Frequency sensitivity, $k_f = \dfrac{\Delta f_1}{V_{m1}} = \dfrac{6.4 \text{ kHz}}{3.2 \text{ volt}} = 2$ kHz/volt

Frequency deviation for $V_m = 8.4$ volt will be determined as

$$\Delta f_2 = k_f V_{m2} = 2 \times 8.4 = 16.8 \text{ kHz.}$$

Frequency deviation for $V_m = 20$ volt will be expressed as

$$\Delta f_3 = k_f V_{m3} = 2 \times 20 = 40 \text{ kHz}$$

It may be observed that the change in modulating frequency made no difference to the frequency deviation because it is independent of the modulating frequency.

The modulation indices can be calculated as under:

$$m_1 = \frac{\Delta f_1}{f_{m1}} = \frac{6.4 \text{ kHz}}{0.5 \text{ kHz}} = 12.8$$

$$m_2 = \frac{\Delta f_2}{f_{m2}} = \frac{16.8 \text{ kHz}}{0.5 \text{ kHz}} = 33.6$$

$$m_3 = \frac{\Delta f_3}{f_{m3}} = \frac{40 \text{ kHz}}{0.2 \text{ kHz}} = 200 \quad \textbf{Ans.}$$

Example 4.28. An FM wave is given by

$$s(t) = 20 \sin (6 \times 10^8 t + 7 \sin 1250t). \textbf{ Determine}$$

(*i*) The carrier and modulating frequencies, the modulation index, and the maximum deviation.

(*ii*) Power dissipated by this FM wave in a 100 ohm resistor.

Solution: (*a*) The standard expression for FM is

$$s(t) = A \sin \left[\omega_c t + m_f \sin (\omega_m t)\right] \quad \text{...(i)}$$

Given expression is

$$s(t) = 20 \sin \left[6 \times 10^8 t + 7 \sin 1250t\right] \quad \text{...(ii)}$$

On comparing equations (*i*) and (*ii*), we obtain

$$f_c = \frac{\omega_c}{2\pi} = \frac{6 \times 10^8}{2\pi} = 95.5 \text{ MHz}$$

$$f_m = \frac{\omega_m}{2\pi} = \frac{1250}{2\pi} = 199 \text{ Hertz}$$

and $$m_f = 7, \Delta f = m_f f_m = 7 \times 199 = 1393 \text{ Hz} \quad \textbf{Ans.}$$

(*ii*) Power dissipated by the given FM wave in 100 ohm resistor can be calculated as under:

$$P = \frac{\left(\frac{A}{\sqrt{2}}\right)^2}{R} = \frac{\left(\frac{20}{\sqrt{2}}\right)^2}{100} = 2 \text{ watt.} \quad \textbf{Ans.}$$

Example 4.29. Find the instantaneous frequency in hertz of each of the following signals:

(*i*) $10 \cos\left(200\pi t + \dfrac{\pi}{3}\right)$

(*ii*) $10 \cos(20\pi t + \pi t^2)$

(*iii*) $\cos 200\pi t \cos(5 \sin 2\pi t) + \sin 200\pi t \sin(5 \sin 2\pi t)$

Solution: (*i*) We have

$$\theta(t) = 200\pi t + \frac{\pi}{3}$$

$$\omega_i = \frac{d\theta}{dt} = 200\pi = 2\pi(100)$$

The instantaneous frequency of the signal is 100 Hz, which is constant.

(ii) $\theta(t) = 20\pi t + \pi t^2$

$$\omega_i = \frac{d\theta}{dt} = 20\pi + 2\pi t = 2\pi(10 + t)$$

The instantaneous frequency of the signal is 10 Hz at $t = 0$ and increases linearly at a rate of 1 Hz/s.

(iii) $\cos 200\pi t \cos(5\sin 2\pi t) + \sin 200\pi t \sin(5 \sin 2\pi t) = \cos(200\pi t - 5\sin 2\pi t)$

$$\theta(t) = 200\pi t - 5 \sin 2\pi t$$

$$\omega_i = \frac{d\theta}{dt} = 200\pi - 10\pi \cos 2\pi t = 2\pi(100 - 5 \cos 2\pi t)$$

The instantaneous frequency of the signal will be 95 Hz at $t = 0$ and oscillate sinusoidally between 95 and 105 Hz.

Example 4.30. Given an angle-modulated signal

$$x_c(t) = 10 \cos[(10^8)\pi t + 5\sin 2\pi(10^3)t]$$

Determine the maximum phase deviation and the maximum frequency deviation.

Solution: Comparing the given $x_c(t)$ with standard FM wave equation, we have

$$\theta(t) = \omega_c t + \phi(t) = (10^8)\pi t + 5 \sin 2\pi(10^3)t$$

and $\phi(t) = 5 \sin 2\pi (10^3)t$

Now, $\phi'(t) = 5(2\pi)(10^3) \cos 2\pi (10^3)t$

Therefore, the maximum phase deviation will be

$$|\phi(t)|_{max} = 5 \text{ rad}$$

and the maximum frequency deviation will be

$$\Delta\omega = |\phi'(t)|_{max} = 5(2\pi)(10^3) \text{ rad/s}$$

or $\Delta f = 5$ kHz.

Example 4.31. An angle-modulated signal is described by

$$x_c(t) = 10\cos[2\pi(10^6)t + 0.1 \sin(10^3)\pi t]$$

(i) considering $x_c(t)$ as a PM signal with $k_p = 10$, obtain $m(t)$

(ii) considering $x_c(t)$ as an FM signal with $k_f = 10\pi$, find $m(t)$,

Solution: (i) We have

$$x_{PM}(t) = A \cos[\omega_c t + k_p m(t)] = 10\cos[2\pi(10^6)t + 10m(t)]$$
$$= 10 \cos[2\pi(10^6)t + 0.1 \sin(10^3)\pi t]$$

Hence, $m(t) = 0.01 \sin(10^3)\pi t$ **Ans.**

(ii) We have $x_{FM}(t) = A\cos\left[\omega_c t + k_f \int_{-\infty}^{t} m(\lambda)\, d\lambda\right] = 10\cos[2\pi(10^6)t + 0.1 \sin(10^3)\pi t]$

Assuming

$$m(t) = a_m \cos(10^3)\pi t$$

we obtain

$$10\pi \int_{-\infty}^{t} m(\lambda)\, d\lambda = 10\pi a_m \int_{-\infty}^{t} \cos(10^3)\pi\lambda\, d\lambda = \frac{a_m}{100}\sin(10^3)\pi t = 0.1 \sin(10^3)\pi t$$

Hence, $a_m = 10$, and

$$m(t) = 10 \cos(10^3)\pi t$$

Example 4.32. Let $m_1(t)$ and $m_2(t)$ be two message signals, and let $x_{c1}(t)$ and $x_{c2}(t)$ be the modulated signals corresponding to $m_1(t)$ and $m_2(t)$, respectively.

(i) Prove that if the modulation is DSB(AM), then $m_1(t) + m_2(t)$ will produce a modulated signal equal to $x_{c1}(t) + x_{c2}(t)$. (This is why AM is sometimes referred to as a linear Modulation.)

(ii) Prove that if the modulation is PM, then the modulated signal produced by $m_1(t) + m_2(t)$ will not be $x_{c1}(t) + x_{c2}(t)$, that is, superposition does not apply to angle-modulated signal. (This is why angle modulation is sometimes referred to as a nonlinear modulation.)

Solution: (i) For DSB (AM), we have

$$m_1(t) \rightarrow x_{c1}(t) = m_1(t) \cos \omega_c t$$
$$m_2(t) \rightarrow x_{c2}(t) = m_2(t) \cos \omega_c t$$

Now,

$$m_1(t) + m_2(t) \rightarrow x_c(t) = [m_1(t) + m_2(t)] \cos \omega_c t$$
$$= m_1(t) \cos \omega_c t + m_2(t) \cos \omega_c(t) = x_{c1}(t) + x_{c2}(t)$$

Therefore, DSB (AM) modulation is a linear modulation.

(ii) For PM, we have

$$m_1(t) \rightarrow x_{c1}(t) = A \cos[\omega_c t + k_p m_1(t)]$$
$$m_2(t) \rightarrow x_{c2}(t) = A \cos[\omega_c t + k_p m_2(t)]$$

Also, $m_1(t) + m_2(t) \rightarrow x_c(t) = A \cos[\omega_c t + k_p[m_1(t) + m_2(t)]\}$

$$\neq x_{c1}(t) + x_{c2}(t)$$

Therefore, PM is a nonlinear modulation.

Example 4.33. A carrier signal is angle-modulated by the sum of two sinusoids as under:

$$x_c(t) = A \cos(\omega_c t + \beta_1 \sin \omega_1 t + \beta_2 \sin \omega_2 t) \quad \text{...(i)}$$

where ω_1 and ω_2 are not harmonically related. Determine the spectrum of $x_c(t)$.

Solution: The given signal $x_c(t)$ can be expressed as under:

$$x_c(t) = A \text{ Re}[e^{j\omega_c t} e^{j(\beta_1 \sin \omega_1 t + \beta_2 \sin \omega_2 t}]$$
$$= A \text{ Re}(e^{j\omega_c t} e^{j\beta_1 \sin \omega_1 t} e^{j\beta_2 \sin \omega_2 t}) \quad \text{...(ii)}$$

We have $e^{j\beta_1 \sin \omega_1 t} = \sum_{n=-\infty}^{\infty} J_n(\beta_1)\, e^{jn\omega_1 t}$

$$e^{j\beta_2 \sin \omega_2 t} = \sum_{m=-\infty}^{\infty} J_m(\beta_1)\, e^{jm\omega_2 t}$$

Substituting these expressions into equation (ii) and taking the real part, we get

$$x_c(t) = A \sum_{n=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} J_n(\beta_1)\, J_m(\beta_2) \cos(\omega_c + n\omega_1 + m\omega_2)\, t \quad \text{...(iii)}$$

Form equation (iii) it maybe observed that the spectrum of $x_c(t)$ consists of four categories:

(1) the carrier line

(2) sideband lines at $\omega_c + n\omega_1$ due to one tone alone

(3) sideband lines at $\omega_c \pm m\omega_2$ due to the other tone alone

(4) sideband lines at $\omega_c \pm n\omega_1 \pm m\omega_2$ because of the and nonlinear property of angle modulation.

Example 4.34. In a tone-modulated angle modulation, the modulated signal $x_c(t)$ is given by

$$x_c(t) = A\cos(\omega_c t + \beta \sin \omega_m t)$$

when $\beta << 1$, we have NB angle modulation.

(*i*) Determine the spectrum of this NB angle-modulated signal.

(*ii*) Compare the result with that of a tone-modulated AM signal.

(*iii*) Discuss the similarities and difference by drawing their phasor representations.

Solution: (*i*) Given $x_c(t) = A\cos(\omega_c t + \beta \sin \omega_m t)$

$$= A\cos\omega_c t\cos(\beta\sin\omega_m t) - A\sin\omega_c t\sin(\beta\sin\omega_m t)$$

Now, when $\beta<<1$, we can write

$$\cos(\beta\sin\omega_m t) \cong 1$$

$$\sin(\beta\sin\omega_m t) \cong \beta\sin\omega_m t$$

Then the NB signal may be approximated as order:

$$x_{NBc}(t) \cong A\cos\omega_c t - \beta A\sin\omega_m t\sin\omega_c t$$

$$= A\cos\omega_c t - \frac{\beta A}{2}\cos(\omega_c - \omega_m)t + \frac{\beta A}{2}\cos(\omega_c + \omega_m)t \quad ...(i)$$

From equation (*i*) it may be observed that the spectrum of $x_{NBc}(t)$ consists of carrier line and a pair of side lines at $\omega_c \pm \omega_m$.

(*ii*) The above result is almost identical to the situation for a tone-modulated AM signal

$$x_{AM}(t) = A\cos\omega_c t + mA\cos\omega_m t\cos\omega_c t$$

$$= A\cos\omega_c t + \frac{mA}{2}\cos(\omega_c - \omega_m)t + \frac{mA}{2}\cos(\omega_c + \omega_m)t \quad ...(ii)$$

where m is the modulation index for AM.

Now, comparisons of equations (*i*) and (*ii*) reveals that the main difference between NB angle modulation and AM is the phase reversal of the lower sideband component.

(*iii*) By using

$$e^{j\beta\sin\omega_m t} \cong 1 + j\beta\sin\omega_m t \text{ for } \beta<<1$$

Equation (*i*) can be written in phasor form as under:

$$x_{NBc}(t) = \text{Re}[A e^{j\omega_c t}(1 + j\beta\sin\omega_m t)]$$

$$= \text{Re}\left[Ae^{j\omega_c t}\left(1 + \frac{\beta}{2}e^{j\omega_m t} - \frac{\beta}{2}e^{-j\omega_m t}\right)\right] \quad ...(iii)$$

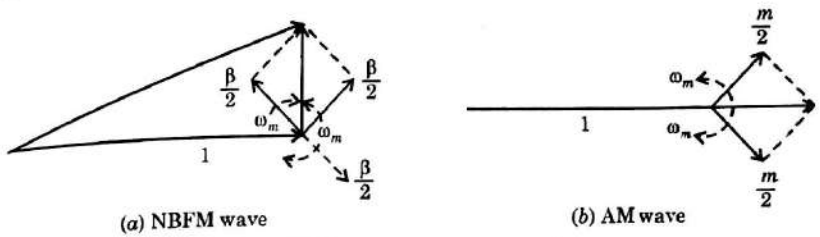
Similarly, equation (*ii*) can be written in phasor form as under:

$$x_{AM}(t) = \text{Re}[A e^{j\omega_c t}(1 + m\cos\omega_m t)]$$

or

$$x_{AM}(t) = \text{Re}\left[A e^{j\omega_c t}\left(1 + \frac{m}{2}e^{j\omega_m t} + \frac{m}{2}e^{-j\omega_m t}\right)\right] \quad ...(iv)$$

By taking the term $Ae^{j\omega_c t}$ as the reference, the phasor representations of equations (*iii*) and (*iv*) have been shown in figure 4.33. From the figure the difference between equations (*iii*) and (*iv*) is quite obvious. In narrowband angle modulation, the modulation is added in quadrature with the carrier, which results in phase fluctuation with little amplitude change. In the AM case, the modulation is added in phase with the carrier, producing amplitude fluctuation with no phase deviation.



(*a*) NBFM wave (*b*) AM wave

Fig. 4.33. *Phasor Representation.*

Example 4.35. Given the angle-modulated signal

$$x_c(t) = 10\cos(2\pi 10^8 t + 200\cos 2\pi 10^3 t)$$

what will be bandwidth of this angle modulated signal?

Solution: The instantaneous frequency is given by

$$\omega_i = 2\pi(10^8) - 4\pi(10^5)\sin 2\pi(10^3)t$$

So $\Delta\omega = 4\pi(10^5)$, $\omega_m = 2\pi(10^3)$, and

$$\beta = \frac{\Delta\omega}{\omega_m} = \frac{4\pi(10^5)}{2\pi(10^3)} = 200$$

Also, we have

Bandwidth $W_B = 2(\beta + 1)\omega_m = 8.04\pi(10^5)$ rad/s

Since $\beta >> 1$, therefore

$W_B \cong 2\Delta\omega = 8\pi(10^5)$ rad/s

or $f_B = 400$ kHz **Ans.**

Example 4.36. A 20 megahertz (MHz) carrier is frequency-modulated by a sinusoidal signal such that the maximum frequency deviation is 100 kHz. Determine the modulation index and the approximate bandwidth of the FM signal if the frequency of the modulating signal is

(*i*) 2 kHz (*ii*) 100 kHz, and (*iii*) 500 kHz.

Given $\Delta f = 100$ kHz, $f_c = 20$ MHz $>> f_m$

Solution: For sinusoidal modulation, $\beta = \Delta f/f_m$.

(*i*) With $f_m = 1$ kHz, $\beta = 100$. This is a WBFM signal, and $f_B \cong 2\Delta f = 200$ kHz.

(*ii*) With $f_m = 100$ kHz, $\beta = 1$. Thus, we have

$$f_B \cong 2(\beta + 1)f_m = 400 \text{ kHz}$$

(*iii*) With $f_m = 500$ kHz, $\beta = 0.2$. This is an NBFM signal, and $f_B \cong 2f_m = 1000$ kHz = 1 MHz.

Example 4.37. Given an angle-modulated signal

$$x_c(t) = 10\cos(\omega_c t + 3\sin\omega_m t)$$

Assume PM and $f_m = 1$ kHz. Calculate the modulation index and find the bandwidth when (*i*) f_m is double and (*ii*) f_m is decreased by one-half.

Solution: Given $x_{PM}(t) = A\cos[\omega_c t + k_p m(t)] = 10\cos(\omega_c t + 3\sin\omega_m t)$

Thus, $m(t) = a_m\sin\omega_m t$, and

$$x_{PM}(t) = 10\cos(\omega_c t + k_p a_m\sin\omega_m t)$$

We have $\beta = k_p a_m = 3$

It may be observed that the value of β is independent of f_m. Now, when $f_m = 1$ kHz, we have

$$f_B = 2(\beta + 1)f_m = 8 \text{ kHz}$$

(*i*) When f_m is doubled, $\beta = 3$, $f_m = 2$kHz, and

$$f_B = 2(3 + 1)2 = 16 \text{ kHz}$$

(*ii*) When f_m is decreased by one-half, $\beta = 3$, $f_m = 0.5$ kHz, and

$$f_B = 2(3 + 1)(0.5) = 4 \text{ kHz}$$

Example 4.38. Repeat problem (4.40) when FM is assumed *i.e.*,

$$x_{FM}(t) = A\cos\left[\omega_c t + k_f \int_{-\infty}^{t} m(\lambda)\, d\lambda\right] = 10\cos(\omega_c t + 3\sin\omega_m t)$$

Solution: Thus $m(t) = a_m \cos\omega_m t$ and

$$x_{FM}(t) = 10\cos\left(\omega_c t + \frac{a_m k_f}{\omega_m}\sin\omega_m t\right)$$

We have $$\beta = \frac{a_m k_f}{\omega_m} = \frac{a_m k_f}{2\pi f_m} = \frac{a_m k_f}{2\pi(10^3)} = 3$$

It may be observed that the value of β is inversely proportional to f_m. Thus, when $f_m = 1$ kHz, we have

$$f_B \cong 2(\beta + 1)f_m = 2(3 + 1)1 = 8 \text{ kHz}$$

(*i*) When f_m is doubled, $\beta = 3/2$, $f_m = 2$kHz, and

$$f_B = 2(\beta + 1)f_m = 2\left(\frac{3}{2} + 1\right)2 = 10 \text{ kHz}$$

(*ii*) When f_m is decreased by one-half, $\beta = 6$, $f_m = 0.5$ kHz, and

$$f_B = 2(\beta + 1)f_m = 2(6 + 1)(0.5) = 7 \text{ kHz} \qquad \textbf{Ans.}$$

Example 4.39. A carrier signal is frequency-modulated with a sinusoidal signal of 2 kHz resulting in a maximum frequency deviation of 5 kHz.

(*i*) Find the bandwidth of the modulated signal.

(*ii*) The amplitude of the modulating sinusoid signal is increased by a factor of 3, and its frequency is lowered to 1 kHz. Determine the maximum frequency deviation and the bandwidth of the new modulated signal.

Solution: (*i*) We have

$$\beta = \frac{k_f a_m}{\omega_m} = \frac{\Delta f}{f_m} = \frac{5(10^3)}{2(10^3)} = 2.5$$

Then, the bandwidth will be

$$f_B = 2(\beta + 1)f_m = 2(2.5 + 1)2 = 14 \text{ kHz}$$

(*ii*) Let β_1 be the new modulation index. Then, we have

$$\beta_1 = \frac{k_f 3a_m}{\frac{1}{2}\omega_m} = 6\frac{k_f a_m}{\omega_m} = 6\beta = 6(2.5) = 15$$

Therefore, we have

$$\Delta f = \beta_1 f_m t = (15)(1) = 15 \text{ kHz}$$
$$f_B = 2(\beta_1 + 1)f_{mt} = 2(15 + 1)(1) = 32 \text{ kHz}$$

Example 4.40. As a matter of fact, in addition to Carson's rule, the following formula is often used to estimate bandwidth of an FM signal:

$$W_B \cong 2(D + 2)\omega_m \qquad \text{for } D > 2$$

where $\omega_M = 2\pi f_M$ and f_M is the highest frequency of the signal in hertz. Calculate the bandwidth, using formula, and compare it to the bandwidth, using Carson's rule for the FM signal with $\Delta f = 75$ kHz and $f_M = 15$ kHz.

Note that commercial FM broadcast stations in the United States are limited to maximum frequency deviation of 75 kHz, and modulating frequencies typically cover 50 Hz to 15 kHz.

Solution: We have $\omega_m = 2\pi f_M$, where $f_M = 15$ kHz, we get

$$D = \frac{\Delta f}{f_M} = \frac{75(10^3)}{15(10^3)} = 5$$

and by using the given formula, the bandwidth will be

$$f_B = 2(D + 2)f_M = 210 \text{ kHz}$$

Using Carson's rule, observe that the bandwidth will be

$$f_B = 2(D + 1)f_M = 180 \text{ kHz}$$

Note: *High-quality FM radio require bandwidth of at least 200 kHz. Hence, it seems that Carson's rule underestimates the bandwidth.*

Example 4.41. Given a frequency multiplier circuit and an NBFM signal

$$x_{NBFM}(t) = A\cos(\omega_c t + \beta\sin\omega_m t)$$

With $\beta < 0.5$ and $f_c = 200$ kHz. Let f_m range from 50 Hz to 15 kHz, and let the maximum frequency deviation Δf at the output be 75 kHz. Determine the required frequency multiplication n and the maximum allowed frequency deviation at the input.

Solution: From the expression

$$\beta = \Delta f / f_m, \text{ we have}$$

$$\beta_{\min} = \frac{75(10^3)}{15(10^3)} = 5$$

and $$\beta_{\max} = \frac{75(10^3)}{50} = 1500$$

If $\beta_1 = 0.5$, where β_1 is the input β, then the required frequency multiplication will be

$$n = \frac{\beta_{\max}}{\beta_1} = \frac{1500}{0.5} = 3000$$

The maximum allowed frequency deviation at the input, denoted Δ_{f1}, will be

$$\Delta f_1 = \frac{\Delta f}{n} = \frac{75(10^3)}{3000} = 25 \text{ Hz} \qquad \textbf{Ans.}$$

Example 4.42. A block diagram of an indirect (Armstrong) FM transmitter has been shown in figure 4.34. Calculate the maximum frequency deviation Δf of the output of the FM transmitter and the carrier frequency f_c if $f_1 = 200$ kHz, $f_{LO} = 10.8$ MHz, $\Delta f_1 = 25$ Hz, $n_1 = 64$, and $n_2 = 48$.

Solution: We have $$\Delta f = (\Delta f_1)(n_1)(n_2) = (25)(64)(48) \text{ Hz} = 76.8 \text{ kHz}$$

$$f_2 = n_1 f_1 = (64)(200)(10^3) = 12.8(10^6) \text{ Hz} = 12.8 \text{ MHz}$$

$$f_3 = f_2 \pm f_{LO} = (12.8 \pm 10.8)(10^6) \text{ Hz} = \begin{cases} 23.6 \text{ MHz} \\ 2.0 \text{ MHz} \end{cases}$$

Thus, when $f_3 = 23.6$ MHz, then, we have

$$f_c = n_2 f_3 = (48)(23.6) = 1132.8 \text{ MHz}$$

When $f_3 = 2$ MHz, then, we have

$$f_c = n_2 f_3 = (48)(2) = 96 \text{ MHz}$$

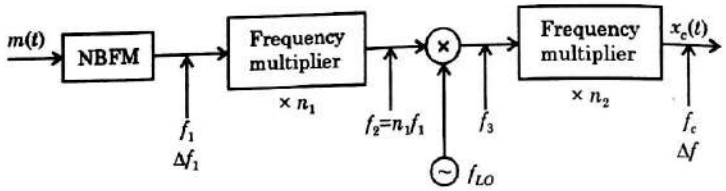


Fig. 4.34. *Block diagram of an indirect FM transmitter.*

Example 4.43. In an Armstrong-type FM generator of figure 4.32, the crystal oscillator frequency is 200 kHz. The maximum phase deviation is limited to 0.2 to avoid distortion. Let f_m range from 50 Hz to 15 kHz. The carrier frequency at the output is 108 MHz, and the maximum frequency deviation is 75 kHz. Select multiplier and mixer oscillator frequencies.

Solution: We have

$$\Delta f_1 = \beta f_m\,(0.2)(50) = 10 \text{ Hz}$$

and
$$\frac{\Delta f}{\Delta f_1} = \frac{75(10^3)}{10} = 7500 = n_1 n_2$$

$$f_2 = n_1 f_1 = n_1(2)(10^5) \text{ Hz}$$

Assuming down conversion, we have

$$f_2 - f_{LO} = \frac{f_c}{n_2}$$

Thus,
$$f_{LO} = n_1 f_1 - \frac{f_c}{n_2} = \frac{7500(2)(10^5) - 108(10^6)}{n_2} = \frac{1392}{n_2}(10^6) \text{ Hz}$$

Letting $n_2 = 150$, we get

$$n_1 = 50 \text{ and } f_{LO} = 9.28 \text{ MHz} \qquad \textbf{Ans.}$$

Example 4.44. A given modulated signal has maximum frequency deviation of 50 Hz for an input sinusoid of unit amplitude and a frequency of 120 Hz. Find the required frequency multiplication factor n to produce a maximum frequency deviation of 20 kHz when the input sinusoid has unit amplitude and a frequency of 240 Hz and the angle modulation used is (*i*) PM and (*ii*) FM.

Solution: (*i*) It may be observed that in sinusoidal PM, the maximum frequency deviation Δf is proportional to f_m. Thus, we have

$$\Delta f_1 = \left(\frac{240}{120}\right)(50) = 100 \text{ Hz}$$

Therefore
$$n = \frac{\Delta f_2}{\Delta f_1} = \frac{20(10^3)}{100} = 200$$

(*ii*) Again we observe that in sinusoidal FM, the maximum frequency deviation Δf is independent of f_m. Thus, we have

$$n = \frac{\Delta f_2}{\Delta f_1} = \frac{20(10^3)}{50} = 400 \qquad \textbf{Ans.}$$

Example 4.45. At low carrier frequencies, it may be possible to generate an FM signal by varying the capacitance of a parallel resonant circuit. Prove that the output $x_c(t)$ of the tuned circuit shown in figure 4.35 is an FM signal if the capacitance has a time dependence of the form

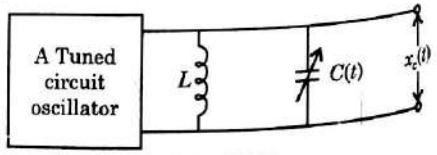


Fig. 4.35.

$$C(t) = C_0 - km(t)$$

and
$$\left|\frac{k}{C_0}\,m(t)\right| << 1$$

Solution: If we assume $km(t)$ is small and slowly varying, then the output frequency ω_i of the oscillator will be given by

$$\omega_i = \frac{1}{\sqrt{LC(t)}} = \frac{1}{\sqrt{L[C_0 - km(t)]}} = \frac{1}{\sqrt{LC_0}}\left[1 - \frac{k}{C_0}\,m(t)\right]^{-1/2}$$

Because $|(k/C_0)m(t)| << 1$, then, we can use the following approximation:

$$(1 - z)^{-1/2} \cong 1 + \frac{1}{2}z$$

and obtain
$$\omega_i \cong \omega_c\left[1 + \frac{1}{2}\frac{k}{C_0}m(t)\right] = \omega_c + k_f m(t)$$

where
$$\omega_c = \frac{1}{\sqrt{LC_0}} \quad \text{and } k_f = \frac{1}{2}\frac{\omega_c k}{C_0}$$

Thus, $x_c(t)$ is an FM signal **Hence Proved.**

Example 4.46. An FM signal

$$x_{FM}(t) = A\cos\left[\omega_c t + k_f \int_{-\infty}^{t} m(\lambda)\,d\lambda\right] \qquad \text{...(i)}$$

is applied to a system consisting of a high-pass RC filter and an envelope detector. Assume that $\omega RC << 1$ in the frequency band occupied by $x_{FM}(t)$. Determine the output signal $y(t)$, assuming that $k_f|m(t)| < \omega_c$ for all t.

Solution: Taking the Fourier transform of both sides of equation (*i*) yields

$$Y(\omega) = \frac{1}{\tau}[X(\omega) - e^{-j\omega\tau}X(\omega)] = \frac{1}{\tau}X(\omega)(1 - e^{-j\omega\tau})$$

If $\omega\tau << 1$, then $1 - e^{-j\omega\tau} \cong j\omega\tau$ and

$$Y(\omega) \cong j\omega X(\omega)$$

which indicates that $y(t)$ is approximately equal to the derivative of $x(t)$ and τ must satisfy the following condition:

$$\tau << \frac{1}{\omega} = \frac{1}{\omega_c + \Delta\omega} \qquad \textbf{Ans.}$$

Example 4.47. The sinusoidal modulating wave

$$m(t) = A_m \cos(2\pi f_m t)$$

is applied to a phase modulator with phase sensitivity k_p. The unmodulated carrier wave has frequency f_c and amplitude A_c.

(*i*) **Determine the spectrum of the resulting phase-modulated signal, assuming that the maximum phasor deviation $\beta_p = k_p A_m$ does not exceed 0.3 radians.**

(*ii*) **Construct a phase diagram for this modulated signal, and compare it with that of the corresponding narrowband FM signal.** *(U.P.S.C., I.E.S., Examination, 1998)*

Solution: (*i*) A phase-modulated wave is expressed as,

$$s(t) = A_c \cos[\omega_c t + k_p m(t)] \qquad \text{...(i)}$$

where, $m(t) = A_m \cos(2\pi f_m t)$

then, $s(t) = A_c \cos[2\pi f_c t + k_p A_m \cos(2\pi f_m t)]$

$= A_c \cos[2\pi f_c t + \beta_p \cos(2\pi f_m t)]$

where $\beta_p = k_p A_m$ known as modulation index for PM.

$= A_c \cos(2\pi f_c t).\cos[\beta_p \cos(2\pi f_m t) - A_c \sin(2\pi f_c t) + \sin[\beta_p \cos(2\pi f_m t)]$

Now, if $\beta_p \le 0.5$, then, we have

$\cos[\beta_p \cos(2\pi f_m t)] \cong 1$

$\sin[\beta_p \cos(2\pi f_m t)] \cong \beta_p \cos(2\pi f_m t)$

Then, we can write equation (*i*) as,

$$s(t) \cong A_c \cos(2\pi f_c t) - \beta_p A_c \sin(2\pi f_c t)\cos(2\pi f_m t)$$

$$= A_c \cos(2\pi f_c t) - \frac{1}{2}\beta_p A_c \sin[2\pi(f_c + f_m)t] - \frac{1}{2}\beta_p A_c \sin[2\pi(f_c - f_m)t] \quad ...(ii)$$

Therefore, spectrum of $s(t)$, will be

$$\cos(2\pi f_c t) \Leftrightarrow \frac{1}{2}[\delta(f - f_c) + \delta(f + f_c)]$$

and $$\sin 2\pi f_c t \Leftrightarrow \frac{1}{2j}[\delta(f - f_c) + \delta(f + f_c)]$$

Using Fourier transform pairs, we have

$$S(f) \cong \frac{1}{2}A_c[\delta(f - f_c) + \delta(f + f_c)] - \frac{1}{4j}\beta_p A_c[\delta(f - f_c - f_m) - \delta(f + f_c + f_m)]$$

$$-\frac{1}{4}\beta_p A_c[\delta(f - f_c + f_m) - \delta(f + f_c - f_m)]$$

(*ii*) The phasor diagram for $s(t)$ is deduced from equation (*ii*), as follows:

Comparing these two phasor diagrams in figure 4.35 and figure 4.36, we see that except for a phase difference the narrowband PM and FM waves have the same form.

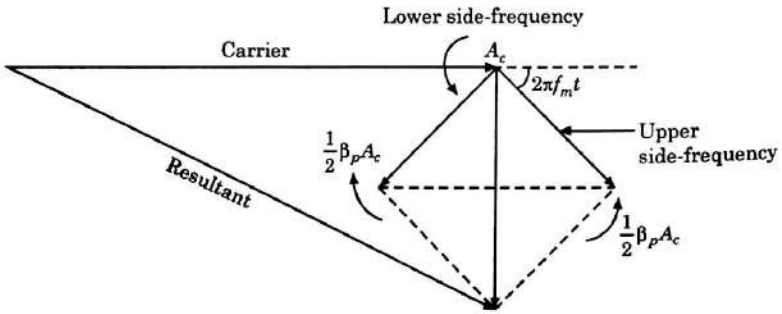


Fig. 4.36.

Phasor diagram for narrow FM has been shown in figure 4.37.

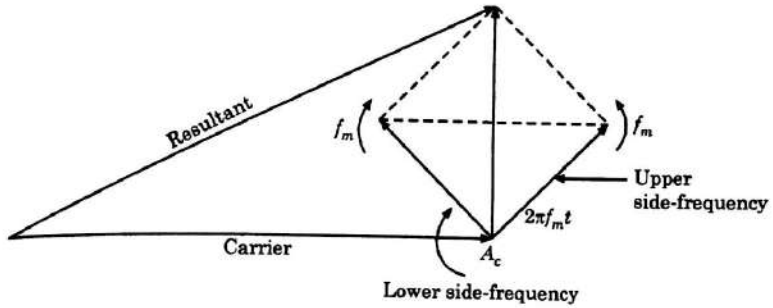


Fig. 4.37.

Example 4.48. An angle-modulated signal is given by

$$x_c(t) = 5\cos[2\pi(10^6)t + 0.2\pi t]$$

Can you identity whether $x_c(t)$ is a PM or and FM signal?

Solution: For angle modulation, the modulated carrier is represented by,

$$x_c(t) = A\cos[\omega_c t + \phi(f)]$$

In case of PM, $\phi(t) = k_p m(y)$,

where $m(t)$ is the message signal and k_p is phase duration constant.

and for FM, $$\frac{d\phi(t)}{dt} = k_f m(t)$$

where k_f is frequency deviation constant.

@sTc

Given $x_c(t) = 5\cos[2\pi(10^6 t) + 0.2\cos 200\pi t)$

Now, if $m(t) = a_m \cos(200\pi t)$

then, $\phi(t) = K_p a_m \cos(200\pi t) = 0.2\cos(200\pi t)$

and then $x_c(t)$ = PM signal.

But, if $m(t) = a_m \sin\omega_m t$

then $$\phi(t) = k_f \int_{t_0}^{f_t} m(\lambda)\,d\lambda + \phi(t_0) = k_f \int_{-\infty}^{t} m(\lambda)\,d\lambda + \phi(-\infty)$$

or $$\phi(t) = k_f \int_{-\infty}^{t} m(\lambda)\,d\lambda\,[\text{At } t_0 = -\infty, \phi(t_0) =]$$

or $$\phi(t) = \frac{k_f a_m}{\omega_m}\cos\omega_m t$$

and then $x_c(t)$ is a FM signal. Hence it can be either PM or an FM signal.

Example 4.49. A baseband signal $m(t)$ modulates a carrier to produce the following angle modulated signal.

$A_c \cos[2\pi \times 10^8 t + k_p m(t)]$, where $m(t)$ is shown in the figure 4.38.

Determine the value of k_p so that the peak-to-peak frequency deviation of the carrier is 100 kHz.

(GATE Examination-1999)

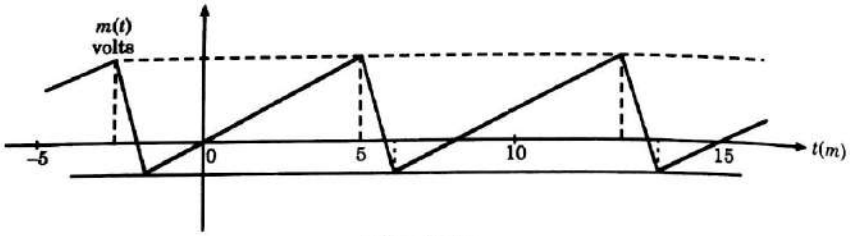


Fig. 4.38.

Solution: Given

$x(t) = A_c \cos [2\pi \times 10^8 t + k_p m(t)]$ as an angle modulated signal.

Since, phase deviation $\phi(t) = k_p m(t)$ *i.e.*, proportional to message signal, therefore, it is an phase-modulated signal.

Now, instantaneous frequency is given by

$$\omega_t = \omega_c + k_p \frac{dm(t)}{dt}$$

Then, maximum frequency deviation will be

$$|\omega_t - \omega_c|_{\max} = k_p \left|\frac{dm(t)}{at}\right|_{\max}$$

or $$100 \times 2\pi = k_p \left|\frac{dm(t)}{at}\right|_{\max} \quad ...(i)$$

from the given figure, we have

$$\left|\frac{dm(t)}{at}\right|_{\max} = \frac{17.5}{7} = 2.5$$

then, from equation (*i*), we get

$$k_p = \frac{100 \times 2\pi}{2.5} = 80\,\pi \qquad \textbf{Ans.}$$

Example 4.50. An angle modulated signal with carrier frequency $\omega_c = 2 \times 10^6$ rad/sec is given by

$$s(t) = \cos 2\pi(2 \times 10^6 t + 30 \sin 150t + 40 \cos 150t)$$

then,

(*i*) Determine the maximum frequency deviation

(*ii*) Find maximum phase deviation, and

(*iii*) Find the bandwidth of $s(t)$.

Solution: (*i*) We know that instantaneous frequency is given by

$$\omega_t = \frac{d}{dt}\theta(t) = \frac{d}{dt}[2\pi(2 \times 10^6 t + 30 \sin 150t + 40 \cos 150t)]$$

$$= 2\pi \times 10^6 + 60\pi \times 150 \,.\, \cos 150t - 80\pi \times 150 \,.\, \sin 150t$$

or $$\Delta\omega = \omega_i - \omega_c = 60\pi \times 150.\ \cos 150t - 80\pi \times 150.\ \sin 150t$$

$$= 3000\pi\,[3 \cos 150t - 4 \sin 150t]$$

$$\Delta\omega = 15000\pi \,.\, \cos (150t + \alpha),$$

Example 4.51. An angle-modulated signal is given as below:

$$s(t) = 20 \cos (\omega_c t + 4 \sin \omega_m t)$$

Assuming as a PM signal and $f_m = 2$ kHz, calculate the modulation index and bandwidth when (*i*) f_m is increased two times (*ii*) f_m is decreased by half.

Solution: A PM signal is given as,

$$s(t) = A_m \cos [\omega_c t + k_p\, m(t)]$$

or $$s(t) = 20 \cos [\omega_c t + 4 \sin \omega_m t] \quad ...(i)$$

Hence, $$m(t) = a_m \sin \omega_m t \quad ...(ii)$$

then, $$s(t) = 20 \cos[\omega_c t + k_p a_m \sin \omega_m t)$$

Comparing equations (*i*) and (*ii*), we get

$$\beta = k_p a_m = 4$$

Also, $$f_B = 2(\beta + 1) f_m = 2(4 + 1).2 = \mathbf{20\ kHz}$$

(*i*) When, $$f_{m1} = 2f_m = 4\text{kHz}$$

$$f_B = 2(4 + 1) \cdot 4 = \mathbf{40\ kHz} \qquad \textbf{Ans.}$$

(*ii*) When, $$f_{m2} = \frac{f_m}{2} = 1 \text{ kHz}$$

$$f_B = 2(\beta + 1) \cdot f_{m2} = 2(4 + 1)\ .\ 1 = 10 \text{ kHz} \qquad \textbf{Ans.}$$

SUMMARY

1. Angle modulation may be defined as the process in which the total phase angle of a carrier wave is varied in accordance with the instantaneous value of the modulating or message signal while keeping the amplitude of the carrier constant.
2. We can vary this phase angle ϕ in two ways and thus there are two types of angle modulation as under:
   (*i*) Phase Modulation (PM)
   (*ii*) Frequency Modulation (FM).
3. Phase modulation (PM) is that type of angle modulation in which the phase angle ϕ is varied linearly with a baseband or modulating signal $x(t)$ about an unmodulated phase angle $(\omega_c t + \theta_0)$.
4. This means that in Phase Modulation, the instantaneous value of the phase angle is equal to the phase angle of the unmodulated carrier $(\omega_c t + \theta_0)$ plus a time-varying component which is proportional to modulating signal $x(t)$.
5. Frequency modulation is that type of angle modulation in which the instantaneous frequency ω_i is varied linearly with a message or baseband signal $x(t)$ about an unmodulated carrier frequency ω_c. This means that the instantaneous value of the angular frequency ω_i will be equal to the carrier frequency ω_c plus a time-varying component proportional to the baseband signal $x(t)$.
6. $s(t) = A \cos \left[\omega_c t + k_f \int_0^t x(t)\, dt\right]$ which is the required general expression for FM wave.
7. The maximum change in instantaneous frequency form the average frequency ω_c is called frequency deviation.
8. This maximum change in instantaneous frequency ω_i from the average or carrier frequency ω_c depends upon the magnitude and sign of $k_f \cdot x(t)$.
9. Therefore,
   Frequency deviation $\Delta\omega = |k_f \cdot x(t)|_{\max}$
10. To get FM by using PM, we first integrate the baseband signal and then apply to the phase modulator.

11. The total variation in frequency from the lowest to the highest point is called carrier swing. Obviously

$$\text{The carrier swing} = 2 \times \text{frequency deviation}$$
$$= 2 \times \Delta\omega$$

12. The amount of frequency deviation or variation depends upon the amplitude (loudness) of the modulating (audio) signal. This means that louder the sound, greater the frequency deviation and vice versa.
13. The frequency deviation is useful in determining the FM signal bandwidth. Since, a maximum frequency deviation of 75 *KHz* is allowed for commercial FM broadcast stations using a band of 88 MHz to 108 MHz, therefore approximate FM channel width is 2 × 75 = 150 kHz. Allowing a 25 kHz guardband on either side, the channel width becomes 2(75 + 25) = 200 kHz. This guardband is meant to prevent interference between adjacent channels.
14. In FM broadcast, the highest audio frequency transmitted is 15 *KHz*.
15. For FM, the modulation index is defined as the ratio of frequency deviation to the modulating frequency.
16. Modulation index, $m_f = \dfrac{\text{Frequency deviation}}{\text{Modulation frequency}}$

or $m_f = \dfrac{\Delta\omega}{\omega_m}$

This modulation index may be greater than unity.

19. The term "per cent modulation" as it is used in reference to FM refers to the ratio of actual frequency deviation to the maximum allowable frequency deviation. Thus 100% modulation corresponds to 75 kHz for the commercial FM broadcast band and 25 kHz for the commercial FM broadcast band and 25 kHz for television.

Per cent modulation $M = \dfrac{\Delta f_{actual}}{\Delta f_{max}} \times 100$.

20. When the value of modulation index m_f is quite large, then in FM a large number of sidebands are produced and hence the bandwidth of FM is sufficiently large. This type of FM system is known as wideband FM.
21. Since the amplitude of FM signal remains unchanged, the power of the FM signal will be same as that of the unmodulated carrier.
22. There are a number of international regulations for frequency modulation prescribed by CCIR (*i.e.* Consultative Committee for International Radio). These regulations must be followed by all the commercial FM broadcast stations to avoid interference problem. They are as follows:
    (*i*) Maximum modulating frequency is 15 kHz
    (*ii*) Maximum frequency deviation is ± 75 kHz
    (*iii*) Frequency stability of the carrier is ± 2 kHz
    (*iv*) Allowable bandwidth per channel = 200 kHz.
23. The modulation index m_f is given as

$$m_f = \frac{\text{Frequency deviation}}{\text{Modulating frequency}} = \frac{\Delta w}{w_m} = \frac{k_f V_m}{w_m}.$$

24. Basically the effective bandwidth is the separation between the two extreme significant side frequencies on either side of the carrier.
25. Fourier analysis reveals that the number of side frequencies which contain a significant amount of power and thus the effective bandwidth of the FM signal is dependent on the modulation index of the modulated wave.
26. For practical purpose, a side frequency is considered to be significant if its amplitude is at least one per cent of the unmodulated carrier amplitude.

27. Carson's rule provides a thumb formula to calculate the bandwidth of a single-tone wideband FM. According to this rule, the FM bandwidth is given as twice the sum of the frequency deviation and the highest modulating frequency.
28. The transmission bandwidth for PM using Carson's rule as

$$BW_{PM} \cong 2(\Delta w)$$
$$\cong 2k_p . V_m . w_m$$

29. For PM, the modulation index m_p will be same as the deviation θ_d is expressed as

$$m_p = k_p . V_m = \theta_d$$

30. The envelope of FM wave or PM wave is constant and is equal to the unmodulated carrier amplitude. On the other hand, the envelope of AM wave is dependent on the modulating signal $x(t)$.
31. The zero crossings (*i.e.* the instants of time at which the waveform changes from negative to a positive value or vice-versa) of a FM wave or a PM wave no longer exhibit a perfect regularity in their spacing like AM wave. Thus this makes the instantaneous frequency of the angle modulated wave depend upon time.
32. FM receivers may be fitted with amplitude limiters to remove the amplitude variations caused by noise.
33. It is possible to reduce noise still further by increasing the frequency-deviation. This is a feature which AM does not have because it is not possible.
34. Standard Frequency Allocations provide a guardband between commercial FM stations. Due to this, there is less adjacent-channel interference than in AM.
35. FM broadcasts operate in the upper VHF and UHF frequency ranges at which there happens to be less noise than in the MF and HF ranges occupied by AM broadcasts.
36. The amplitude of the FM wave is constant. It is thus independent of the modulation depth, whereas in AM, modulation depth governs the transmitted power. This permits the use of low-level modulation in FM transmitter and use of efficient class *C* amplifiers in all stages following the modulator.
37. A much wider channel typically 200 kHz is required in FM as against only 10 kHz in AM broadcast. This forms serious limitation of FM.
38. FM transmitting and receiving equipments particularly used for modulation and demodulation tend to be more complex and hence more costly.
39. The FM modulator circuits used for generating FM signals may be put into two categories as under.
    (*i*) The direct method or parameter variation method.
    (*ii*) The indirect method or the Armstrong method.
40. In direct method or parameter variation method, the baseband or modulating signal directly modulates the carrier. The carrier signal is generated with the help of an oscillator circuit. This oscillator circuit uses a parallel tuned L-C circuit. Thus the frequency of oscillation of the carrier generation is governed by the expression

$$w_c = \frac{1}{\sqrt{LC}}$$

41. The varactor diode is a semiconductor diode whose junction capacitance changes with d.c. bias voltage. This varactor diode is connected in shunt with the tuned circuit of the carrier oscillator.
42. In direct method of FM generation, it is not easy to get a high order stability in carrier frequency. This is due to the fact that generation of carrier signal is directly affected by the baseband or modulating signal. The baseband signal directly controls the tank circuit of the carrier generator and thus a stable oscillator circuit (*i.e.* crystal oscillator) cannot be used. This means that the carrier generation cannot be of high stability which is a necessary requirement.
43. The non-linearity of the varactor diode produces a frequency variation due to harmonics of the modulating or baseband signal and therefore the FM signal is distorted. We will have to take the proper cares to keep this type of distortion minimum.

44. In Armstrong method of FM generation, we can get very high frequency stability since in this case the crystal oscillator may be used as a carrier frequency generator.
45. The working principle of Armstrong method is to generate a narrowband FM (NBFM) indirectly by utilizing the phase-modulation technique and then changing this narrowband FM into a wideband FM.
46. It converts the frequency-modulated (FM) signal into a corresponding amplitude modulated (AM) signal with the help of frequency dependent circuits, *i.e.*, the circuits whose output voltage depends upon the input frequency. These circuits are generally known as **frequency discriminators.**
47. The original modulating or baseband signal is recovered from this AM signal with the help of the linear diode detector or envelope detector.
48. The principle of operation of slope detectors depends upon the slope of the frequency response characteristics of a frequency selective network. The two main FM detectors, which use detuned resonant circuits come under this category, are as under:
    (*i*) Single-tuned detector circuit or simple slope detector.
    (*ii*) Stagger-tuned detector circuit or balanced slope detector.
49. The following two circuits come under this category:
    (*i*) Foster-Seeley detector
    (*ii*) Ratio detector.

SHORT QUESTIONS WITH ANSWERS

Q.1. What do you mean by angle modulations?

Ans. Angle modulation may be defined as the process in which the total phase angle of a carrier wave is varied in accordance with the instantaneous value of the modulating or message signal while keeping the amplitude of the carrier constant.

Q.2. What are the types of angle modulation?

Ans. There are two types of angle modulation as under:
(*i*) Phase Modulation (PM)
(*ii*) Frequency Modulation (FM)

Q.3. Define phase modulation?

Ans. Phase modulation (PM) is that type of angle modulation in which the phase angle ϕ is varied linearly with a baseband or modulating signal $x(t)$ about an unmodulated phase angle $(\omega_c t + \theta_0)$. This means that in Phase Modulation, the instantaneous value of the phase angle is equal to the phase angle of the unmodulated carrier $(\omega_c t + \theta_0)$ plus a time-varying component which is proportional to modulating signal $x(t)$.

Q.4. Define frequency modulation?

Ans. Frequency modulation is that type of angle modulation in which the instantaneous frequency ω_i is varied linearly with a message or baseband signal $x(t)$ about an unmodulated carrier frequency ω_c. This means that the instantaneous value of the angular frequency ω_i will be equal to the carrier frequency ω_c plus a time-varying component proportional to the baseband signal $x(t)$.

Q.5. What is frequency deviation?

Ans. The instantaneous frequency of FM signal varies with time around the carrier frequency ω_c. This means that the instantaneous frequency of FM signal varies according to the modulating signal.

The maximum change in instantaneous frequency form the average frequency ω_c is called **frequency deviation.**

Q.6. What is carrier signal?

Ans. The total variation in frequency from the lowest to the highest point is called carrier swing. Obviously

$$\text{The carrier swing} = 2 \times \text{frequency deviation}$$
$$= 2 \times \Delta\omega$$

Q.7. Define modulation index for FM?

Ans. For *FM*, the modulation index is defined as the ratio of frequency deviation to the modulating frequency.

Mathematically,

$$\text{Modulation index, } m_f = \frac{\text{Frequency deviation}}{\text{Modulation frequency}} = \frac{\Delta\omega}{\omega_m}$$

This modulation index may be greater than unity.

Q.8. What is per cent modulation an for FM?

Ans. The term "per cent modulation" as it is used in reference to FM refers to the ratio of actual frequency deviation to the maximum allowable frequency deviation. Thus 100% modulation corresponds to 75 kHz for the commercial FM broacast band and 25 kHz for the commercial FM broadcast band and 25 kHz for television.

$$\text{Per cent modulation } M = \frac{\Delta f_{actual}}{\Delta f_{max}} \times 100$$

Q.9. What is sideband FM?

Ans. When the value of modulation index m_f is quite large, then in FM a large number of sidebands are produced and hence the bandwidth of FM is sufficiently large. This type of FM system is known as wideband FM.

Q.10. What is transmission BW for FM?

Ans. For n sidebands the bandwidth of FM wave is given by

$$BW = 2\, nw_m \text{ radians/sec.} = 2nf_m \text{ Hz.}$$

Q.11. What is Carson's rule?

Ans. Carson's rule provides a thumb formula to calculate the bandwidth of a single-tone wideband FM. According to this rule, the FM bandwidth is given as twice the sum of the frequency deviation and the highest modulating frequency. However, it must be remembered that this rule is just an approximation.

Mathematically,

$$BW = 2\,(\Delta w + w_m)$$

Q.12. What are disadvantages of FM order?

Ans. A much wider channel typically 200 kHz is required in FM as against only 10 kHz in AM broadcast. This forms serious limitation of FM.

FM transmitting and receiving equipments particularly used for modulation and demodulation tend to be more complex and hence more costly.

REVIEW QUESTIONS

1. Explain the term instantaneous frequency in case of angle-modulation.
2. Define the following terms for FM wave:
   (*i*) Carrier swing
   (*ii*) Frequency deviation
   (*iii*) Per cent modulation
3. Derive an expression for a single-tone frequency-modulated wave.
4. Explain the difference between narrowband FM and wideband FM.
5. Derive an expression for a single-tone narrowband frequency-modulated wave.
6. What is meant by the term 'significant sidebands' in frequency-modulated wave.
7. What are the principal merits and limitations of FM.
8. Explain the salient features of wideband FM system.

9. State the typical applications of narrowband FM system.
10. Draw the circuit diagram of varactor-diode modulator and explain its working.
11. Explain the Armstrong method for the generation of wideband FM.

NUMERICAL PROBLEMS

1. A single-tone FM signal is given by
$$v(t) = 10 \sin (16\pi \times 10^6 t + 20 \sin 2\pi \times 10^3 t) \text{ volts.}$$
Determine the modulation index, modulating frequency, frequency deviation, carrier frequency and the power of the FM signal.

2. An FM wave, modulated to a depth of 8, generates a bandwidth of 180 kHz. Find the frequency deviation.

3. A carrier signal $10 \cos 8\pi\, 10^6 t$ is angle modulated by a modulating signal $5 \cos 30\pi\, 10^3 t$.
   (i) Find the bandwidth for frequency modulation assuming $k_f = 15$ kHz per volt.
   (ii) Assuming the same bandwidth, find k_p for phase modulation.
   (iii) Find the change in the bandwidth for frequency and phase modulation if the modulating signal becomes $5 \cos 10\pi.\, 10^3 t$.

4. Determine the carrier swing the highest and the lowest frequencies attained and the modulation index of the FM signal generated by frequency modulating a 101.6 MHz carrier with an 8 kHz. Sine wave causing a frequency deviation of 40 kHz. [Ans. 80 kHz, 101.64 MHz, 101.56 MHz, 5]

5. Calculate the frequency deviation and carrier swing of a frequency-modulated wave which was produced by modulating a 50.400 MHz carrier. The highest frequency reached by the FM wave is 50.406 MHz. Then calculate the lowest frequency reached by the FM wave.

6. In an FM system, the frequency-deviation is 4 kHz when the audio modulating frequency is 200 Hz and the audio modulating voltage is 4V. Compute the modulation index. Also compute the frequency deviation and modulating index if
   (i) AF voltage is increased to 8 V and modulating frequency is increased to 800 Hz.
   (ii) AF voltage is increased to 12 V and modulating frequency is decreased to 100 Hz.

7. An angle-modulated signal is given by
$$x_c(t) = 5 \cos[2\pi(10^6)\, t + 0.2 \cos 200\pi t]$$
Can you identify whether $x_c(t)$ is a PM or an FM signal ?
[Ans. No. It can be either a PM or an FM signal]

8. The frequency multiplier is a nonlinear device followed by a bandpass filter, as shown in figure 4.39. Suppose that the nonlinear device is an ideal square-law device with input-output characteristics.
$$e_o(t) = a e_{i2}(t)$$
Find the output $y(t)$ if the input is an FM signal given by
$$e_i(t) = A \cos (\omega_c t + \beta \sin \omega_m t)$$

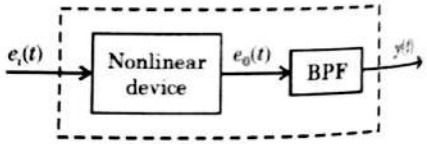


Fig. 4.39.

[Ans. $Y(t) = A \cos (2\omega_c t + 2\beta \sin \omega_m t)$ where $A' = \frac{1}{2} a A^2$. This result indicates that a square-law device can be used as a frequency doubler]

9. A given has maximum frequency deviation of 25 Hz for a modulating sinusoid of unit amplitude and a frequency of 100 kHz. Find the required value of frequency multiplication n to produce a maximum frequency deviation of 20 kHz when the modulating sinusoid has unit amplitude and a frequency of 200 Hz. [Ans. $n = 800$]

10. Assume that the 1.8 MHz signal in figure 4.40 is derived from the 200 kHz oscillator (multiplication by 4) and that the 200 kHz oscillator drift is 0.1 Hz.

(i) Find the drift in the 10.8 MHz signal.
(ii) Find the drift in the carrier of the resulting FM signal. [Ans. (a) ± 54 Hz, (b) 48 Hz]

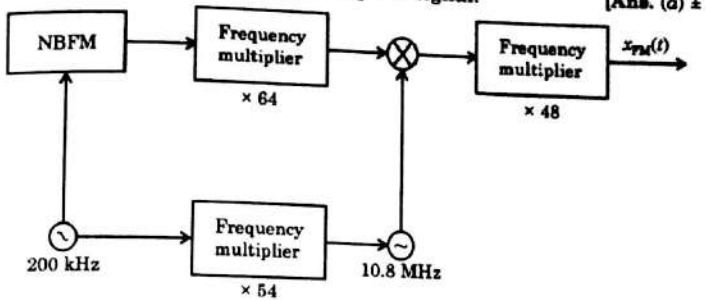


Fig. 4.40.

11. A block diagram of a typical FM receiver, covering the broadcast range of 88 to 108 MHz, is shown in figure 4.41. The amplifier frequency is 10.7 MHz. The limiter is used to remove the amplitude fluctuations caused by channel imperfection. The FM receiver is tuned to a carrier frequency of 100 MHz.

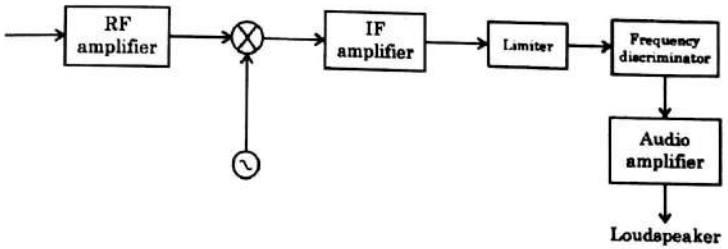


Fig. 4.41. *FM Receiver.*

(i) A 10 kHz audio signal frequency modulates a 100 MHz carrier, producing $\beta = 0.2$. Find the bandwidths required for the RF and IF amplifiers and for the audio amplifier.
(ii) Repeat (a) if $\beta = 5$.

[Ans. (i) RF and IF amplifiers: 24 kHz, audio amplifier: 10 kHz
(ii) RF and IF amplifier: 120 kHz, audio amplifier: 10 kHz]

□□□

Download from
www.dahiyastech.blogspot.com

CHAPTER 5

Probability, Random Signals & Random Process

Inside this Chapter

5.1. Introduction 300
5.2. Basic Definitions Related to Probability 301
5.3. Probability 302
5.4. Properties of Probability 302
5.5. Conditional Probability 303
5.6. Probability of Statistically Independent Events 303
5.7. Random Variables 307
5.8. Probability Function or Probability Distribution of a Discrete Random Variable 308
5.9. Cumulative Distribution Function (CDF) 308
5.10. Probability Density Function (PDF) 309
5.11. Joint Cumulative Distribution Function 317
5.12. The Joint Probability Density Function 318
5.13. Marginal Densities 321
5.14. Conditional Probability Density Function 322
5.15. Statistical Averages of Random Variables 323
5.16. Uniform Distribution 328
5.17. Gaussian or Normal Distribution 330
5.18. Rayleigh Distribution 332
5.19. Random Process 334
5.20. Sum of Random Processes 338
5.21. Correlation Function 339
5.22. Spectral Densities 341
5.23. Response of Linear Systems to Random Inputs 343

5.1. Introduction

In second chapter, we disscused how to apply Fourier transform to analyze some signals. All these signals were described by some fixed mathematical equations. Such type of signals are called **deterministic signals**. The behaviour of such type of signals as well as processing through linear time invariant (LTI) systems can be determined with the help of mathematical models. These

mathematical models represent the complete behaviour of the signal at every instant of time. Hence, for such deterministic signals, there is no uncertainty about the value at any instant of time. This means that these signals are predictable.

There is one other class of signals, the behaviour of which cannot be predicted. Such type of signals are called **random signals**. These signals are called random signals because the precise value of these signals cannot be predicted in advance before they actually occur. The examples of random signals are the noise interferences in communication systems. This means that the noise interference during transmission is totally unpredictable. In the same way, the noise generated by the receiver itself is random. Even some other signals which are not noise signals are also random signals. These signals cannot be modelled mathematically. Actually the electromagnetic interference is the major source of random noise.

In the receiver, the thermal noise is caused by the random motion of the electrons. Although the random signals are not predictable in advance precisely, they can be described in terms of its statistical properties. It is possible to analyze the random signals statistically with the help of probability theory.

In fact, the probability theory is very essential mathematical tool in the design of digital communication systems.

5.2. Basic Definitions Related to Probability

Before we discuss the concept of probability and its applications in the field of communication, we must discuss few basic terms which are related to probability theory.

5.2.1. Experiment

An experiment is defined as the process which is conducted to get some results. If the same experiment is performed repeatedly under the same conditions, similar results are expected. But there are few experiments which do not produce the same result as we have stated above. The theory of probability mainly deals with such type of experiments. In fact, the theory of probability is applied to such type of experiments to predict the possibility of a particular output. An experiment is sometimes called **trial**. As an example, throw of a coin is an experiment or trial. This trial results in two outcomes namely **Head** and **Tail**. Similarly, drawing a card from a well shuffled pack is the trial or experiment and it results in 52 possible outcomes (*i.e.* cards). Hence, each experiment or a trial has an outcome and the possibility of this outcome can be predicted with the help of probability theory.

The outcomes of an event are called **equally** likely, if any one of them cannot be expected to occur in preference to another. As an example, the tossing of a coin results in two outcomes, Head and Tail. Both have same possibility of 50%. Such type of outcomes are called **equally** likely outcomes. On the otherhand, if the box contains two black and four white balls, then getting a white or black ball is not equally likely. This is possible that we may get white balls sequentially.

5.2.2. Sample Space

A set of all possible outcomes of an experiment *or* trial is called the **sample space** of that experiment. It is generally denoted by 'S'. The total number of outcomes in a sample space is denoted by $n(s)$. As an example, the tossing of a coin has two outcomes. Hence, we may write its sample space as:

$$S = \{H, T\}$$

where, $H \rightarrow$ Head

and $T \rightarrow$ Tail

Now, we consider another example. If three coins are tossed simultaneously, then each experiment has two outcomes namely H and T. Hence, the maximum possible number of outcomes will be eight.

So that $$S = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$$

$$n(s) = 8$$

5.2.3. Event

The expected subset of the sample space *or* happening is called an event. As an example, let us consider an experiment of throwing a cubic die. In this case, the sample space S will be as

$$S = \{1, 2, 3, 4, 5, 6\}$$

Now, if we want the number '3' to be an outcome or an even number, *i.e.*, {2, 4, 6}, then this subset is called an **event**. This is denoted by letter 'E'. Hence event E is a subset of the sample space 'S'. If event E has only one outcome, then it is called an elementary event. On the other hand, if event E does not contain any out come, then it is called a null event. If $E = S$, then an event contains all the outcomes. Such as event is called a **certain event**. It always occurs, no matter what so ever is the outcome.

Two events A and B are called independent events if happening of event A has nothing to do with happening of B. As an example, let us consider an experiment of tossing a coin two times. In this case, occurrence of head in the first throw has nothing to do with the occurrence of head in the second throw. Hence, these two events are independent events since their outcomes are independent. On the other hand, if the outcome of one event is affected by other, then these events are called **dependent events**.

5.3. Probability

Probability may be defined as the study of random experiments. In any random experiment, there is always an uncertainty that a particular event will occur *or* not. As a measure of probability of occurrence of an event, a number between 0 to 1 is assigned. If it is sure that an event will occur, then we can say that its probability is 100% or 1. If it is not sure that an event will occur, then we can say that its probability is 0% or 0. If it is not sure whether the event will occur *or* not, then its probability is between 0 and 1.

Let us consider an example. The probability of occurrence of 28th February in a year is 1 since it is certain to occur every year. On the other hand, the probability of occurrence of 30th February in a year is 0 since it never comes. Again, the probability of occurrence of 29th February in a year is neither 0 not 1. It is always between 0 and 1. Actually, it is 1/4 since it occurs every leap year *i.e.*, once in four years.

Therefore, from above discussion, we can write a mathematical expression for probability as

Probability of an event A,

$$P(A) = \frac{\text{Number of possible favourable outcomes}}{\text{Total number of outcomes}}$$

For example, let us consider the probability of getting an even number in tossing of a die.

In tossing of a die, an even number can occur in 3 ways out of 6 equally likely ways.

Therefore, $$P(A) = \frac{3}{6} = \frac{1}{2} = 0.5 = 50\%$$

5.4. Properties of Probability

As discussed earlier, the sample space S contains all possible outcomes of an experiment. Also, an event is the subset of the sample space. Two events are said to be mutually exclusive if the occurrence of one of them precludes the occurrence of other. For example, in tossing of a coin, events Head and Tail are mutually exclusive. In throw of a die, the occurrence of number '4' will automatically exclude the occurrence of numbers 1, 2, 3 ,5 and 6. If an event contains all the outcomes then it is called certain event. Then, probability of this event is unity, *i.e.*,

$$P(A) = P(s) = 1$$

The properties of probability may be listed as under :

***Property* 1:** The probability of a certain event is unity *i.e.*,

$$P(A) = 1 \qquad ...(5.1)$$

***Property* 2:** The probability of any event is always less than *or* equal to 1 and non-negative. Mathematically,

$$0 \le P(A) \le 1 \qquad ...(5.2)$$

***Property* 3:** If A and B are two mutually exclusive events, then

$$P(A + B) = P(A) + P(B) \qquad ...(5.3)$$

***Property* 4:** If A is any event, then the probability of not happening of A is

$$P(\overline{A}) = 1 - P(A) \qquad ...(5.4)$$

where $\overline{A}$ represents the complement of event A.

***Property* 5:** If A and B are any two events (not mutually exclussive events), then

$$P(A + B) = P(A) + P(B) - P(AB) \qquad ...(5.5)$$

where $P(AB)$ is called the probability of events A and B both occurring simultaneously. Such an event is called joint event of A and B, and the probability $P(AB)$ is called the joint probability.

Now, if events A and B are mutually exclusive, then the joint probability, $P(AB) = 0$.

5.5. Conditional Probability

The concept of conditional probability is used in conditional occurrences of the events. Let us consider an experiment which involves two events A and B. Now the probability of event B, given that event A has occurred, is represented by $P(B/A)$. Similarly, $P(A/B)$ represents probability of event A given that event B has already occurred. Therefore, $P(B/A)$ and $P(A/B)$ are called **conditional probabilities**. These conditional probabilities may be defined in terms of their independent and joint probabilities as under:

The conditional probability of event B given that event A has already happened.

$$P(B/A) = \frac{P(AB)}{P(A)} \qquad ...(5.6)$$

where $P(AB)$ is the joint probability of A and B.

Similarly, the conditional probability of event A given that event B has already happened.

$$P(A/B) = \frac{P(AB)}{P(B)} \qquad ...(5.7)$$

At this point, it may be noted that the joint probability has commutative property which states that

$$P(AB) = P(BA) \qquad ...(5.8)$$

5.6. Probability of Statistically Independent Events

If A and B are two events in an experiment, and possibility of occurrence of event B does not depend upon occurrence of event A, then these two events A and B are known as statistically independent events. The probability of event B, given that event A has already happened is expressed as

$$P(B/A) = \frac{P(AB)}{P(A)} \qquad ...(5.9)$$

Again, since occurrence of event B does not depend on the occurrence of event A, then the probability of event B will be same as conditional probability $P(B/A)$. Mathematically,

$$P(B/A) = P(B) \qquad ...(5.10)$$

Now, putting the value of $P(B/A)$ from equation (5.10) in equation (5.9), we have

$$P(AB) = P(A)\,P(B) \qquad ...(5.11)$$

The probability of event A, given that event B has already happened, is expressed as:

$$P(A/B) = \frac{P(AB)}{P(B)} \qquad ...(5.12)$$

Again, since events A and B are statistically independent, then the probability of event A will be same as conditional probability $P(A/B)$. Mathematically,

$$P(A/B) = P(A) \quad \text{...(5.13)}$$

Using equations (5.12) and 5.13), we have

$$P(AB) = P(A)\,P(B) \quad \text{...(5.14)}$$

which is same as equation (5.11).

Example 5.1. A box contains 3 red, 4 white and 5 black balls. One ball is drawn at random. Find the probability that it is (a) red (b) not Black (c) black or white.

Solution: Let R, W and B be the events of drawing a red, a white and a black ball respectively.

(a)
$$P(R) = \frac{\text{Ways of choosing a red ball}}{\text{Total no. of ways of choosing a ball}}$$

$$P(R) = \frac{3}{3+4+5} = \frac{3}{12} = \frac{1}{4}$$

(b) Probability of not getting a black ball is $P(\overline{B})$.

$$P(\overline{B}) = 1 - P(\overline{B})$$

But
$$P(B) = \frac{5}{3+4+5} = \frac{5}{12}$$

Therefore,
$$P(\overline{B}) = 1 - \frac{5}{12} = \frac{7}{12}$$

(c) Probability of getting a black or white ball will be $P(B + W)$.

$$\therefore \quad P(B + W) = P(B) + P(W) \qquad [\because B \text{ and } W \text{ are two mutually exclusive events}]$$

$$P(B + W) = \frac{5}{12} + \frac{4}{12}$$

$$P(B + W) = \frac{9}{12} = \frac{3}{4}$$

Example 5.2. Two dies are thrown simultaneously. Find the probability of getting a 5.

Solution: We know that each die has six faces. Therefore, the total number of possible outcomes $6 \times 6 = 36$.

Each outcome is having an equal probability of occurrence. Therefore, probability of any outcome is 1/36.

We have 4 different combinations which give a 5. They are (4, 1), (3, 2), (2, 3), (1, 4).

Hence the probability of the overall event may be given as

$$P(A) = P(5) = \frac{1}{36} + \frac{1}{36} + \frac{1}{36} + \frac{1}{36} = \frac{4}{36} = \frac{1}{9} \quad \textbf{Ans.}$$

Example 5.3. Two cards are drawn at random from an ordinary deck of 52 playing cards. Obtain the probability that one card is a Diamond and the other card is a Heart.

Solution: Since two cards are drawn, there are $\frac{52 \times 51}{1 \times 2} = 1326$ ways of drawing 2 cards from a deck of 52 cards.

Again, because there are 13 diamonds and 13 Hearts, there are $13 \times 13 = 169$ ways of drawing a Diamond and a Heart. Therefore, the probability of

$$P(A) = \frac{\text{The number of ways of drawing a diamond and a heart}}{\text{The total number of ways of drawing two cards}}$$

$$P(A) = \frac{169}{1326} = \frac{13}{102} \quad \textbf{Ans.}$$

Example 5.4. A box contain 4 white and 3 black balls. Three balls are drawn from the box in succession. Find the probability that the first two balls are white and the third is black.

Solution: Let A_1 be the event that the first ball is a white ball.

$$P(A_1) = \frac{4}{7}$$

Since there are four white balls in the box out of total, 7 balls.

Let A_2 be the event that the second ball is also a white ball.

$$\therefore \quad P\left(\frac{A_2}{A_1}\right) = \frac{3}{6}$$

Since, after the first drawing, there remain 3 white balls out of a total of 6 balls.

Let A_3 be the event that the third ball is black.

$$\therefore \quad P\left(\frac{A_3}{A_1 A_2}\right) = \frac{3}{5}$$

Since after the second drawing, there are 3 black balls out of a total of 5 balls.

Therefore the required probability is

$$P = P(A_1 + A_2 A_3) = P(A_1)\,P(A_2/A_1)\,P(A_3/A_1A_2) = \frac{4}{7} \times \frac{3}{6} \times \frac{3}{5} = \frac{6}{35} \quad \textbf{Ans.}$$

Example 5.5. A box contains 5 white, 3 red and 2 black balls. Three balls are drawn in succession. Find the probability that the balls will be of different colours.

Solution: Let W, R, B be the events of drawing white, red and black balls, respectively. It may be observed that there are six possible ways of getting the balls of different colours as follow:

$$WRB,\ WBR,\ RWB,\ RBW,\ BWR,\ BRW$$

Now, the probability of first combination WRB will be

$$P(WRB) = \frac{5}{10}, \frac{3}{9}, \frac{2}{8}$$

Here, it may be noted that

$\frac{5}{10}$ is the probability of getting first ball white.

$\frac{3}{9}$ is the probability of getting second ball red.

$\frac{2}{8}$ is the probability of getting third ball black.

Similarly, probabilities for the other combinations may be calculated.

Since all the six combinations (events) are mutually exclusive, the required probability is

$$P = P(WRB + WBR + RWB + RBW + BWR + BRW)$$
$$= P(WRB) + P(WBR) + P(RWB) + P(RBW) + P(BWR) + P(BRW)$$
$$P = \left(\frac{5}{10} \times \frac{3}{9} \times \frac{2}{8}\right) + \left(\frac{5}{10} \times \frac{2}{9} \times \frac{3}{8}\right) + \left(\frac{3}{10} \times \frac{5}{9} \times \frac{2}{8}\right) + \left(\frac{3}{10} \times \frac{2}{9} \times \frac{5}{8}\right)$$
$$+ \left(\frac{2}{10} \times \frac{5}{9} \times \frac{3}{8}\right) + \left(\frac{2}{10} \times \frac{3}{9} \times \frac{5}{8}\right)$$
$$P = \frac{1}{4} \quad \textbf{Ans.}$$

Example 5.6. Three students A, B and C are given a problem in Maths. The probabilities of their solving the problem are $\frac{3}{4}$, $\frac{2}{3}$ and $\frac{1}{4}$ respectively. Determine the probability that the problem is solved if all of them try to solve the problem. *(PTU, 1999)*

Solution: The probabilities that A, B and C will not be able to solve the problem are $\left(1-\frac{3}{4}\right), \left(1-\frac{2}{3}\right)$ and $\left(1-\frac{1}{4}\right)$ or $\frac{1}{4}, \frac{1}{3}$ and $\frac{3}{4}$ respectively.

Therefore, the probability that no one will solve the problem

$$= \frac{1}{4} \times \frac{1}{3} \times \frac{3}{4} = \frac{3}{48}$$

Hence the probability that the problem will be solved

$$= 1 - \frac{3}{48} = \frac{45}{48} = \frac{15}{16} \qquad \textbf{Ans.}$$

Example 5.7. A box contains three white balls W_1, W_2, W_3 and two red balls R_1, R_2. Two balls are drawn in succession at random. Find the probability that the first drawn ball is white and the second ball is red.

Solution: Probability of event W_1 *i.e.* first ball is white will be

$$= \frac{\text{No. of white balls}}{\text{Total no. of balls}} = \frac{3}{5}$$

Now, if a white ball is drawn, there remain two white and two red balls.

Now, the conditional probability of the event of second ball red with first white

$$= P\left(\frac{R_2}{W_1}\right) = \frac{2}{4}$$

Hence, the probability of the event that the second ball is red with first white, will be

$$= P[W_1, R_2] = P\,(\text{first ball white, second red})$$

$$= P\left(\frac{R_2}{W_1}\right) \cdot P(W_1) = \frac{2}{4} \times \frac{3}{5} = \frac{6}{20} = \frac{3}{10} \qquad \textbf{Ans.}$$

Example 5.8. In a coin-tossing experiment, if the coin has head, one die is thrown and the result is recorded. But, if the coin has tail, two dice are thrown and their sum is recorded. Find the probability that the recorded number is 2.

Solution: We know that in a coin-tossing experiment, the probability of getting head $= \frac{1}{2}$.

Similarly, the probability of getting tail $= 1 - \frac{1}{2} = \frac{1}{2}$

Now, the probability of getting head with recorded number 2.

$$= \left(\frac{1}{2}\right)\left(\frac{1}{6}\right) = \frac{1}{12} \qquad \text{...(i)}$$

Similarly, the probability of getting tail with recorded number 2, will be

$$= \left(\frac{1}{2}\right)\left(\frac{1}{36}\right) = \frac{1}{72} \qquad \text{...(ii)}$$

Therefore, the required probability of getting recorded number 2 will be the sum of the equations (i) & (ii)

$$= \frac{1}{12} + \frac{1}{72} = \frac{6+1}{72} = \frac{7}{72} \quad \textbf{Ans.}$$

Example 5.9. Two weak students in programming write a program. Their chance of writing a program correctly are $\frac{1}{8}$ and $\frac{1}{12}$. Now, if the probability of making a common error is $\frac{1}{10001}$ and they get the same answer, then find the probability that their program is correct.

Solution: It is given that the probability of two students A and B writing the program correctly are $\frac{1}{8}$ and $\frac{1}{12}$ respectively *i.e.*, $P(A) = \frac{1}{8}$

and $P(B) = \frac{1}{12}$

Therefore, the probability of making common error $= \frac{1}{10001}$

Hence the required probability = Probability (Program is correct)

$$= \frac{\left(\frac{1}{8}\right)\left(\frac{1}{12}\right)}{\left(\frac{1}{8}\right)\left(\frac{1}{12}\right) + \left(1-\frac{1}{8}\right)\left(1-\frac{1}{12}\right)\left(\frac{1}{10001}\right)} = 0.9924 \qquad \textbf{Ans.}$$

5.7. Random Variables

As discussed earlier, an event is the possible outcome of an experiment. The range of all the possible outcomes of an experiment is called the **sample space** 'S'. When a trial *or* experiment is performed, any one sample point in the sample space is the outcome of the trial. This means that a sample point always lies in the sample space 'S'. On the other hand, an event may correspond to a single sample point *or* set of sample points.

As an example, in an experiment of tossing a die, the sample space contains six sample points. Every time, the trial is performed, the outcome is any one sample point (number 1 to 6) in the sample space. In every trial, the outcome (sample point) will occur randomly. This has no fixed output. Hence, the outcome of a trial *or* experiment is a variable which can take values over the set of sample points.

Therefore, from above discussion, we can define a random variable as under:

A function which can take on any value from the sample space and its range is some set of real numbers is called a **random variable** of the experiment. Random variables are denoted by upper case letters such as X, Y etc. and the values taken by them are denoted by lower case letters with subscripts such as x_1, x_2, y_1, y_2 etc.

Random variables may be classified as under:

1. Discrete random variables
2. Continuous random variables.

5.7.1. Discrete Random Variables

A discrete random variable may be defined as the random variable which can take on only finite number of values in a finite observation interval. This means that the discrete random variable has countable number of distinct values.

For example, let us consider an experiment of tossing three coins simultaneously. In this case, there are eight possible outcomes. This will constitute the sample space S. Let the number of heads be the random variable X. Sample space S and random variable X may be written as

$$S = \{HHH \quad HHT \quad HTH \quad THH \quad HTT \quad THT \quad TTH \quad TTT\}$$

$$X = \{x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad x_6 \quad x_7 \quad x_8\}$$

$$\{3 \quad 2 \quad 2 \quad 2 \quad 1 \quad 1 \quad 1 \quad 0\}$$

Hence, from above example, the concept of random variable is quite clear. From this example, it is also clear that random variable X takes on only a finite number of values *i.e.*, 8. Therefore, it is a discrete random variable.

5.7.2. Continuous Random Variables

A random variable that takes on an infinite number of values is called a continuous random variable. Actually, there are several physical system (experiments) that generate continuous outputs or outcomes. Such systems generate infinte number of outputs or outcomes within the finite period. Continuous random variables may be used to define the outputs of such systems.

As an example, the noise voltage generated by an electronic amplifier has a continuous amplitude. This means that sample space S of the noise voltage amplitude is continuous. Therefore, in this case, the random variable X has a continuous range of values.

5.8. Probability Function or Probability Distribution of a Discrete Random Variable

Let X be a discrete random variable and also let $x_1, x_2, x_3, \ldots$ be the values that X can take.

Then $$P(X = x_j) = f(x_j) \quad \text{...(5.15)}$$
where $j = 1, 2, 3, \ldots$
will be the probability of x_j.

This $f(x_j)$ or simply $f(x)$ is called the *probability function* or *probability distribution of the discrete random variable.*

Example 5.10. In an experiment, three coins are tossed simultaneously. If the number of heads is the random variable, find the probability function for this random variable.

Solution: In the three tosses of coins, there are eight possible outcomes. This is the sample space S. Let the random variable, number of heads, be X.

Since, in the given experiment, the random variable X, number of heads, takes only finite values, this is a discrete random variable.

Writing the sample space S and values of discrete random variable X as

$$S = \{HHH \quad HHT \quad HTH \quad THH \quad HTT \quad THT \quad TTH \quad TTT\}$$
$$X = \{x_1 \quad x_2 \quad x_3 \quad x_4 \quad x_5 \quad x_6 \quad x_7 \quad x_8\}$$
$$\{3 \quad 2 \quad 2 \quad 2 \quad 1 \quad 1 \quad 1 \quad 0\}$$

Since the coin is fair, the probability of each of 8 possible outcomes, will be 1/8.

Now, $$P(X = 0) = P(x_8) = \frac{1}{8}$$
$$P(X = 1) = P(x_5) + P(x_6) + P(x_7)$$
or $$P(X = 1) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8}$$
$$P(X = 2) = P(x_2) + P(x_3) + P(x_4)$$
or $$P(X = 1) = \frac{1}{8} + \frac{1}{8} + \frac{1}{8} = \frac{3}{8}$$
$$P(X = 3) = P(x_1) = \frac{1}{8}$$

Therefore, the probability function for discrete random variable X is as given below:

X	0	1	2	3
$f(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

5.9. Cumulative Distribution Function (CDF)

The Comulative Distribution Function (*CDF*) of a random variable 'X' may be defined as the probability that a random variable 'X' takes a value less than or equal to x. Here x is the dummy variable. In other words, the comulative distribution function (*CDF*) provides probabilistic description of a random variable.

Let us consider the probability of the event $X \le x$. The probability of this event may be denoted as $P(X \le x)$.

Now, according to the definition, the cumulative distribution function (*CDF*) may be written as

$$CDF: F_X(x) = P(X \le x) \quad \text{...(5.16)}$$

where $F_X(x)$ is called cumulative distribution function (*CDF*) of a random variable X.

From equation (5.16), it may be observed that $F_X(x)$ is a function of dummy variable 'x' *i.e.* the value taken by random variable X.

It may also be noted that for any value of 'x', the *CDF* represents a probability. The *CDF* may be defined for discrete as well as continuous random variables. Because cumulative distribution function (*CDF*) basically represents the probability of random variable 'X' for event $X \le x$, it is also called probability distribution function of the random variable *or* simply distribution function of the random variable. The *CDF* is sometimes also called cumulative probability distribution function.

5.9.1. Properties of Cumulative Distribution Function (CDF)

The properties of *CDF* may be listed as under:

***Property* 1:** Since cumulative distribution function (CDF) is the probability distribution function *i.e.* it is defined as the probability of event $(X \le x)$, its value is always between 0 and 1. This means that *CDF* is bounded between 0 and 1. Mathematically,

$$0 \le F_X(x) \le 1 \quad \text{...(5.17)}$$

***Property* 2:** $$F_X(-\infty) = 0 \quad \text{...(5.18)}$$
and $$F_X(\infty) = 1 \quad \text{...(5.19)}$$

For first case, $x = -\infty$ means no possible event. Due to this fact, $P(X \le -\infty)$ will always be zero. Therefore $F_X(-\infty) = 0$.

For second case, $x = \infty$ means $P(X \le \infty)$. Since $P(X \le \infty)$ includes probability of all possible events and the probability of a certain event is '1' therefore

$$F_X(\infty) = 1$$

***Property* 3:** $$F_X(x_1) \le F_X(x_2) \quad \text{if } x_1 \le x_2 \quad \text{...(5.20)}$$

The above property states that the *CDF*, $F_X(x)$ is a monotone non-decreasing function of x.

5.9.2. Cumulative Distribution Function (CDF) for Discrete Random Variables

If 'X' is a discrete random variable, then it takes on values at discrete points.

The Cumulative Distribution Function (CDF) is expressed as

$$F_X(x) = P(X \le x) \quad \text{...(5.21)}$$

Now, if X is a discrete random variable, then it can take on values as $x_1, x_2, x_3 \ldots, x_n$. This means that x may be expressed interms of $x_1, x_2, \ldots, x_n$.

Figure 5.1 shows how random variable can be assigned values.

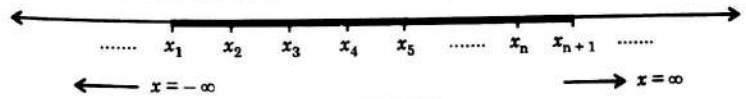


Fig. 5.1.

For a discrete random variable, the Cumulative Distribution Function (CDF) for a complete range of x may be expressed as

$$\boxed{\begin{aligned} F_X(x) &= 0 && \text{for } -\infty \le x < x_1 \\ &= \sum_{j=1}^{n} P(X = x_j) && \text{for } x_1 \le x \le x_n \\ &= 1 && \text{for } x_n < x < \infty \end{aligned}} \quad \text{...(5.22)}$$

From above equation, it is clear that *CDF* of a discrete variable at any certain event is equal to the summuation of the probabilities of random variable upto that event.

As x goes from $-\infty$ to ∞, $F_X(x)$ looks like a stair case with upward steps of height $P(X = x_j)$ at each $x = x_j$.

The *CDF* remains constant between the two events or steps.

5.10. Probability Density Function (PDF)

The derivative of cumulative distribution function (CDF) with respect to some dummy variable is known as Probability Density Function (PDF). Probability density function (PDF) is generally denoted by $f_X(x)$. Mathematically, PDF may be expressed as

$$PDF: f_X(x) = \frac{d}{dx} F_X(x) \qquad ...(5.23)$$

where x is a dummy variable.

Probability density function (*PDF*) is the more convenient representation for **continuous random variable.**

5.10.1. Properties of Probability Density Function (PDF)

In this article, we shall discuss a number of properties of PDF.

***Property* 1:** Probability Density Function (PDF) is always nonzero for all values of x.

Mathematically,

$$f_X(x) \geq 0 \qquad \text{for all values of } x \qquad ...(5.24)$$

Proof:

We know that cumulative distribution function (CDF) increases monotonically. Therefore, the derivative of cumulative distribution function (CDF) will always be positive. Since PDF is obtained by taking derivative of CDF, therefore PDF will always be positive.

***Property* 2:** The area under the PDF curve is always equal to unity. Mathematically,

$$\int_{-\infty}^{\infty} f_X(x)\, dx = 1 \qquad ...(5.25)$$

Proof: The expression for PDF is given as

$$f_X(x) = \frac{d}{dx} F_X(x)$$

Integrating both sides of above equation, we have

$$\int_{-\infty}^{\infty} f_X(x)\, dx = \int_{-\infty}^{\infty} \left[\frac{d}{dx} F_X(x)\right] dx$$

$$= [F_X(x)]_{-\infty}^{\infty}$$

$$= [F_X(\infty) - F_X(-\infty)] = 1 - 0 = 1$$

***Property* 3:** The cumulative distribution function (CDF) may be obtained by integrating probability density function (PDF). Mathematically,

$$F_X(x) = \int_{-\infty}^{x} f_X(x)\, dx \qquad ...(5.26)$$

Proof: The expression for Probability Density Function (PDF) is given as

$$f_X(x) = \frac{d}{dx} F_X(x)$$

Integrating both sides of the above equation, we have

$$\int_{-\infty}^{x} f_X(x)\, dx = \int_{-\infty}^{x} \left[\frac{d}{dx} F_X(x)\right] dx$$

or $$\int_{-\infty}^{x} f_X(x)\, dx = [F_X(x)]_{-\infty}^{x} = [F_X(x) - F_X(-\infty)]$$

or $$= F_X(x) - 0 = F_X(x)$$

Hence, we have $$F_X(x) = \int_{-\infty}^{x} f_X(x)\, dx \qquad \textbf{Hence Proved.}$$

***Property* 4:** Probability of the event $\{x_1 < X \leq x_2\}$ is simply given by the area under the Probability Density Function (PDF) curve in the range $x_1 < X \leq x_2$.

Mathematically,

$$P(x_1 < X \leq x_2) = \int_{x_1}^{x_2} f_X(x)\, dx \qquad ...(5.27)$$

Example 5.11. Probability density function (PDF) is given by the expression $f_X(x) = a\, e^{-b|x|}$. Here X is a random variable whose values lie in the range $x = -\infty$ to $x = +\infty$. Determine the following:

(*i*) The relationship between a and b.

(*ii*) Cumulative Distribution Function (CDF)

(*iii*) The probability that outcome lies between 1 and 2. *(A.M.I.E., Examination, 1997)*

Solution: (*i*) We know that the area under the *PDF* curve is always equal to 1.

$$\int_{-\infty}^{\infty} f_X(x)\, dx = 1,$$

Given that $$f_X(x) = a\, e^{-b|x|}$$

Therefore, $$\int_{-\infty}^{\infty} a\, e^{-b|x|}\, dx = 1$$

or $$\int_{-\infty}^{0} a\, e^{-b(-x)}\, dx + \int_{0}^{\infty} a\, e^{-bx}\, dx = 1$$

or $$\int_{-\infty}^{0} a\, e^{-b(-x)}\, dx + \int_{0}^{\infty} a\, e^{-bx}\, dx = 1$$

or $$\frac{a}{b}\left[e^{bx}\right]_{-\infty}^{0} - \frac{a}{b}\left[e^{-bx}\right]_{0}^{\infty} = 1$$

or $$\frac{a}{b}\,[e^0 - e^{-\infty} - e^{-\infty} + e^0] = 1$$

or $$\frac{a}{b}\,[1 - 0 - 0 + 1] = 1$$

or $$\frac{a}{b}\,[1 + 1] = 1$$

or $$\frac{2a}{b} = 1$$

or $$b = 2a$$

This is required relationship between a and b.

(*ii*) We know that cumulative distribution function (CDF) may be obtained by integrating PDF.

CDF: $$F_X(x) = \int_{-\infty}^{x} f_X(x)\, dx = \int_{-\infty}^{x} a\, e^{-b|x|}\, dx$$

Now, we shall break the above expression for $x < 0$, $x = 0$, $x > 0$.

$$\text{Therefore, } F_X(x) = \underbrace{\int_{-\infty}^{x} a e^{-b(-x)}\, dx}_{\text{This factor for } x<0} + \underbrace{\int_{x=0} a e^{-bx} dx + \int_{0^+}^{x} a e^{-bx}\, dx}_{\text{These two factors for } x \geq 0}$$

Hence, for $x < 0$, we have

$$\int_{-\infty}^{x} a e^{bx}\, dx = \frac{a}{b}\left[e^{bx}\right]_{-\infty}^{x} = \frac{a}{b}\left[e^{bx} - e^{-\infty}\right] = \frac{a}{b} e^{bx} = \frac{1}{2} e^{bx}$$

Similary, for $x \geq 0$, we have

$$\int_{x=0} a e^{b}\, dx + \int_{0^+}^{x} a e^{-bx}\, dx$$

$$= \frac{a}{b}[e^0] - \frac{a}{b}\left[e^{-bx}\right]_{0^+}^{x}$$

$$= \frac{a}{b} \times 1 - \frac{a}{b}\left[e^{-bx} - e^{0^+}\right]$$

$$= \frac{a}{b} - \frac{a}{b}\left[e^{-bx} - 1\right] = \frac{a}{b}\left[1 - e^{-bx} + 1\right] \qquad \left[\because e^{0^+} \equiv 1\right]$$

$$= \frac{a}{b}\left[2 - e^{-bx}\right] = \frac{a}{2a}\left[2 - e^{-bx}\right] \qquad (\because b = 2a)$$

$$= \frac{1}{2}(2 - e^{-bx}) = 1 - \frac{1}{2} e^{-bx}$$

Therefore, the expression for *CDF* will be given by

$$F_X(x) = \begin{cases} \frac{1}{2} e^{bx} & \text{for } x < 0 \\ 1 - \frac{1}{2} e^{-bx} & \text{for } x \geq 0 \end{cases}$$

Plot of cumulative distribution function (CDF) with respect to x:

We know from previous discussion that for a CDF $F_X(x)$

$F_X(x)$ at $x = -\infty = 0$

and $F_X(x)$ at $x = \infty = 1$

Also, $F_X(x)$ at $x = 0 = 1 - \frac{1}{2} e^{-bx}$

$$= 1 - \frac{1}{2} e^{-b.0} = 1 - \frac{1}{2} e^{0} = \frac{1}{2}$$

Hence, from the above information we can have an approximate plot of $F_X(x)$ as:

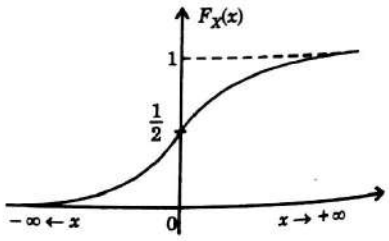


Fig. 5.2. *Plot of CDF*

(*iii*) We know that the probability that outcome lies between x_1 and x_2, is given by the expression

$$P(x_1 < X \leq x_2) = \int_{x_1}^{x_2} f_X(x)\, dx$$

Given that $x_1 = 1$ and $x_2 = 2$

Therefore, $$P(1 < X \leq 2) = \int_{1}^{2} a e^{-b|x|}\, dx$$

For the interval 1 to 2, $e^{-b|x|} = e^{-bx}$

Hence, $$P(1 < X \leq 2) = \int_{1}^{2} a e^{-bx}\, dx = -\frac{a}{b}\left[e^{-bx}\right]_1^2 = -\frac{a}{b}\left[e^{-2b} - e^{-b}\right]$$

or $$P(1 < X \leq 2) = \frac{a}{b}\left[e^{-b} - e^{-2b}\right]$$

But since $2a = b$, So $\frac{a}{b} = \frac{1}{2}$

$\therefore$ $$P(1 < X \leq 2) = \frac{1}{2}\left[e^{-b} - e^{-2b}\right] \qquad \textbf{Ans.}$$

Example 5.12. The CDF for a certain random variable is given as

$$F_X(x) = \begin{cases} 0 & -\infty < x \leq 0 \\ kx^2 & 0 < x \leq 10 \\ 100\,\text{k} & 10 < x < \infty \end{cases}$$

(*i*) Find the value of k;
(*ii*) Find the value of $P(X \leq 5)$;
(*iii*) Find the value of $P(5 < X \leq 7)$;
(*iv*) Find the expression for PDF.

Solution: (*i*) The given *CDF* is

$$F_X(x) = \begin{cases} 0 & -\infty < x \leq 0 \\ kx^2 & 0 < x \leq 10 \\ 100\text{k} & 10 < x < \infty \end{cases} \qquad \text{...(i)}$$

We know that the expression for CDF is given as

$$F_X(x) = \begin{cases} 0 & -\infty < x \leq x_1 \\ \sum_{j=1}^{n} P(X = x_j) & x_1 < x \leq x_n \\ 1 & x_n < x < \infty \end{cases} \qquad \text{...(ii)}$$

Comparing the two equations (*i*) & (*ii*) we have

$$F_X(x) = 100\,\text{k} = 1 \qquad \text{for } 10 < x < \infty$$

or $$k = \frac{1}{100} \qquad \textbf{Ans.}$$

(*ii*) Since CDF is expressed as

$$F_X(x) = P(X \leq x)$$

or $$P(X \leq x) = F_X(x)$$

Therefore, for $x = 5$, we have

$$P(X \leq 5) = F_X(5)$$

But $$F_X(x) = \frac{1}{100} \times x^2$$

Therefore, $$P(X \leq 5) = \frac{1}{100}(5)^2$$

$$P(X \leq 5) = \frac{1}{100} \times 25 = \frac{1}{4} \qquad \textbf{Ans.}$$

(*iii*) The PDF is expressed as

$$f_X(x) = \frac{d}{dx} F_X(x)$$

Since
$$F_X(x) = \begin{cases} 0 & -\infty < x \le 0 \\ \dfrac{x^2}{100} & \text{for } 0 < x \le 10 \\ 1 & \text{for } 10 < x < \infty \end{cases}$$

Now on differentiating $F_X(x)$ with respect to x, we have

PDF, $f_X(x) = 0$ for $-\infty < x \le 0,\ 10 < x < \infty$

and
$$f_X(x) = \begin{cases} \dfrac{d}{dx}\left[\dfrac{x^2}{100}\right] & \text{for } 0 < x \le 10 \\ \dfrac{x}{50} & \text{for } 0 < x \le 10 \end{cases}$$

(*iv*) We know that $P(x_1 < X \le x_2) = \int_{x_1}^{x_2} f_X(x)\, dx$

Here, $x_1 = 5,\ x_2 = 7$

$$P(5 < X \le 7) = \int_5^7 \frac{x}{50}\, dx \qquad [\because f_X(x) = \frac{x}{50}, \text{ for } 0 < x \le 10]$$

or
$$P(5 < X \le 7) = \frac{1}{50}\left[\frac{x^2}{2}\right]_5^7 = \frac{1}{100}\,[49 - 25]$$

$$P(5 < X \le 7) = \frac{1}{100} \times 24 = \frac{24}{100} \qquad \textbf{Ans.}$$

Example 5.13. A continuous random variable has a probability density function (PDF) expressed as

$$f_X(x) = 2\,e^{-2x} \qquad \text{for } x \ge 0$$

Determine the probability that it will take a value between 1 and 3.

Solution: We know that the relation between probability and PDF is expressed as

$$P(x_1 < X \le x_2) = \int_{x_1}^{x_2} f_X(x)\, dx$$

Now the probability that a random variable will be between 1 and 3 is expressed as

$$P(1 < X \le 3) = \int_1^3 f_X(x)\, dx = \int_1^3 2\,e^{-2x}\, dx$$

or
$$P(1 < X \le 3) = 2\left[\frac{e^{-2x}}{-2}\right]_1^3 = 2\left[-\frac{1}{2}(e^{-6} - e^{-2})\right] = -[-0.1328] = 0.1328 \qquad \textbf{Ans.}$$

Example 5.14. A random process provides measurements x between the values 0 and 1 with a probability density function (PDF) given as

$$f_X(x) = 12x^3 - 21x^2 + 10x \quad \text{for } 0 \le x \le 1 = 0, \text{ otherwise}$$

Determine the followings:

(*i*) $P\left[X \le \frac{1}{2}\right]$ and $P\left[X > \frac{1}{2}\right]$

(*ii*) **Obtain a number k such that $P[X \le k] = \frac{1}{2}$.** *(Karnataka University, 1999)*

Solution: (*i*) We know that

$$P(x_1 < X \le x_2) = \int_{x_1}^{x_2} f_X(x)\, dx$$

Therefore,

$$P(X \le \tfrac{1}{2}) = \int_0^{1/2} f_X(x)\, dx = \int_0^{1/2} (12x^3 - 21x^2 + 10x)\, dx$$

or
$$P(X \le \tfrac{1}{2}) = \left[\frac{12x^4}{4} - \frac{21x^3}{3} + \frac{10x^2}{2}\right]_0^{1/2} = \left[3x^4 - 7x^3 + 5x^2\right]_0^{1/2}$$

or
$$P(X \le \tfrac{1}{2}) = \left[3\left(\frac{1}{2}\right)^4 - 7\cdot\left(\frac{1}{2}\right)^3 + 5\left(\frac{1}{2}\right)^2 - 0\right] = \frac{9}{16}$$

Also,

$$P(X > \tfrac{1}{2}) = 1 - P(X \le \tfrac{1}{2}) = 1 - \frac{9}{16} = \frac{7}{16}$$

(*ii*) Since

$$P(x_1 < X \le x_2) = \int_{x_1}^{x_2} f_X(x)\, dx$$

Therefore,
$$P(X \le k) = \int_0^k f_X(x)\, dx$$

But given that $P(X \le k) = \frac{1}{2}$

Hence,
$$\frac{1}{2} = \int_0^k (12x^3 - 21x^2 + 10x)\, dx = \left[\frac{12x^4}{4} - \frac{21x^3}{3} + \frac{10x^2}{2}\right]_0^k$$

or
$$\frac{1}{2} = \left[3x^4 - 7x^3 + 5x^2\right]_0^k = 3k^4 - 7k^3 + 5k^2$$

or $6k^4 - 14k^3 + 10k^2 - 1 = 0$ ...(*i*)

The required answer will be the root of equation (*i*) which lies between 0 and 1.

Therefore, solving equation (*i*) for a root between 0 and 1, we get

$$k = 0.453 \qquad \textbf{Ans.}$$

Example 5.15. Determine whether the function given by expression

$$f_X(x) = \begin{cases} 0 & \text{for } x < 2 \\ \frac{1}{18}(3+2x) & \text{for } 2 \le x \le 4 \\ 0 & \text{for } x > 4 \end{cases}$$

is a density function? *(RGP University, MP-1999)*

Solution: If any given function is a probability density function (PDF), then it must satisfy following two basic conditions:

(i) $f_X(x) \ge 0$ for every x

(ii) $\int_{-\infty}^{\infty} f_X(x)\,dx = 1$

The given function is $= 0$ for $x < 2$

$$f_X(x) = \frac{1}{18}(3+2x) \text{ for } 2 \le x \le 4 = 0 \quad \text{for } x > 4$$

From this expression, it may be observed that first condition is satisfied.

For second condition, we have

$$\int_{-\infty}^{\infty} f_X(x)\,dx = \int_{-\infty}^{2} 0\,dx + \int_{2}^{4} \frac{1}{18}(3+2x)\,dx + \int_{4}^{\infty} 0\,dx$$

or
$$\int_{-\infty}^{\infty} f_X(x)\,dx = 0 + \frac{1}{18}\left[3x + x^2\right]_2^4 + 0 = \frac{1}{18}\left[\{3(4)+16\} - \{3(2)+4\}\right]$$

$$= \frac{1}{18}[12+16-6-4] = \frac{1}{18}[28-10] = \frac{1}{18} \times 18 = 1$$

Therefore, second condition is also satisfied.

This means that the given function is a probability function (PDF).

Example 5.16. Determine the constant k such that the function $f_X(x)$ given by the expression

$$f_X(x) = \begin{cases} \frac{1}{k} & \text{for } a \le x < b \\ 0 & \text{elsewhere} \end{cases}$$

is a probability density function (PDF). Also, find the commulative distribution function (CDF) of the random variable X if k satisfies the conditions for $f_X(x)$ to be a probability density function (PDF).

Solution: Since it is given that $f_X(x)$ is a density function, therefore, it will satisfy the following requirement:

$$\int_{-\infty}^{\infty} f_X(x)\,dx = 1$$

Putting the value of $I(x)$ in the given limit, we get

$$\int_a^b \frac{1}{k}\,dx = 1 \quad \text{or} \quad \frac{1}{k}\int_a^b dx = 1$$

or
$$\frac{1}{k}[x]_a^b = 1 \quad \text{or} \quad k = [b-a] \qquad \textbf{Ans.}$$

We know that CDF is given as

$$F_X(x) = \int_{-\infty}^{x} f_X(x)\,dx = \int_a^x f_X(x)\,dx = \int_a^x \frac{1}{k}\,dx = \int_a^x \frac{1}{b-a}\,dx$$

or
$$F_X(x) = \frac{1}{b-a}\int_a^x dx = \frac{1}{b-a}[x]_a^x = \frac{1}{b-a}\cdot x - a = \frac{x-a}{b-a} \qquad \textbf{Ans.}$$

Example 5.17. Determine whether the following function is commulative distribution function (CDF):

$$F_X(x) = \begin{cases} 0 & \text{for } x < a \\ \frac{1}{2}\left(\frac{x}{a}+1\right) & \text{for } -a \le x \le a \\ 1 & \text{for } x > a \end{cases}$$

(WBTU, Kolkata-2003)

Solution: We know that the probability density function (PDF) is given as

$$f_X(x) = \frac{d}{dx}F_X(x)$$

$$f_X(x) = \begin{cases} \frac{1}{2a} & \text{for } -a \le x \le a \\ 0 & \text{elsewhere} \end{cases}$$

The required condition for a function $f_X(x)$ to be PDF is

$$\int_{-\infty}^{\infty} f_X(x)\,dx = 1 \qquad \text{...(i)}$$

Therefore,

$$\int_{-\infty}^{\infty} f_X(x)\,dx = \int_{-\infty}^{-a} f_X(x)\,dx + \int_{-a}^{a} f_X(x)\,dx + \int_{a}^{\infty} f_X(x)\,dx$$

$$\int_{-\infty}^{\infty} f_X(x)\,dx = \int_{-\infty}^{-a} 0\cdot dx + \int_{-a}^{a} \frac{1}{2a}\,dx + \int_{a}^{\infty} 0\cdot dx = 0 + \frac{1}{2a}[x]_{-a}^{a} + 0 = \frac{1}{2a}[a-(-a)] = 1$$

Hence, the function $f_X(x)$ satisfies the condition of a PDF. This means that $F_X(x)$ is a commulative distribution function (CDF). **Ans.**

5.11. Joint Cumulative Distribution Function

We may also define the outcome of an experiment by two random variables. In earlier article, we discussed the, cumulative distribution function (CDF) for a single random variable. In this article, we shall discuss the CDF for two random variables X and Y which is known as combined CDF or simply joint Distribution Function.

The joint Distribution Function or joint CDF $F_{XY}(x, y)$ of two random variables X and Y is defined as the probability that the random variable X is less than *or* equal to a specified value x and that the random variable y is less than *or* equal to a specified value y.

The joint Cumulative Distribution Function may be defined systematically as:

The joint Cumulative Distribution Function $F_{XY}(x, y)$ may be defined as the probability that the outcome of an experiment will result in a sample point lying inside the range $(-\infty < X \le x, -\infty < Y \le y)$ of the joint sample space. The joint sample space is the combined sample space of X and Y. Mathematically,

$$F_{XY}(x, y) = P(X \le x, Y \le y) \quad ...(5.28)$$

5.11.1. Properties of Joint Cumulative Distribution Function

The Joint Cumulative Distribution Function has the following properties:

***Property* 1:** The Joint Cumulative Distribution Function is a non-negative function.

Mathematically,

$$F_{XY}(x, y) \ge 0 \quad ...(5.29)$$

The Joint Cumulative Distribution Function is basically defined as the probability in the Joint sample space of random variables. We know that the probability always lies between 0 and 1. Therefore, the joint Cumulative Distribution Function also lies between 0 and 1 and hence non-negative.

***Property* 2:** The Joint Cumulative Distribution Function is a monotone non-decreasing function of both x and y.

***Property* 3:** The Joint Cumulative Distribution Function is always continuous everywhere in the xy-plane.

5.12. The Joint Probability Density Function

The Joint Probability Density Function *or* simply Joint PDF is the PDF of two or more random variables.

The joint PDF of any two random variables X and Y may be defined as the partial derivative of the joint cumulative distribution function $F_{XY}(x, y)$ with respect to the dummy variables x and y. Mathematically,

$$f_{XY}(x, y) = \frac{\partial^2 F_{XY}(x, y)}{\partial x \, \partial y} \quad ...(5.30)$$

Here, we take partial derivative since the differentiation is with respect to two variables x and y.

5.12.1. Properties of Joint PDF

The Joint PDF has the following properties:

***Property* 1:** The Joint PDF is non-negative. Mathematically,

$$f_{XY}(x, y) \ge 0 \quad ...(5.31)$$

Because the joint PDF is a derivative of a non-negative function (Joint CDF), therefore its value is always positive.

***Property* 2:** The total volume under the surface of joint PDF is equal to unity.

Mathematically,

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x, y)\, dx\, dy = 1 \quad ...(5.32)$$

Because $f_{XY}(x, y)$ is a two-dimensional function, it represents a surface in the x, y plane. Hence, its integration is called as volume under the surface.

***Property* 3:** The Joint PDF is continuous everywhere because joint CDF is continuous.

5.12.2. The Relationship between Joint *PDF* and Probability

The relationship between probability and PDF over a certain interval is expressed as

$$P(x_1 < X \le x_2) = \int_{x_1}^{x_2} f_X(x)\, dx \quad ...(5.33)$$

The above relationship may be extended for two random variables X and Y. Mathematically,

$$P(x_1 < X \le x_2, y_1 < Y \le y_2) = \int_{y_1}^{y_2}\int_{x_1}^{x_2} f_{XY}(x, y)\, dx\, dy \quad ...(5.34)$$

In above expression, the double integral represents the volume between x-y palne and the surface $f_{XY}(x, y)$.

Now, if the two random variables X and Y are statistically independent, then joint PDF of these two random variables becomes a product of two separate PDFs. Mathematically,

$$f_{XY}(x, y) = f_X(x)\, f_Y(y) \quad ...(5.35)$$

Putting this value of $f_{XY}(x, y)$ from equation in equation, we have

$$P(x_1 < X \le x_2, y_1 < Y \le y_2) = \int_{y_1}^{y_2}\int_{x_1}^{x_2} f_X(x)\, f_Y(y)\, dx\, dy \quad ...(5.36)$$

$$= \int_{y_1}^{y_2} f_Y(y)\, dy + \int_{x_1}^{x_2} f_X(x)\, dx \quad ...(5.37)$$

The above equation provides the relationship between probability PDF of statistically independent random variables X and Y.

Example 5.18. The joint PDF of two random variables is expressed as $F_{XY}(x, y)$. If the random variables X and Y are statistically indendent, then show that

$$f_{XY}(x, y) = F_X(x)\, F_Y(y)$$

Here $f_X(x)$ is the PDF of random variable X and $f_Y(y)$ is the PDF of random variable Y.

Solution: We know that the joint *PDF* may be defined in terms of joint CDF as

$$f_{XY}(x, y) = \frac{\partial^2 F_{XY}(x, y)}{\partial x \, \partial y}$$

Now, if random variables X and Y are statistically independent, then

$$f_{XY}(x, y) = \frac{d}{dx} F_{XY}(x, y) \frac{d}{dy} F_{XY}(x, y) = f_X(x)\, f_Y(y)$$

Example 5.19. A joint probability density function of two random variables X and Y is given as

$$f_{XY}(x, y) = \begin{cases} e^{-(x+y)} & \text{for } x \ge 0, y \ge 0 \\ 0 & \text{otherwise} \end{cases}$$

Determine the followings:

(i) $P(X < 1)$

(ii) $P(X > Y)$

(iii) $P(X + Y < 1)$

Solution: Given that joint probability density function is given as

$$f_{XY}(x, y) = e^{-(x+y)} = e^{-x} \cdot e^{-y} = f_X(x) \cdot f_Y(y)$$

Hence, X and Y are independent random variables.

Therefore, $f_X(x) = e^{-x}$ for $x \geq 0$

and $f_Y(y) = e^{-y}$ for $y \geq 0$

Now,

(i) $$P(X < 1) = \int_0^1 e^{-x}\,dx = \left[\frac{e^{-x}}{-1}\right]_0^1 = -\left[e^{-x}\right]_0^1$$

or $$P(X < 1) = -(e^{-1} - e^{-0}) = -\left(\frac{1}{e} - 1\right) = 1 - \frac{1}{e} \qquad \textbf{Ans.}$$

(ii) $$P(X > Y) = 1 - P(X < Y)$$

or $$P(X < Y) = \int_0^\infty \left[\int_0^y f_{XY}(x, y)\,dx\right] dy = \int_0^\infty e^{-y}\left[\frac{e^{-x}}{-1}\right]_0^y dy$$

or $$P(X < Y) = -\int_0^\infty e^{-y}\cdot(e^{-y} - 1)\,dy = -\left(\frac{e^{-2y}}{-2} + e^{-y}\right)_0^\infty = \left(\frac{1}{2}e^{-2y} - e^{-y}\right)_0^\infty$$

$$= \left[\left(\frac{1}{2}e^{-\infty} - e^{-\infty}\right) - \left(\frac{1}{2}e^{-0} - e^{-0}\right)\right] = \frac{1}{2}\cdot 0 - 0 - \frac{1}{2}\cdot 1 + 1$$

or $$P(X < Y) = 0 - 0 - \frac{1}{2} + 1 = \frac{1}{2}$$

Hence, $$P(X > Y) = 1 - P(X < Y) = 1 - \frac{1}{2} = \frac{1}{2}$$

(iii) $$P(X + Y < 1) = \int_0^1 \left[\int_0^{1-x} e^{-(x+y)}\,dy\right] dx$$

or $$P(X + Y < 1) = \int_0^1 e^{-x}\left[\frac{e^{-y}}{-1}\right]_0^{1-x} dx = -\int_0^1 e^{-x}\cdot[e^{-(1-x)} - 1)]\,dx$$

or $$P(X + Y < 1) = -\int_0^1 [e^{-1} - e^{-x}]\,dx = 1 - \frac{2}{e} \qquad \textbf{Ans.}$$

Example 5.20. The joint probability density function of two random variables X and Y is given as

$$f_{XY}(x, y) = \begin{cases} C(2x + y) & \text{for } 0 \leq x \leq 2,\ 0 \leq y \leq 3 \\ 0 & \text{elsewhere} \end{cases}$$

Determine the value of constant C.

Solution: For a joint *PDF*, we know that

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x, y)\,dxdy = 1$$

Putting the value of $f_{XY}(x, y)$, we get

$$\int_0^3\int_0^2 C(2x + y)\,dxdy = 1$$

or $$C\int_0^3\left[\frac{2x^2}{2} + yx\right]_0^2 dy = 1 \quad C\int_0^3 (2y + 4)\,dy = 1$$

or $$C\left(\frac{2y^2}{2} + 4y\right)_0^3 = 1 \quad C(y^2 + 4y)_0^3 = 1$$

or $$C(9 + 12 - 0 - 0) = 1 \quad 21\cdot C = 1 = \frac{1}{21} \qquad \textbf{Ans.}$$

5.13. Marginal Densities

distribution function

In case, when the probability density functions $f_X(x)$ and $f_Y(y)$ for any single random variable are obtained from joint PDF, then $f_X(x)$ and $f_Y(y)$ are called marginal PDF *or* simply marginal densities.

We can obtain the marginal *PDFs* as follows:

Since, the CDF for a single random variable may be obtained from its PDF as

$$F_X(x) = \int_{-\infty}^{x} f_X(x)\,dx \qquad \text{...(5.38)}$$

The above expression may be extended for any two random variable, X and Y as

$$F_{XY}(x, y) = P(X \leq x,\ Y \leq y) \qquad \text{...(5.39)}$$

$$F_{XY}(x, y) = \int_{-\infty}^{y}\int_{-\infty}^{x} f_{XY}(x, y)\,dx\,dy \qquad \text{...(5.40)}$$

Now, if we have to find *CDF* of random variable X, then the value of Y does not matter. Therefore,

$$F_X(x) = P(X \leq x) = P(X \leq x, -\infty < y \leq \infty) \qquad \text{...(5.41)}$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{x} f_{XY}(x, y)\,dx\,dy \qquad \text{...(5.42)}$$

For a single random variable, the PDF may be obtained by differentiating CDF.

$$f_X(x) = \frac{d}{dx}F_X(x) \qquad \text{...(5.43)}$$

In this equation differentiation is with respect to x because y varies from $-\infty$ to $+\infty$.

or $$f_X(x) = \frac{d}{dx}\left[\int_{-\infty}^{\infty}\int_{-\infty}^{x} f_{XY}(x, y)\,dx\,dy\right] \qquad \text{...(5.44)}$$

or $$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)\,dx\,dy \qquad \text{...(5.45)}$$

Here, dx is cancelled due to differentiation. Therefore, the above equation shows that the PDF $f_X(x)$ may be obtained by integrating joint PDF $f_{XY}(x, y)$ with respect to undersired random variable y from $-\infty$ to ∞.

This $f_X(x)$ is then called marginal PDF. In the same way, the marginal PDF for random variable y may be obtained. It is expressed as

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y)\, dx \qquad ...(5.46)$$

Therefore the marginal densities are:

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)\, dy \qquad ...(5.47)$$

and

$$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y)\, dx \qquad ...(5.48)$$

5.14. Conditional Probability Density Function

Out of the two random variables, one variable may take a fixed value. In this case, the PDF is called *conditional*. As an example, out of the two continuous random variables X and Y, let $X = x$.

Then we may find the conditional, *PDF* of Y given that $X = x$ as,

$$f_Y(y/x) = \frac{f_{XY}(x, y)}{f_X(x)} \qquad ...(5.49)$$

where $f_X(x)$ is the marginal density of random variable X.

Similarly, we may find the conditional *PDF* of X given that $Y = y$. Then

$$f_X(x/y) = \frac{f_{XY}(x, y)}{f_Y(y)} \qquad ...(5.50)$$

5.14. 1. Properties of Conditional PDF

The conditional PDF has following properties:

***Property* 1:** A conditional *PDF* is always non-negative function. Mathematically,

$$f_X(x/y) \geq 0 \qquad ...(5.51)$$

and

$$f_Y(y/x) \geq 0 \qquad ...(5.52)$$

The conditional PDF is the ratio of two PDFs which are non-negative. Therefore, conditional PDF is also non-negative.

***Property* 2:** The are under the conditional PDF is equal to unity. Mathematically,

$$\int_{-\infty}^{\infty} f_X(x/y)\, dx = 1 \qquad ...(5.53)$$

and

$$\int_{-\infty}^{\infty} f_Y(y/x)\, dy = 1 \qquad ...(5.54)$$

***Property* 3:** If two random variables X and Y are statistically independent, then the conditional PDF reduces to marginal density.

Mathematically,

$$f_Y(y/x) = f_Y(y) \qquad ...(5.55)$$

and

$$f_X(x/y) = f_X(x) \qquad ...(5.56)$$

5.15. Statistical Averages of Random Variables

We know that the Probability Density Function (PDF) gives some kind of information about the random variable. However, the interpretation of this information is quite complex. There are some other measures or numbers which give more useful and quick information about the random variable. Collectively these characteristic numbers or measures are known as **statistical averages**. Some useful and convenient informations like mean or average, moments, standard deviation, variance etc. may be obtained for random variables. These factors are special characteristics the Probability Density Function (PDF) and actually they depend on the particular type of PDF.

Now we shall discuss these statistical averages.

5.15.1. Mean or Average

The mean or average of any random variable is expressed by the summation of the values of random variables X weighted by their probabilities. Mean value of a random variable is denoted by m_x.

Mean value is also known as expected value of random variable X.

$$m_x = E(X) \qquad ...(5.57]$$

where $E[\ ]$ represents expection operator.

Therefore both m_x and $E(X)$ have the same meaning.

Also, in general term, mean or average is given by

$$\text{Mean or average of } X = \frac{\text{Arithmatic sum of all values of } X}{\text{Total number of values of } X}$$

5.15.2. Mean Value of Discrete Random Variable

Let the discrete random variable X be take the following values

$$X = \{x_1, x_2, x_3 \ldots x_n\} \qquad ...(5.58)$$

Then the mean or average value m_x is expressed as

$$m_x = x_1P(x_1) + x_2P(x_2) + x_3P(x_3) + \ldots + x_nP(x_n) \qquad ...(5.59)$$

Therefore, the mean value of a discrete random variable X is equal to the summation of values of X weighted by probability of that particular value.

Hence, mean value of discrete random variable

$$m_x = E(X) = \overline{X} = \sum_{i=1}^{n} x_i . P(x_i) \qquad ...(5.60)$$

Here $\overline{X}$ is also the notation for mean values.

5.15.3. Mean Values of Continuous Random Variables

If random variable X becomes continuous, the sample points $x_1, x_2, x_3 \ldots, x_n$ becomes quite close to each other such that $(x_2 - x_1 \cong 0)$. Therefore, summation in equation 5.60 for discrete random variable converts to an integration over the complete range of $x(-\infty < x < \infty)$.

For continuous random variable X, the mean or average value is expressed as

$$m_x = \int_{-\infty}^{\infty} x f_X(x)\, dx \qquad ...(5.61)$$

where $f_X(x)$ = Probability density function (PDF)

Also, if a function $g(x)$ transforms random variable X into another random variable, then mean or average of $g(x)$ may be expressed as

$$m = \overline{g(x)} = E[g(x)] \quad ...(5.62)$$

$$m = \int_{-\infty}^{\infty} g(x)\, f_X(x)\, dx \quad ...(5.63)$$

Example 5.21. The random variable Z is the function of another random variable X in such a way that $Z = \cos(X)$ and X is uniformly distributed in the interval $(-\pi, \pi)$ *i.e.*,

$$f_X(x) = \begin{cases} \dfrac{1}{2\pi} & \text{for } -\pi < x < \pi \\ 0 & \text{otherwise} \end{cases}$$

Determine the mean value of Z.

Solution: The given random variable is

$$Z = \cos(X)$$

Now, assuming $g(x) = \cos(X)$. The above equation may be written as

$$Z = g(x)$$

We know that the mean value of random variable $Z = g(x)$ is expressed as

$$m = \int_{-\infty}^{\infty} g(x)\, f_X(x)\, dx$$

Putting the values of $g(x)$ and $f_X(x)$, we have

$$m = \int_{-\pi}^{\pi} \cos x \left(\frac{1}{2\pi}\right) dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} \cos x\, dx$$

or $$m = \frac{1}{2\pi}[-\sin x]_{-\pi}^{\pi} = \frac{1}{2\pi}[0] = 0 \qquad \textbf{Ans.}$$

Thus it may be observed from above that since the given function is $Z = \cos(X)$, it has zero mean value over the period of one cycle.

5.15.4. Moments and Variance

The nth moment of any random variable X may be defined as the mean value of X^n.

i.e.

$$g(x) = X^n \quad ...(5.64)$$

then the mean is written as

$$\text{mean} = \overline{g(x)} \quad ...(5.65)$$

$$\overline{X^n} = \int_{-\infty}^{\infty} x^n\, f_X(x)\, dx = E(X^n) \quad ...(5.66)$$

$$\overline{X^n} = \int_{-\infty}^{\infty} x^n\, f_X(x)\, dx$$

Putting $n = 1$, the above equation becomes

$$\overline{X} = E(X) = \int_{-\infty}^{\infty} x\, f_X(x)\, dx \quad ...(5.67)$$

or $$\overline{X} = m_x \quad ...(5.68)$$

Hence, the first moment of random variable X will be same as its mean value.

If $n = 2$, then equation (5.66) becomes

$$\overline{X^2} = E(X^2) = \int_{-\infty}^{\infty} x^2\, f_X(x)\, dx \quad ...(5.69)$$

where $\overline{X^2}$ is known as mean square value of random variable X.

Similarly, the central moments are the moments of the difference between random variable X and its mean m_x. Therefore, the nth central moment may be given as

$$E[(X - m_x)^n] = \int_{-\infty}^{\infty} (x - m_x)^n\, f_X(x)\, dx \quad ...(5.70)$$

The second central moment for $n = 2$, is known as **variance** of random variable X *i.e.*,

$$\text{Variance } [X] = E[(X - m_x)^2] \quad ...(5.71)$$

$$= \int_{-\infty}^{\infty} (x - m_x)^2\, f_X(x)\, dx \quad ...(5.72)$$

Hence, it is clear that the variance provides an indication about randomness of the random variable.

Variance is generally represented by σ_x^2 *i.e.*,

$$\sigma_x^2 = \text{Variance } (X) \quad ...(5.73)$$

$$\sigma_x^2 = \text{Var } (X)$$

or $$\sigma_x^2 = E[(X - m_x)^2] \quad ...(5.74)$$

or $$\sigma_x^2 = \int_{-\infty}^{\infty} (x - m_x)^2\, f_X(x)\, dx \quad ...(5.75)$$

Expanding equation (5.74), we get

$$\sigma_x^2 = E[X^2 - 2\, m_x X + m_x^2] \quad ...(5.76)$$

$$\because (X - m_x)^2 = X^2 + m_x^2 - 2\, m_x X$$

The $E[\,]$ operator represents mean or average value, therefore it is linear. Thus,

$$\sigma_x^2 = E(X^2) - 2\, m_x E(X) + m_x^2 \quad ...(5.77)$$

or $$\sigma_x^2 = E(X^2) - 2\, m_x\, m_x + m_x^2$$

or $$\sigma_x^2 = E(X^2) - m_x^2 \quad ...(5.78)$$

or $$\sigma_x^2 = \text{mean square value} - \text{square of mean value}$$

The above equation may also be written as

$$\text{Variance} = \sigma_x^2 \quad ...(5.79)$$

$$= \overline{X^2} - m_x^2 \quad ...(5.80)$$

The square root of variance is known as **standard deviation** of a random variable X. Standard deviation gives the measure of spread observed over the values of X related to mean value. Therefore,

$$\text{Standard deviation} = \sqrt{\text{Variance}} \quad ...(5.81)$$

$$= \sqrt{\sigma_x^2} = \sigma_x \quad ...(5.82)$$

Thus, standard deviation $= \sigma_x$

$$= \sqrt{E(X^2) - m_x^2} \qquad \text{...(5.83)}$$

Example 5.22. The probability density function (PDF) of a continuous random variable is of the form

$$f_X(x) = \frac{1}{2} e^{-|x|} \qquad \text{for } -\infty < x < \infty$$

Determine the mean of the random variables.

Solution: We know that for a continuous random variable mean is expressed as

$$m_x = \int_{-\infty}^{\infty} x f_X(x)\, dx = \int_{-\infty}^{\infty} x \cdot \frac{1}{2} e^{-|x|}\, dx = \int_{-\infty}^{\infty} \frac{x}{2} e^{-|x|}\, dx$$

Let $\quad \frac{x}{2} e^{-|x|} = f(x)$

Then $\quad m_x = \int_{-\infty}^{\infty} f(x)\, dx$

But $f(x)$ is an odd function since

$$f(-x) = -\frac{x}{2} e^{-|x|} = -f(x)$$

Therefore $\int_{-\infty}^{\infty} f(x)\, dx = 0$

Thus $\quad m_x = 0 \qquad$ **Ans.**

Example 5.23. The probability density function (PDF) of a continuous random variable X in the range (–3, 3) is defined as follow:

$$f_X(x) = \begin{cases} = \frac{1}{16}(3+x)^2 & \text{for } -3 \le x \le -1 \\ = \frac{1}{16}(2-6x)^2 & \text{for } -1 \le x \le 1 \\ = \frac{1}{16}(3-x)^2 & \text{for } 1 \le x \le 3 \end{cases}$$

Verify that the area under the curve is unity. Also prove that the mean is zero.

(A.M.I.E., Examination, 1998)

Solution: We know that the area under the PDF curve is expressed as

$$\text{area} = \int_{-\infty}^{\infty} f_X(x)\, dx = \int_{-3}^{3} f_X(x)\, dx$$

$$\text{area} = \int_{-3}^{-1} f_X(x)\, dx \int_{-1}^{1} f_X(x)\, dx \int_{1}^{2} f_X(x)\, dx$$

or

$$\text{area} = \int_{-3}^{1} \frac{1}{16}(3+x)^2\, dx + \int_{-1}^{1} \frac{1}{16}(2-6x^2)\, dx + \int_{1}^{3} \frac{1}{16}(3-x)^2\, dx$$

$$\text{area} = \frac{1}{48}[8 + 32 + 8] = \frac{1}{48} \times 48 = 1$$

Therefore, the area under the PDF curve is unity.

We know that for a continuous random variable, the mean is expressed as

$$m_x = \int_{-\infty}^{\infty} x f_X(x)\, dx$$

where $f_X(x)$ = probability density function (PDF)

Now

$$m_x = \int_{-3}^{3} x \cdot f_X(x)\, dx = \int_{-3}^{1} \frac{1}{16} \cdot x(3+x)^2\, dx + \int_{-1}^{1} \frac{1}{16} \cdot x(6-2x^2)\, dx + \int_{1}^{3} \frac{1}{16} \cdot x(3-x)^2\, dx$$

$$m_x = \frac{1}{16}\int_{-3}^{-1} (9x + 6x^2 + x^3)\, dx + \frac{1}{16}\int_{-1}^{1} (6x - 2x^3)\, dx + \frac{1}{16}\int_{1}^{3} (9x - 6x^2 + x^3)\, dx = 0$$

Hence, mean value is zero. **Ans.**

Example 5.24. A random variable X has the uniform distribution given by

$$f_X(x) = \begin{cases} \frac{1}{2\pi} & \text{for } 0 \le x \le 2\pi \\ = 0 & \text{otherwise} \end{cases}$$

Determine m_x, $\overline{X^2}$ and σ_x. *(BPTU, Orissa-2004)*

Solution: We know that the mean value of a continuous random variable is expressed as

$$m_x = \int_{-\infty}^{\infty} x f_X(x)\, dx = \int_{0}^{2\pi} x \cdot \frac{1}{2\pi}\, dx$$

$$m_x = \frac{1}{2\pi}\int_{0}^{2\pi} x\, dx = \frac{1}{2\pi}\left[\frac{x^2}{2}\right]_0^{2\pi} = \frac{1}{4\pi}\left[x^2\right]_0^{2\pi} = \frac{1}{4\pi}[4\pi^2 - 0] = \pi \qquad \textbf{Ans.}$$

We know that the mean square value of a continuous variable is given as

$$\overline{X^2} = E(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x)\, dx = \int_{0}^{2\pi} x^2 \frac{1}{2\pi}\, dx = \frac{1}{2\pi}\int_{0}^{2\pi} x^2\, dx$$

or

$$\overline{X^2} = \frac{1}{2\pi}\left[\frac{x^3}{3}\right]_0^{2\pi} = \frac{1}{6\pi}\left[x^3\right]_0^{2\pi} = \frac{1}{6\pi}[8\pi^3 - 0] = \frac{1}{6\pi} \cdot 8\pi^3 = \frac{4}{3}\pi^2 \qquad \textbf{Ans.}$$

Variance of a continuous random variable is expressed as

$$\sigma_x^2 = E[X^2] - m_x^2 = \frac{4}{3}\pi^2 - \pi^2 = \frac{1}{3}\pi^2$$

Also the standard deviation σ_x is given as the square root of variance *i.e.*,

$$\sigma_x = \sqrt{\sigma_x^2} = \sqrt{\frac{1}{3}\pi^2} = \frac{\pi}{\sqrt{3}} \qquad \textbf{Ans.}$$

Example 5.25. The probability density function (PDF) of a random variable X is given as $f_X(x)$. A random variable Y is defined as $Y = aX + b$ where $a < 0$. Obtain PDF of Y in terms of the PDF of X.

Solution: We know that the Probability Density Function (PDF) may be expressed as

$$P(x < X \le x + dx) = f_X(x)\, dx$$

Similarly for random variable Y

$$P(y < Y \le y + dy) = f_Y(y)\, dy$$

Because the probabilities of random variables X and Y are same, we may equate the above two equations. Thus we get

$$f_Y(y)\, dy = f_X(x)\, dx = f_X(x) \cdot \frac{dx}{dy} \qquad \text{...(i)}$$

Also given that $\quad y = ax + b$

thus $\quad \dfrac{dy}{dx} = a = \dfrac{1}{a} \quad$ and $\quad x = \dfrac{y-b}{a}$

Putting all these values in equation (*i*), we get

$$f_Y(y) = \frac{1}{a} f_X\left(\frac{y-b}{a}\right)$$

This is the required expression of PDF of Y in terms of PDF of X.

5.16. Uniform Distribution

The Probability Density Function (PDF) of an uniform distribution is expressed as

$$\textit{Uniform PDF: } f_X(x) = 0 \qquad \text{for } x < m - \frac{A}{2} \;\&\; x > m + \frac{A}{2}$$

$$= \frac{1}{A} \qquad \text{for } \left(m - \frac{A}{2}\right) \le x \le \left(m + \frac{A}{2}\right) \qquad \text{...(5.84)}$$

Figure 5.3 shows the PDF of an uniformly distributed random variable.

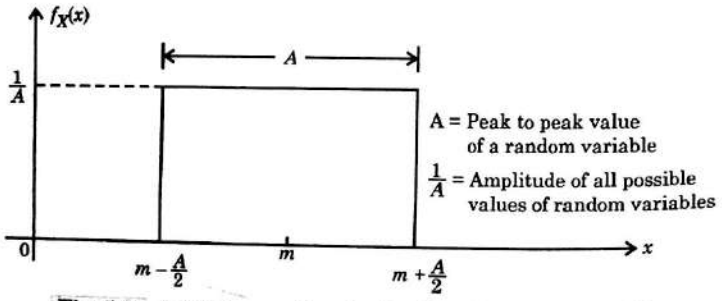


Fig. 5.3. *PDF of an uniformly distributed random variable*

In above figure, the peak to peak value is A and amplitude is uniform *i.e.* A. From figure 5.3, it is also clear that the random variable and its PDF are continuous. Therefore uniform distribution is utilised for continuous random variable.

Also, the value of PDF, $f_X(x)$ is same for all possible values of a random variable. Thus this distribution is known as **Uniform Distribution.**

Example 5.26. Prove that the mean and variance of a random variable X having an uniform distribution in the inerval $[a, b]$ are

$$m_x = \frac{a+b}{2} \text{ and } \sigma_x^2 = \frac{(a-b)^2}{12}$$

Solution: An uniform distribution in the interval $[a, b]$ is sketched in the figure (5.4).

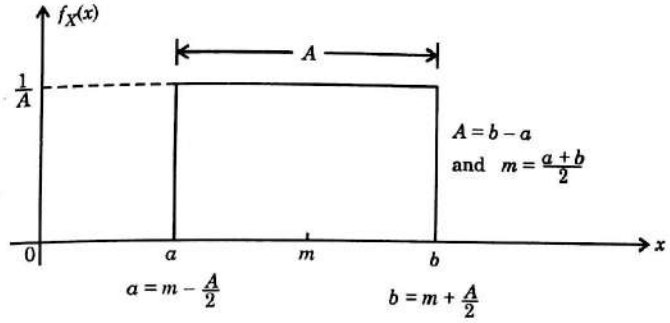


Fig. 5.4. *Uniform distribution having interval [a, b].*

We know that the mean value of a continuous random variable is expressed as

$$m_x = \int_{-\infty}^{\infty} x\, f_X(x)\, dx = \int_a^b x \cdot \frac{1}{A}\, dx$$

From figure 5.4 it may be observed that the value of A will be the difference between upper and lower limits of interval *i.e.*,

$$A = b - a$$

With this value of A, the mean value will become

$$m_x = \int_a^b x \cdot \frac{1}{b-a}\, dx = \frac{1}{b-a}\int_a^b x\, dx = \frac{1}{2(b-a)}\left[\frac{x^2}{2}\right]_a^b = \frac{1}{2(b-a)}\left[x^2\right]_a^b$$

or $$m_x = \frac{1}{2(b-a)}[b^2 - a^2] = \frac{1}{2(b-a)} \cdot (b-a)(b+a) = \frac{b+a}{2} = \frac{a+b}{2} \qquad \text{...(i)}$$

Also, the variance is expressed as

$$\sigma_x^2 = \int_{-\infty}^{\infty} (x - m_x)^2 f_X(x)\, dx = \int_a^b (x - m_x)^2 \cdot \frac{1}{b-a}\, dx \qquad \left[\because f_X(x) = \frac{1}{b-a}\right]$$

Now, let $x - m_x = y$, then $dx = dy$ and limits will become

if $\quad x = a$, then $y = a - m_x$

and if $\quad x = b$, then $y = b - m_x$

So that, $$\sigma_x^2 = \int_{a-m_x}^{b-m_x} y^2 \cdot \frac{1}{b-a}\, dy = \frac{1}{b-a}\left[\frac{y^3}{3}\right]_{a-m_x}^{b-m_x}$$

or $$\sigma_x^2 = \frac{1}{3(b-a)}[(b-m_x)^3 - (a-m_x)^3]$$

Putting the value of $m_x = \frac{a+b}{2}$ in above equation, we get

$$\sigma_x^2 = \frac{1}{3(b-a)}\left[\left(b - \frac{a+b}{2}\right)^3 - \left(a - \frac{a+b}{2}\right)^3\right]$$

or
$$\sigma_x^2 = \frac{1}{3(b-a)}\left[\left(\frac{b-a}{2}\right)^3 - \left(\frac{a-b}{2}\right)^3\right]$$

writing $(b-a) = -(a-b)$, we have

$$\sigma_x 2 = -\frac{1}{3(a-b)}\left[\left(-\frac{a-b}{2}\right)^3 - \left(\frac{a-b}{2}\right)^3\right]$$

$$\sigma_x^2 = -\frac{1}{3(a-b)} \cdot -\frac{(a-b)^3}{4} = \frac{(a-b)^2}{12} \quad ...(ii)$$

Because $(a-b)^2 = (b-a)^2$, the above equation may be written as

$$\sigma_x^2 = \frac{(a-b)^2}{12}$$

But $A = b - a$,

Therefore, $\sigma_x^2 = \frac{A^2}{12}$ ...(iii)

From figure 5.4, we have

$$a = m - \frac{A}{2}$$

and
$$b = m + \frac{A}{2}$$

thus
$$a + b = m - \frac{A}{2} + m + \frac{A}{2} = 2m$$

or
$$\frac{a+b}{2} = m$$

Putting this value in equation (*i*), we get
Mean value of uniform distribution

$$m_x = \frac{a+b}{2} = m$$

Using equations (*ii*) and (*iii*), we get
Variance of Uniform Distribution

$$\sigma_x^2 = \frac{(a-b)^2}{12} = \frac{(b-a)^2}{12} = \frac{A^2}{12} \quad \textbf{Ans.}$$

5.17. Gaussian or Normal Distribution

Gaussian Distribution is also known as Normal Distribution. It is defined for a continuous random variable. The PDF for a Gaussian random variable is expressed as

Gaussian PDF

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-m)^2/2\sigma^2} \quad ...(5.85)$$

In this equation

m = mean value of the random variable

σ^2 = variance of the random variable

Figure 5.5 shows the plot of Gaussian PDF $f_X(x)$ with respect to x.

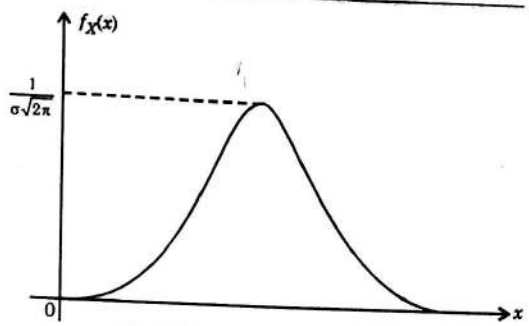


Fig. 5.5. *Plot of Gaussian PDF.*

5.17.1. Properties of Gaussian PDF

The several properties of Gaussian PDF are as under:

Property 1: The peak value occurs at $x = m$ *i.e.* mean value.
Mathematically,

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} \text{ at } x = m \text{ (mean value)} \quad ...(5.86)$$

This value may be obtained by putting $x = m$ in equation (5.85).

Property 2: The plot of Gaussian *PDF* exhibit even symmetry around mean value.
Mathematically,

$$f_X(m - \sigma) = f_X(m + \sigma) \quad ...(5.87)$$

Property 3: The area under the PDF curve is 1/2 for all values of x below mean value and 1/2 for all values of x above mean value. Mathematically,

$$P(X \le m) = P(X > m) = \frac{1}{2} \quad ...(5.88)$$

Property 4: As $\sigma \to 0$, the Gaussian function approaches to δ (impulse) function located at $x = m$. This is because the area under the PDF curve is always one (unity). Also the area of impulse function is also unity.

Importance

The Gaussian distribution is used for continuous random variables. We know that the random motion of the thermaly agitated electrons produces thermal noise. This thermal noise has Gaussian distribution. The random errors in the experimental measurements create the measured values to have Gaussian distribution about the true value. The Gaussian distribution is very important in the analysis of communication and statistical systems.

Example 5.27. Find out the Cumulative Distribution Function (CDF) of the Gaussian Random variable.

Solution: We know that the Cumulative Distribution Function (CDF) may be obtained by integrating *PDF* as

CDF:

$$F_X(x) = \int_{-\infty}^{x} f_X(x)\, dx \quad ...(i)$$

But Gaussian PDF is expressed as

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-m)^2/2\sigma^2}$$

Putting this value of $f_X(x)$ in equation (*i*)

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-m)^2/2\sigma^2}\, dx \qquad ...(ii)$$

Let $\frac{m-x}{\sigma\sqrt{2}} = z$ in the above equation

then $-\frac{dx}{\sigma\sqrt{2}} = dz$

or $dx = -\sigma\sqrt{2}\, dz$

The limits will be

when $x \to -\infty$; then $z \to +\infty$

when $x \to x$; then $z \to \frac{m-x}{\sigma\sqrt{2}}$

Putting all these values in equation (*ii*)

$$F_X(x) = \int_{\infty}^{\frac{m-x}{\sigma\sqrt{2}}} \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-z^2}\left(-\sigma\sqrt{2}\, dz\right) = -\frac{1}{\sqrt{\pi}} \int_{\infty}^{\frac{m-x}{\sigma\sqrt{2}}} e^{-z^2}\, dz$$

Interchanging the limits, we get

$$= \frac{1}{\sqrt{\pi}} \int_{\frac{m-x}{\sigma\sqrt{2}}}^{\infty} e^{-z^2}\, dz$$

Rearranging, we get

$$F_X(x) = \frac{1}{2} \cdot \frac{2}{\sqrt{\pi}} \int_{\frac{m-x}{\sigma\sqrt{2}}}^{\infty} e^{-z^2}\, dz \qquad ...(iii)$$

We know that the error function is given by

$$\text{erfc}\,(u) = \frac{2}{\sqrt{\pi}} \int_{u}^{\infty} e^{-z^2}\, dz \qquad ...(iv)$$

Here $u = \frac{m-x}{\sigma\sqrt{2}}$

Using equations (*iii*) and (*iv*), we get

Gaussian *CDF*: $F_X(x) = \frac{1}{2}\,\text{erfc}\left(\frac{m-x}{\sigma\sqrt{2}}\right)$

5.18. Rayleigh Distribution

Rayleigh Distribution is used for continuous random variables. It is produced from two Gaussian random variables. Let X and Y be independent Gaussian random variables having mean value zero and variance σ^2. Mathematically,

$$m_x = m_y = 0 \qquad ...(5.89)$$

and

$$\sigma_x^2 = \sigma^2 \qquad ...(5.90)$$

Figure 5.6 shows the rectangular to polar conversion.

From figure 5.6, it may be observed that independent Gaussian random variables X and Y are related to Rayleigh Distribution random variables R and ϕ such that

$$R = \sqrt{X^2 + Y^2} \qquad ...(5.91)$$

and

$$\phi = \tan^{-1}\left(\frac{Y}{X}\right) \qquad ...(5.92)$$

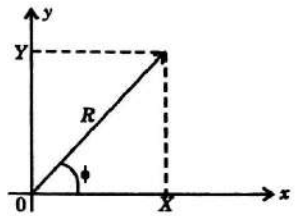


Fig. 5.6. *Rectangular to polar conversion*

From these equations, it is evident that R will always be positive because the variables X and Y are squared.

Also, the range of ϕ will be form 0 to 2π because it repeats after 2π radians. This means that ϕ completes one circle in figure 5.6 when it goes from 0 to 2π. After transformation of Gaussian random variables X and Y into R, it is found that R is having Rayleigh Probability Density Function (PDF) which is expressed as

$$f_R(r) = \frac{r}{\sigma^2} e^{-r^2/2\sigma^2} \qquad ...(5.93)$$

Figure 5.7 shows the Rayleigh PDF defined by above equation.

Therefore, from *PDF* curve, it is clear that

$$f_R(r) = 0 \qquad \text{for } r < 0 \qquad ...(5.94)$$

$$f_R(r) = \frac{r}{\sigma^2} e^{-r^2/2\sigma^2} \qquad \text{for } r \geq 0 \qquad ...(5.95)$$

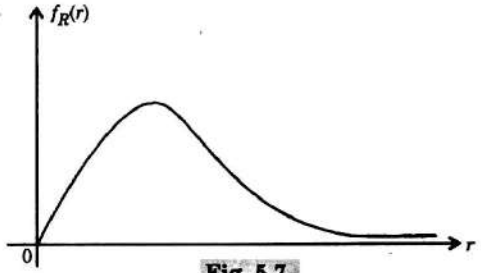


Fig. 5.7.

This means that R has always positive value and never goes for negative value.

Importance

Rayleigh Distribution is always used for modelling of statistics of signals transmitted through radio channels such as cellular radio.

Example 5.28. If Probability distribution function (PDF) of Rayleigh distribution is given by the expression

$$f_R(r) = \frac{r}{\sigma^2} e^{-r^2/2\sigma^2}$$

then above that the cumulative distribution function (CDF) of Rayleigh Distribution will be

$$F_R(r) = P(R \leq r) = \left(1 - e^{-r^2/2\sigma^2}\right) u(r)$$

Solution: We know that *CDF* may be obtained from PDF as

$$F_X(x) = \int_{-\infty}^{x} f_X(x)\, dx$$

where

$$F_R(r) = \int_{-\infty}^{x} f_R(r)\, dr$$

Now since r never becomes negative, therefore integration limits will be from 0 to r.

Thus,

$$F_R(r) = \int_{0}^{r} \frac{r}{\sigma^2} e^{-r^2/2\sigma^2}\, dr \qquad ...(i)$$

Putting

$$\frac{r^2}{2\sigma^2} = t$$

So that $\frac{2r\,dr}{2\sigma^2} = dt$ or $\frac{r\,dr}{\sigma^2} = dt$

when $r = 0$ then $t = 0$

and when $r = r$ then $t = \frac{r^2}{2\sigma^2}$

Putting all these values in equation (*i*), we get

$$F_R(r) = \int_0^{r^2/2\sigma^2} e^{-t}\,dt = \left[-e^{-t}\right]_0^{r^2/2\sigma^2}$$

$$= \left[-e^{-r^2/2\sigma^2} + e^0\right] = \left[-e^{-r^2/2\sigma^2} + 1\right]$$

or $F_R(r) = 1 - e^{-r^2/2\sigma^2}$

Since we know that $r \geq 0$ always, therefore the above equation may be multiplied by unit step function $u(r)$.

Hence Rayleight CDF:

$$F_R(r) = \left(1 - e^{-r^2/2\sigma^2}\right) u(r)$$

5.19. Random Process

Let there be a random experiment E having outcome λ from the sample sapce S. This means that $\lambda \in S$. Thus every-time an experiment is conducted, the outcome λ will be one of the sample point in sample space. If this outcome λ is associated with time, then a function of λ and time t is formed i.e., $X(\lambda, t)$. Then the function $X(\lambda, t)$ is known as **random process**. Hence, when any random experiment E is given a time dimension, then each outcome appears at some certain time and the random experiment will be converted to Random Process. A random process is the function of two variables λ and t.

Here λ belongs to sample sapce S and $-\infty < t < \infty$. At some specific time say $t = t_0$, $X(\lambda, t_0)$ is a random variable whose value depends upon λ. Also, for a specific outcome λ_i, there is a single time function $X(\lambda_i, t_0)$. This function is known as sample function whereas the collection of all the sample functions is known an **ensemble**.

We can represent random process as $X(t)$ *i.e.* here λ is dropped.

Therefore, a random process $X(t)$ may represent several functions of t and λ, a single time function, a random variable or a single number.

In practical life, there are various physical phenomenon which may be represented with the help of random process.

As an example, let us consider a set of voltage waveforms generated by thermal electron motion in a large number of identical resistors.

We can represent the voltage produced at each instant by the resistor as v. Then the random process $X(v, t)$ denotes the output voltages from the set of resistors.

Figure 5.8 shows the waveforms produced by these resistors randomly and different parameters of random process.

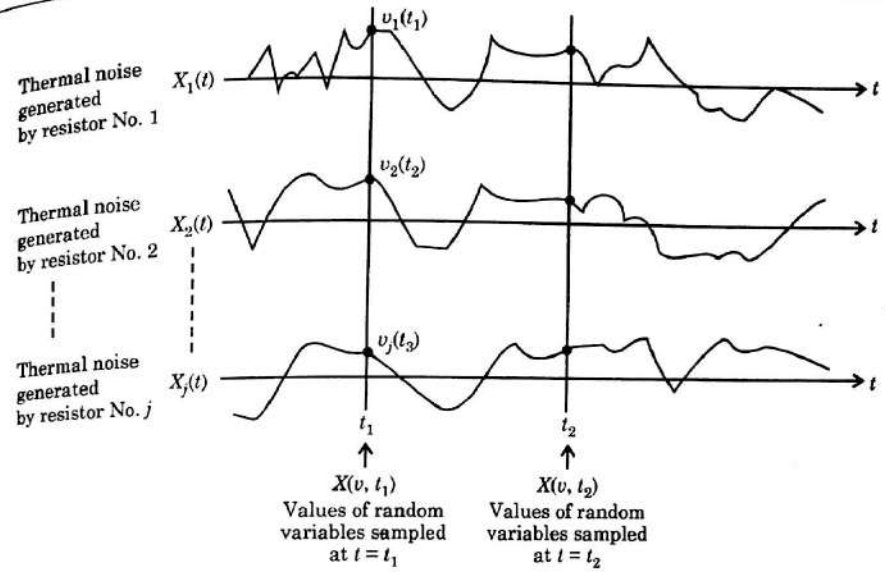


Fig. 5.8.

5.19.1. Ensemble and Time Averages

We know that the random variables may be specified with the help of different probability distributions. Random proceesses cannot be completely specified with the help of PDFs. Generally ensemble and time statistics are used to specify the random processes. Most generally, statistical averages such as mean $m_x(t)$ and autocorrelation function $R_\tau(t_1, t_2)$ are used to describe random process.

5.19.2. Ensemble Averages

In case of ensemble averages, the average is taken over the ensemble of waveforms, keeping the time constant or fixed.

As an example, if we desire ensemble mean value in random process of figure 5.8, then we will have to take mean of $v_1(t_1)$, $v_2(t_1)$, $v_3(t_1)$ ..., $v_i(t_1)$.

This ensemble mean is taken at time $t = t_1$.

In the same way, we can take ensemble mean at time $t = t_2$ also. The ensemble mean at $t = t_1$ and $t = t_2$ need not be same. Similarly, ensemble mean at other values of t may also be evaluated. Thus ensemble mean becomes function of time t. Therefore it is represented as $m_x(t)$.

We know that the mean value for a continuous random variable is expressed as

$$m_x = \int_{-\infty}^{\infty} x f_X(x)\,dx \qquad ...(5.96)$$

Using this expression, the ensemble mean may be defined as

$$m_x(t) = E\,[X(v, t] \qquad ...(5.97)$$

Hence, Ensemble mean value

$$m_x(t) = \int_{-\infty}^{\infty} x f_X(x, t)\,dx \qquad ...(5.98)$$

In this equation, the time t is treated as constant and the integration is done with respect to x. The next ensemble average is the autocorrelation function which may be expressed as

$$R_\tau(t_1, t_2) = E[X(v, t_1) \cdot X(v, t_2)]$$
$$= E[X(t_1) \cdot X(t_2)] \quad ...(5.99)$$
$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2 f_{x_1} f_{x_2}(x_1, x_2)\, dx_1\, dx_2 \quad ...(5.100)$$

In this equation, the value of $R_\tau(t_1, t_2)$ represents the similarity of amplitudes at time t_1 and t_2 and is obtained by multiplying the amplitudes of sample functions at t_1 and t_2 and taking mean of this product over an ensemble.

Thus the autocorrelation function gives the information about frequency content of the process.

5.19.3. Time Averages

When the statistical averages are taken along the time; they are known as time averages. As an example, we may define time mean value of a sample function $x(t)$ as

Time mean value: $$<m_x> = \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} X(t)\, dt \quad ...(5.101)$$

The equation (5.101) is a standard equation to find average or mean value of a function.

Let us consider the random process shown in figure (5.8). We can obtain time mean of waveform $X_1(t)$ by integrating it over a long period. The time mean of $X_1(t)$ and $X_2(t)$ need not be same. Thus the collection of time means becomes a function of the sample function $X_i(t)$.

The autocorrelation function may be expressed using time averaging as

$$<R_\tau(\varepsilon)> = \lim_{T\to\infty} \frac{1}{T}\int_{-T/2}^{T/2} x(t) \cdot x(t+\tau)\, dt \quad ...(5.102)$$

5.19.4. Stationary and Non-Stationary Random Process

A random process $X(t)$ is called stationary if its statistics are not affected by any shift in the time origin. We may define stationary process in terms of ensemble averages as:

(*i*) The ensemble mean is independent of time. Mathematically,

$$m_x(t) = m_x(t_1) = m_x(t_2) = m_x(t_3) = ... = \text{constant at all the time instants} \quad ...(5.103)$$

(*ii*) The autocorrelation function $R_\tau(t_1, t_2)$ depends only upon the time difference $t_2 - t_1$.

$$R_\tau(t_1, t_2) = R_\tau(t_1 + t, t_2 + t) \quad ...(5.104)$$

Therefore, the autocorrelation function of any stationary process is a function of time-difference and is given as

$$R_\tau(t_1, t_2) = R_\tau(t_2 - t_1) \quad ...(5.105)$$
$$= R_\tau(\rho)$$

where $\rho = t_2 - t_1$

Thus the system of resistors producing thermal noise voltages represents a stationary process due to the fact that the noise generated by resistor does not depend on time.

5.19.5. Wide Sense Stationary (*i.e.*, Weakly Stationary) Process

The process may not be stationary in strict sense, still the mean and autocorrelation functions are independent of shift of time region. Such type of process is known as **wide sense stationary process.**

All the processes in practice, are non-stationary since every process has some start and end. This means that the statistics of such process are dependent upon time. A true stationary process should start at $t = -\infty$ and should not stop till $t = \infty$. Such type of process is not possible practically.

The process may appear stationary over a certain period of time. Then this will be side sense stationary process.

5.19.6. Ergodic Process

A random process is known as ergodic process if the time-averages are equal to ensemble averages. Hence for a ergodic process, we have

$$m_x = <m_x> \quad ...(5.106)$$
$$R_\tau(t_1, t_2) = <R_\tau(\varepsilon)> \quad ..(5.107)$$

For any ergodic process, all ensemble averages will be equal to corresponding time averages for any particular sample function. The time averages are not a function of time. When time and ensemble averages are the same, it implies that ensemble averages also are not a function of time.

Thus an ergodic process is always stationary but converse in not true.

Ergodicity of the process may be defined in terms of some statistical averages like mean and autocorrelation.

The random process is ergodic in the mean if

$$m_x = <m_x>$$

and variance of $<m_x> \to 0$ as $T \to \infty$

Similarly, the random process is ergodic in the autocorrelation if

$$R_\tau(t_1, t_2) = <R_\tau(\varepsilon)> \quad ...(5.108)$$

and variance of $<R_\tau(\varepsilon)> \to 0$ as $T \to \infty$

Hence, the Ergodicity of any random process may be determined by evaluating statistical averages of single sample function. This means that a single sample function represents entire random process.

5.19.7. Gaussion Process

Let us consider a random process denoted by $X(t)$ for an interval which starts at a time $t = 0$ and lasts until $t = T$.

Now, let us weigh the random process $X(t)$ by any function $g(t)$ and then integrate the product $X(t)\, g(t)$ over this observation time-interval $(0 \le t \le T)$.

Thus, Gassuion we obtain a random variable Y defined as

$$Y = \int_0^T X(t)\, g(t)\, dt \quad ...(5.109)$$

@sTc

where Y is a linear function of $X(t)$.

Also, the random variable Y depends upon the argument function $X(t)\, g(t)$ over the observation time interval 0 to T. Here, we assume that the weighing function $g(t)$ is such that the mean square value of the random variable

$$Y = \int_0^T X(t)\, g(t)\, dt \quad \text{is finite.} \quad ...(5.110)$$

Now, if the random variable Y is a Gaussian distributed random variable for each value of function $g(t)$, then the process $X(t)$ is known as a **Gaussian process**. Also, a random variable Y has a Gaussian distribution if its probability density function has the following form:

$$f_Y(y) = \frac{1}{\sqrt{2\pi}\,\sigma_y} e^{\{[-(y - mY)^2]/2\sigma Y\}^2} \quad ...(5.111)$$

Here, m_Y is the mean and σ_Y^2 is the variance of the random variable Y. A Gaussian distributed random variable is completely characterized by spectifying its mean and the variance.

Figure 5.9. exhibits a plot of Gaussian distributed probability density function (pdf) for $m_Y = 0$ and $\sigma_Y^2 = 1$.

It may be noted that a Gaussian process has two main features or advantages:

(i) Firstly, the Gaussian process has many properties which make analytic results possible.

(ii) Secondly, the random process produced by physical phenomena are often such that a Gaussian model is an appropriate one.

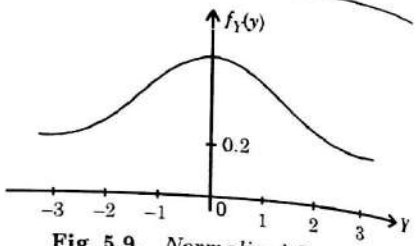


Fig. 5.9. *Normalized Gaussion distributed pdf of a random variably.*

Also the use of a Gaussian model to describe physical phenomena is generally confirmed by the experiments. Thus, the Gaussian process is very important in the study of communication systems.

A Gaussian process has the following important properties:

(i) If a Gaussian process $X(t)$ is made to apply to a stable linear filter, then the random process $Y(t)$ produced at the output of the filter is also Gaussian.

(ii) Let us consider the samples of random process $X(t_1), X(t_2), \ldots, X(t_n)$, obtained by observing a random process $X(t)$ at time $t = t_1, t = t_2, \ldots, t = t_n$.

Now, if the random process $X(t)$ is Gaussian, then this set of random variables are jointly Gaussian for any value of n, with their n-fold joint *pdf* being completely deterimined by specifying the set of means.

Here, means used for more than one mean (or average), *i.e.*

$$m_X(t_i) = E[X(t_i)],\ i = 1, 2, \ldots, n$$

and the set of autocorrelation function

$$C_X(t_j, t_i) = E[\{X(t_j) - m_X(t_j)\}\{X(t_i) - m_X(t_i)\}] \quad \ldots(5.112)$$

for $i = j = 1, 2, 3 \ldots, n$

(iii) Also, if a Gaussian process is wide-sense stationary (WSS), then the process is also stationary in the strict sense.

(iv) If the set of random variables $X(t_1), X(t_2) \ldots, X(t_n)$, obtained by sampling a Gaussian process $X(t)$ at times $t = t_1, t = t_2, \ldots, t = t_n$, are not correlated *i.e.*,

$$E[\{X(t_j) - m_X(t_j)\}\{X(t_i) - m_X(t_i)\}] = 0 \quad \ldots(5.113)$$

for $i \neq j$

then this set of random variables are statistically independent.

5.20. Sum of Random Processes

Now, we shall discuss the sum of two random processes.

Let us consider two WSS processes $X(t)$ and $Y(t)$ with zero means. The sum of random processes is another random process denoted as $Z(t)$. Also, the statistical parameters of $Z(t)$ may be determined in terms of random processes $X(t)$ and $Y(t)$.

Thus, the random process $Z(t)$ may be expressed as

$$Z(t) = X(t) + Y(t) \quad \ldots(5.114)$$

and the autocorrelation function of $Z(t)$ be found as

$$R_Z(\tau) = E[Z(t)\, Z(t-\tau)]$$
$$= E[\{X(t) + Y(t)\}\{X(t-\tau)\, Y(t-\tau)\}]$$
$$= E[X(t)\, X(t-\tau) + X(t)\, Y(t-\tau) + X(t-\tau)\, Y(t) + Y(t)\, Y(t-\tau)]$$

or $R_Z(\tau) = E[X(t)\, X(t-\tau)] + E[X(t)\, Y(t-\tau)] + E[X(t-\tau)\, Y(t)] + E[Y(t)\, Y(t-\tau)]$

or $R_Z(\tau) = R_X(\tau) + R_{XY}(\tau) + R_{YX}(\tau) + R_Y(\tau) \quad \ldots(5.115)$

The power spectral density (psd) of random process $Z(t)$ may be obtained by taking Fourier transform of both sides of equation (5.115)

Hence, we have $S_Z(\omega) = S_X(\omega) + S_{XY}(\omega) + S_{YX}(\omega) + S_Y(\omega) \quad \ldots(5.116)$

The power spectral density (psd) of the sum of two WSS random processes is equal to the sum of the individual power spectral densities (psds) plus the sum of the cross power spectral densities (psds) $S_{XY}(\omega)$ and $S_{YX}(\omega)$.

If random processes $X(t)$ and $Y(t)$ are non-correlated then the cross power spectral densities $S_{XY}(\omega)$ and $S_{YX}(\omega)$ are zero. Now equation (5.116) can be written as

$$S_Z(\omega) = S_X(\omega) + S_Y(\omega)$$

Hence, the power spectral density (psd) of the sum of two uncorrelated WSS random processes is equal to the sum of their individual power spectral densities.

5.21. Correlation Function

The correlation function provides a measure of similarity or coherence between a given signal (process) and a replica of the same signal or other signal (process) by a variable amount.

5.21.1. Autocorrelation Function

Autocorrelation function may be defined as a measure of similarity between a signal or process and its replica by a variable amount.

The autocorrelation function of a stationary process $X(t)$ may be defined as

$$R_{X(t_j - t_i)} = E[X(t_j)\, X(t_i)] \text{ for any } t_j \text{ and } t_i \quad \ldots(5.117)$$

where $X(t_j)$ and $X(t_i)$ are the random variables obtained by observing the process $X(t)$ at times t_j and t_i, respectively.

Also, the autocorrelation function depends only on the time difference $(t_j - t_i)$.

Using $\tau = t_j - t_i$ in equation (5.117), we get

$$R_X(\tau) = E[X(t)\, X(t-\tau)] \quad \ldots(5.118)$$

The expressions for $X(t)$ and $X(t-\tau)$ are viewed as random variables. The variable τ is known as **time-lag or time-delay parameter**. The autocorrelation function for stationary random process is independent of a shift of time origin.

5.21.2. Properties of Autocorrelation Function of Random Process

The autocorrelation function $R_X(\tau)$ of a wide-sense stationary (WSS) process has the following important properties:

(i) The autocorrelation function of WSS process $X(t)$ is always an even function of time-lag. Autocorrelation function $R_X(\tau)$ satistifies the following mathematical relationship, *i.e.*

$$R_X(\tau) = R_X(-\tau) \quad \ldots(5.119)$$

Proof: According to the definition of autocorrelation function of WSS random process, we have

$$R_X(\tau) = E[X(t)\, X(t-\tau)] \quad \ldots(5.120)$$

Substituting $\tau = -\tau$, in equation (5.120), we get

$$R_X(-\tau) = E[X(t)\, X(t+\tau)]$$

or

$$R_X(-\tau) = E[X(t+\tau)\, X(t)] \quad \ldots(5.121)$$

For WSS process, we know that

$$X(t)\, X(t-\tau) = X(t+\tau)\, X(t) \quad \ldots(5.122)$$

From last three equations, we conclude that

$$R_X(\tau) = R_X(-\tau)$$

(*ii*) The mean square value of a WSS process is equal to the autocorrelation function of the random process for zero time-lag. It may be written as

$$R_X(0) = R_X(\tau)\big|_{\tau=0} = E[X(t)\,X(t-\tau)]\big|_{\tau=0}$$

or $$R_X(0) = E[X(t)\,X(t-0)] = E[X^2(t)]$$

or $$R_X(0) = E[X^2(t)] \qquad ...(5.123)$$

(*iii*) The autocorrelation function of a WSS random process has the maximum magnitude at zero time-lag ($\tau = 0$). This property states that

$$|R_X(\tau)| = \text{magnitude of } R_X(\tau) = R_X(0) \qquad ...(5.124)$$

Proof: We know that the mean square value of the difference between $X(t)$ and $X(t-\tau)$ is always non-negative, *i.e.*

$$E[\{X(t) - X(t-\tau)\}^2] \geq 0$$

or $$E[X^2(t) - 2X(t)\,X(t-\tau) + X^2(t-\tau)] \geq 0 \qquad ...(5.125)$$

Since expectation operation is a linear operation, we can write above equation as

$$E[X^2(t)] - 2E[X(t)\,X(t-\tau)] + E[X^2(t-\tau)] \geq 0 \qquad ...(5.126)$$

Also, for a WSS random process, we have

$$E[X^2(t)] = E[X^2(t-\tau)] = R_X(0) \qquad ...(5.127)$$

and $$E[X(t)\,X(t-\tau)] = R_X(\tau) \qquad ...(5.128)$$

Substituting equations (5.131) and (5.127) in equation (5.128), we get

$$|R_X(\tau)| \leq R_X(0) \qquad ...(5.129)$$

Thus, the autocorrelation functions $R_X(\tau)$ provide means of decreasing the interdependance of two random variables obtained by observing a random process $X(t)$ τ seconds apart. The autocorrelation function is a measure of the rate of the fluctuation of random process. Various types of random processes have been illustrated in figure 5.10.

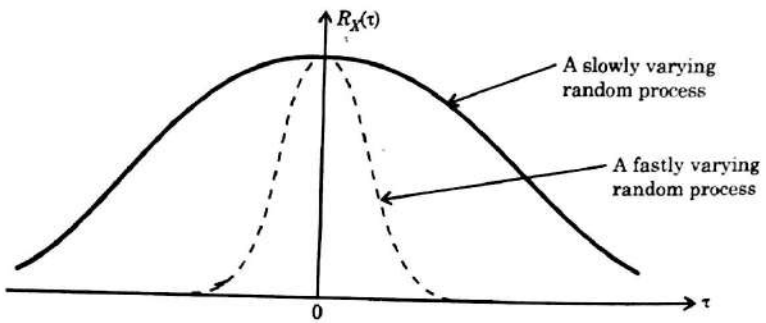


Fig. 5.10. *Illustration of autocorrelation functions of slowly and rapidly varying random processes.*

5.21.3. Cross-Correlation Functions

An autocorrelation function is determined for single random process but cross correlation function is determined for two random processes. Let us consider two random processes $X(t)$ and $X(t)$ with autocorrelation functions $R_X(t, u)$ and $R_Y(t, u)$, respectively. There will be two cross-correlation functions between two random process $X(t)$ and $Y(t)$. These cross-correlations may be defined as

$$R_{XY}(t, u) = E[X(t)\,Y(u)] \qquad ...(5.130)$$

and $$R_{YX}(t, u) = E[Y(t)\,X(u)] \qquad ...(5.131)$$

where t and u are the two values of time at which the processes are observed. The cross-correlation function of two random processes $X(t)$ and $Y(t)$ may be expressed conveniently in matrix form as

$$R(t, u) = \begin{bmatrix} R_X(t,u) & R_{XY}(t,u) \\ R_{YX}(t,u) & R_Y(t,u) \end{bmatrix} \qquad ...(5.132)$$

$R(t, u)$ is called the correlation matrix of random processes $X(t)$ and $Y(t)$.

Equation (5.137) may also be written as

$$R(t, u) = \begin{bmatrix} R_X(t-u) & R_{XY}(t-u) \\ R_{YX}(t-u) & R_Y(t-u) \end{bmatrix} \qquad ...(5.133)$$

If both random processes $X(t)$ and $Y(t)$ are WSS processes, then equation (5.132) can be written as equation (5.133). In addition, both random processes $X(t)$ and $Y(t)$ are said to be jointly wide-sense stationary.

It may be noted that in general, the cross-correlation function is not an even function of τ like an autocorrelation function. It is having following symmetry relation

$$R_{XY}(\tau) = R_{YX}(-\tau) \qquad ...(5.134)$$

Moreover, cross-correlation does not have a maximum at the origin like the autocorrelation function. The two random processes $X(t)$ and $Y(t)$ are said to be incoherent or orthogonal if cross-correlation function of $X(t)$ and $Y(t)$ is zero.

$$R_{XY}(\tau) = 0 \qquad ...(5.135)$$

The two random processes are said to be non-correlated, if their cross-correlation functions $R_{XY}(\tau)$ are equal to the multiplication of mean values.

$$R_{XY}(\tau) = E[X(t)]\,E[Y(t)] \qquad ...(5.136)$$

The incoherent or orthogonal processes are non-correlated processes with $E[X(t)]$ and or $E[Y(t)] = 0$.

5.22. Spectral Densities

As a matter of fact, spectral densities are used to represent random process in frequency-domain. In this article, we shall discuss frequency-domain description of random processes. Frequency-domain methods are very powerful tools for analysis of various types of signals. We have already discussed some powerful tools such as Fourier series, Fourier transform, Laplace transform for the analysis of deterministic signals. In this article, let us extend these tools to the stationary random processes.

5.22.1. Power Spectral Density (psd)

Power spectral density $S_X(\omega)$ of a wide-sense stationary (WSS) random process $X(t)$ may be defined as

$$S_X(\omega) = \text{Fourier transform } \{R_X(\tau)\}$$

$$= \int_{-\infty}^{\infty} R_X(\tau)\, e^{-j\omega\tau} d\tau \qquad ...(5.137)$$

where $R_X(\tau)$ = Autocorrelation function of random process $X(t)$.

Thus, we can say that the Fourier transform of autocorrelation function $R_X(\tau)$ of a random process $X(t)$ is called power spectral density (psd) of random process. Equation (5.137) was intially developed by Albert Einstein. Later it was developed by *N.* Weiner and Khenchine. Hence, this equation is known as Einstein-Weiner Khenchine (EWK) equation. Power spectral density (psd) is used to measure the energy of the signal in the frequency band $[\omega, \omega + d\omega]$.

Cross Power Spectral Density (CPSD).

The cross power spectral density of two jointly WSS random processes $X(t)$ and $Y(t)$ may be defined as

$$S_{XY}(\omega) = \text{Fourier transform } \{R_{XY}(\tau)\}$$

$$= \int_{-\infty}^{\infty} R_{XY}(\tau)\, e^{-j\omega\tau} d\tau \qquad ...(5.138)$$

where $R_{XY}(\tau)$ = cross-correlation function of random processes $X(t)$ and $Y(t)$.

Properties of Power Spectral Density

Power spectral density (psd) of a random process has the following main properties.

(i) *Power spectral density of a random process is a real function of frequency* ω.

Proof: We know that the autocorrelation function of a random process is an even function, *i.e.*

$$R_X(\tau) = R_X(-\tau)$$

Therefore,
$$S_X(\omega) = \int_{-\infty}^{\infty} R_X(\tau)\, e^{-j\omega\tau} d\tau$$

or
$$S_X(\omega) = \int_{-\infty}^{\infty} R_X(\tau)\, [\cos \omega\tau - j \sin \omega\tau]\, d\tau$$

$$S_X(\omega) = \int_{-\infty}^{\infty} R_X(\tau) \cos \omega\tau\, d\tau - j \int_{-\infty}^{\infty} R_X(\tau) \sin \omega\tau\, d\tau \qquad ...(5.139)$$

Now let $\int_{-\infty}^{\infty} R_X(\tau) \cos \omega\tau\, d\tau = I$ ...(5.140)

Substituting $t = -\tau$ in equation (5.140), we get

$$\int_{-\infty}^{\infty} R_X(-t) \sin(-\omega t)\,(-dt) = I$$

or
$$-\int_{-\infty}^{\infty} R_X(-t) \sin \omega t\, dt = I \qquad ...(5.141)$$

Using the even property of autocorrelation function, we obtain

$$-\int_{-\infty}^{\infty} R_X(\tau) \sin \omega\tau\, d\tau = I \qquad ...(5.142)$$

After adding equations (5.141) and (5.142), we get

$$I = 0 \qquad ...(5.143)$$

Substituting $\int_{-\infty}^{\infty} R_X(\tau) \sin \omega\tau\, d\tau = 0$ in equation (5.143), we get

$$S_X(\omega) = \int_{-\infty}^{\infty} R_X(\tau) \cos \omega\tau\, d\tau \qquad ...(5.144)$$

Here $S_X(\omega)$ is a real function of ω.

(ii) Power spectral density (psd) of a random process $X(t)$ is an even function of frequency ω, *i.e.*

$$S_X(\omega) = S_X(-\omega)$$

(iii) Power spectral density (psd) of a random process $X(t)$ is a non-negative function of ω, *i.e.*,

$$S_X(\omega) \geq 0 \text{ for all } \omega$$

5.22.2. Energy Spectral Density (ESD)

The energy spectral density (psd) $\psi_X(\omega)$ may be defined as a measure of density of the energy contained in random process $X(t)$ in Joules per Hertz. It may be noted that since the amplitude spectrum of a real-valued random process $X(t)$ is an even function of ω, the energy spectral density of such a signal is symmetrical about the vertical axis passing through the origin.

Thus, total energy of the random process $X(t)$ is defined as

$$E = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Psi_X(\omega)\, d\omega \qquad ...(5.145)$$

5.23. Response of Linear Systems to Random Inputs

In this article, we shall discuss about response of linear systems to random inputs. This means that we shall study transmission of random process through linear systems or linear filters.

Suppose that a random process $X(t)$ is applied to the input of a linear time-invariant (LTI) system having an impulse response $h(t)$. Let the input random process $X(t)$ produces a random process $Y(t)$ at the system output. The block diagram of this LTI system has been shown in figure 5.11. Here, we shall determine the mean and the auto-correlation function of the output random process $Y(t)$ in terms of input random process $X(t)$. Also, we are assuming that $X(t)$ is a wide-sense stationary (WSS) process.

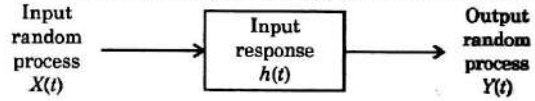


Fig. 5.11. *Determination of response of a linear system to a random input.*

The mean of the output random process is expressed as

$$m_Y(t) = E[Y(t)] = E\left[\int_{-\infty}^{\infty} h(\tau)\, X(t-\tau)\, d\tau\right]$$

or
$$m_Y(t) = \int_{-\infty}^{\infty} h(\tau)\, E[X(t-\tau)]\, d\tau = \int_{-\infty}^{\infty} h(\tau)\, m_X(t-\tau)]\, d\tau \qquad ...(5.146)$$

when the input random process $X(t)$ is wide-sense stationary, the mean $m_X(t)$ is a constant m_X. Therefore, equation (5.146) may be written as

$$m_Y = \int_{-\infty}^{\infty} h(\tau)\, m_X\, d\tau = m_X \int_{-\infty}^{\infty} h(\tau)\, d\tau = m_X H(0) \qquad ...(5.147)$$

where $H(0)$ = zero-frequency response of the system.

In other words, we can say that the mean of the output random process of a stable linear time-invariant (LTI) system is equal to the mean of the input random process multiplied by the zero-frequency response of the system.

We shall now determine the autocorrelation function of the output random process $Y(t)$. Let t and u be two values of time at which output process is observed. Autocorrelation function of output random process is defined as

$$R_Y(t, u) = E[Y(t)\, Y(u)] \qquad ...(5.148)$$

By using convolution intergal, equation (5.148) can be written as

$$R_Y(t, u) = E\left[\int_{-\infty}^{\infty} h(\tau_1)\, X(t-\tau_1)\, d\tau_1 \int_{-\infty}^{\infty} h(\tau_2)\, X(u-\tau_2)\, d\tau_2\right] \quad ...(5.149)$$

Assuming $E[X^2(t)]$ is finite for all t and the system is stable, we can interchange the order of the expectation operator and integration w.r.t. τ_1 and τ_2 to get

$$R_Y(t, u) = \int_{-\infty}^{\infty} h(\tau_1)\, d\tau_1 \int_{-\infty}^{\infty} h(\tau_2)\, d\tau_2\, E[X(t-\tau_1)\, X(u-\tau_2)]$$

$$= \int_{-\infty}^{\infty} h(\tau_1)\, d\tau_1 \int_{-\infty}^{\infty} h(\tau_2)\, d\tau_2\, R_X(t-\tau_1, u-\tau_2) \quad ...(5.150)$$

When the input $X(t)$ is a wide-sense stationary process, the autocorrelation function of $X(t)$ is only a function of the difference between the observation time $t-\tau_1$ and $u-\tau_2$. Substituting $\tau = t-u$ in equation (5.150), we get

$$R_Y(\tau) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\tau_1)\, h(\tau_2)\, R_X(\tau-\tau_1+\tau_2)\, d\tau_1\, d\tau_2 \quad ...(5.151)$$

$R_Y(\tau)$ is the function of time lag τ.

By combining the result of equation (5.147) with equation (5.151), we see that if the input to a stable LTI system is a WSS process, then the output of the linear time-invariant (LTI) system is also a WSS process.

Putting $\tau = 0$ in equation (5.147), we get

$$R_Y(0) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} h(\tau_1)\, h(\tau_2)\, R_X(\tau_2-\tau_1)\, d\tau_1\, d\tau_2 \quad ...(5.152)$$

Also we know that

$$R_Y(0) = E[Y^2(t)] \quad ...(5.153)$$

Example 5.29. A random variable X follows Gaussian probability density function with parameters μ and σ. Find its probability distribution function.

Solution: We know that the relationship between probability density function (pdf) and the probability distribution function of a random variable X is expressed as

$$F_X(z) = \int_{-\infty}^{z} f_x(x)\, dx \quad ...(i)$$

Since we have given that the random variable X follows Gaussian probability density function then

$$F_X(z) = \int_{-\infty}^{z} \frac{1}{\sqrt{2\pi\sigma}}\, e^{-(x-\mu)/2\sigma^2}\, dx \quad ...(ii)$$

Substituting $\frac{x-\mu}{\sigma} = y$ in equation (*ii*), we get

$$F_x(z) = \int_{-\infty}^{[(Z-\mu)/\sigma]} \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{-y^2/2}(\sigma dy) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{[(z-\mu)/\sigma]} e^{-y^2/2} dy$$

or $$F_x(z) = \frac{1}{\sqrt{2}}\left[\int_{-\infty}^{0} e^{-y^2/2}\, dy + \int_{0}^{(z-\mu)/2} e^{-y^2/2} dy\right]$$

$$= \left[-\int_{0}^{\infty} e^{-y^2/2} dy + \int_{0}^{(z-\mu)} e^{-y^2/2} dy\right] \quad ...(iii)$$

Also error function is defined as

$$erf(x) = \frac{1}{\sqrt{2\pi}} \int_{0}^{x} e^{-y^2/2}\, dy \quad ...(iv)$$

and $$erf(\infty) = \frac{1}{\sqrt{2\pi}} \int_{0}^{\infty} e^{-y^2/2} dy = \frac{1}{2} \quad ...(v)$$

Putting equations (*iv*) and (*v*) in equation (*iii*), we get

$$F_x(z) = \frac{1}{2} + erf\left(\frac{z-\mu}{\sigma}\right) \quad \textbf{Ans.}$$

Example 5.30. Prove that for a Gaussian random variable X,

$$P[x_1 \le X \le x_2] = erf\left(\frac{x_2-\mu}{\sigma}\right) - erf\left(\frac{x_2-\mu}{\sigma}\right)$$

here erf (x) = Error function

μ and σ = parameters of Gaussian probability density function and x_1, x_2 = any real numbers.

Solution: We know that

$$P[x_1 \le X \le x_2] = \int_{x_1}^{x_2} f_X(x)\, dx \quad ...(i)$$

here $$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma}\, e^{(x-\mu)^2/2\sigma^2} \quad ...(ii)$$

Substituting equation (*ii*) in equation (*i*), we get

$$P[x_1 \le X \le x_2] = \int_{x_1}^{x_2} \frac{1}{\sqrt{2\pi}\,\sigma}\, e^{(x-\mu)^2/2\sigma^2}\, dx = \frac{1}{\sqrt{2\pi}\,\sigma}\int_{x_1}^{x_2} e^{(x-\mu)^2/2\sigma^2}\, dx \quad ...(iii)$$

Putting $\left(\frac{x-\mu}{\sigma}\right) = y$ in equation (*iii*), we get

$$P[x_1 \le X \le x_2] = \frac{1}{\sqrt{2\pi}\,\sigma} \int_{[(x_1-\mu)/\sigma)}^{[(x_2-\mu)/\sigma]} e^{y^2/2}\sigma\, dy = \frac{1}{\sqrt{2\pi}} \int_{[(x_1-\mu)/\sigma]}^{[(x_2-\mu)/\sigma]} -e^{y^2/2}\, dy$$

$$= \frac{1}{\sqrt{2\pi}}\left[-\int_{[(x_1-\mu)/\sigma]}^{0} e^{y^2/2}dy + \int_{0}^{[(x_2-\mu)/\sigma]} e^{y^2/2}dy\right]$$

$$= \frac{1}{\sqrt{2\pi}}\left[-\int_{0}^{(x_1-\mu)/\sigma} e^{y^2/2}dy + \int_{0}^{(x_2-\mu)/\sigma} e^{y^2/2}\,dy\right]$$

$$= -\,erf\left(\frac{x_1-\mu}{\sigma}\right) + erf\left(\frac{x^2-\mu}{\sigma}\right) \qquad \textbf{Ans.}$$

Example 5.31. The joint probability density function of two random variables X and Y is given by $f_{XY}(x, y) = C(1 - x - y)$ for values of x and y for which (x, y) lies within the triangle as shown in figure 5.12.

Determine:

(i) the value of constant C

(ii) the marginal probability density function $f_X(x)$ and $f_Y(y)$. *(Kerla University, 1991)*

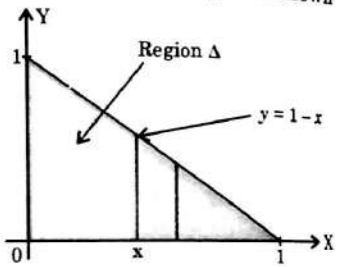


Fig. 5.12.

Solution: (i) The constant C may be found by using following condition

$$\iint_{\Delta} f_{XY}(x, y)\, dx\, dy = 1 \qquad ...(i)$$

The triangle is bounded by the region $x = 0$, $x + y = 1$.

Equation (i) may be written as

$$\int_0^1 \int_0^{1-x} C(1 - x - y)\, dx\, dy = 1$$

or $$\int_0^1 \frac{C(1-x)^2}{2}\, dx = 1$$

@sTc

or $$\frac{C}{6} = 1 \quad \text{or} \quad C = 6$$

(ii) The marginal probability density function $f_X(x)$ is expressed as

$$f_X(x) = \int_{-\infty}^{+\infty} f_{XY}(x, y)\, dy = \int_0^{1-x} 6(1 - x - y)\, dy = 6\left[-\frac{y^2}{2}\right]_0^{1-x}$$

or $$f_X(x) = \begin{cases} -3(1-x)^2, & \text{for } 0 \le x \le 1 \\ 0, & \text{for } x < 0,\ x \ge 1 \end{cases}$$

And $$f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y)\, dx = \int_0^{1-y} 6(1 - x - y)\, dx = 6\left[-\frac{x^2}{2}\right]_0^{1-y}$$

or $$f_Y(y) = \begin{cases} -3(1-y)^2, & \text{for } 0 \le y \le 1 \\ 0, & \text{for } y < 0,\ y \ge 1 \end{cases} \qquad \textbf{Ans.}$$

Example 5.32. The marginal probability density functions of two random variables X and Y are given below

$$f_X(x) = \begin{cases} 3(1-x)^2, & \text{for } 0 \le x \le 1 \\ 0, & \text{for } x < 0,\ x > 1 \end{cases}$$

and $$f_Y(y) = \begin{cases} 3(1-y)^2, & \text{for } 0 \le y \le 1 \\ 0, & \text{for } y < 0,\ y > 1 \end{cases}$$

Determine:

(i) Mean of random variables X and Y

(ii) Variance of random variables X and Y. *(WBTU, Kolkata-2003)*

Solution: (i) Mean of $$X = \mu_X = E(X) = \int_{-\infty}^{\infty} x f_X(x)\, dx$$

$$E(X) = \int_0^1 x\, 3(1-x)^2\, dx = \frac{1}{4}$$

and the mean of $Y = \mu_Y$

$$E(Y) = \int_{-\infty}^{\infty} y f_Y(y)\, dy = \int_0^1 y\, 3(1-y)^2\, dy = \frac{1}{4}$$

(b) Variance of $$X = \sigma_X^2 = E(X^2) - \mu_X^2 \qquad ...(i)$$

Now, $E(X^2)$ may be found as

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x)\, dx = \int_0^1 x^2\, 3(1-x)^2\, dx = \frac{1}{10}$$

Similarly, we can find $E(Y^2)$ as

$$E(Y^2) = \int_{-\infty}^{\infty} y^2 f_Y(y)\, dy = \int_0^1 y^2\, 3(1-y)^2\, dy = \frac{1}{10}$$

Substituting the values $E[X^2]$ and μ_X in equation, we get

$$\sigma_X^2 = E(X^2) - \mu_X^2 = \frac{1}{10} - \left(\frac{1}{4}\right)^2 = \frac{8-5}{80} = \frac{3}{80}$$

Similarly, variance of Y is determined as

$$\sigma_Y^2 = E(Y^2) - \mu_Y^2 \qquad ...(ii)$$

Putting the values of $E(Y^2)$ and μ_Y in equation (ii), we get

$$\sigma_Y^2 = \frac{1}{10} - \left(\frac{1}{4}\right)^2 = \frac{3}{80}$$

Example 5.33. Given two random processes $X(t)$ and $Y(t)$ as

$$X(t) = Z_1(t) + 3Z_2(t - \tau)$$
$$Y(t) = Z_2(t + \tau) + 3Z_1(t - \tau)$$

here $Z_1(t)$ and $Z_2(t)$ are independent white noise processes each with a variance equal to 0.5. Determine,

(i) autocorrelation functions of $X(t)$ and $Y(t)$,

(ii) cross-correlation function of $X(t)$ and $Y(t)$. *(U.P.S.C., I.E.S., Examination, 1997)*

Solution: (a) Autocorrelation function of $X(t)$ is expressed as

$$R_X(t_1, t_0) = E[X(t_1)\, X(t_0)] = E[\{Z_1(t_1) + 3Z_2(t_1 - \tau)\}\{Z_1(t_0)\, 3Z_2(t_0 - \tau)\}]$$
$$= E[Z_1(t_1)\, Z_1(t_0)] + 3E[Z_1(t_1) - Z_2(t_0 - \tau)] + 3E[Z_2(t_1 - \tau)\, Z_1(t_0)] + 9[Z_2(t_1 - \tau)\, Z_2(t_0 - \tau)] \quad ...(i)$$

But we know that

$$E[Z_1(t_1)\, Z_2(t_0 - \tau] = 0$$

Since, $Z_1(t)$ and $Z_2(t)$ are independent.

Similarly, $E[Z_2(t_1 - \tau)\, Z_1(t_0)] = 0$

For $t_1 \neq t_0$, $E[Z_1(t_1)\, Z_1(t_0)] = 0$

If $t_1 = t_0$ $E[Z_1(t_1)\, Z_1(t_0)] = E\,[Z_1^2(t_1)] = \sigma^2 + \mu^2 = 0.5\,\delta(t_1 - t_0)$ ...(ii)

Similarly $E[Z_2(t_1 - \tau)\, Z_2(t_0 - \tau)] = 0$ for $t_1 \neq t_0 = 0.5$ for $t_1 = t_0 = 0.5\,\delta(t_1 - t_0)$ ...(iii)

Substituting equations (ii) and (iii) in equation (i), we get

$$R_X(t_1, t_0) = 0.5\,\delta(t_1 - t_0) + 4.5\,\delta(t_1 - t_0) \quad ...(iv)$$

Similarly, $R_X(t_1, t_0) = 0.5\,\delta(t_1 - t_0) + 4.5\,\delta(t_1 - t_0)$ ...(v)

From equations (iv) and (v), we see that the random processes $X(t)$ and $Y(t)$ are wide-sense stationary (WSS).

(ii) Cross-correlation function of random processes $X(t)$ and $Y(t)$ is expressed as

$$R_{XY}(t_1, t_0) = E[X(t_1)\, Y(t_0)] = E[\{Z_1(t_1) + 3Z_2(t_1)\}\{Z_2(t_0 + \tau) + 3Z_1(t_0 - \tau)\}]$$
$$= E[Z_1(t_1)\, Z_2(t_0 + \tau)] + 3E[Z_1(t_1)\, Z_1(t_0 - \tau)]$$
$$+ 3E[Z_2(t_1) Z_2(t_0 + t)] + 9E[Z_2(t_1)\, Z_1(t_0 - \tau) \quad ...(vi)$$

Now $E[Z_1(t_1)\, Z_2(t_0)] = 0$ (Since $Z_1(t_1)$ and $Z_2(t_0)$ are independent random variables)

Therefore, $$E[Z_1(t_1) Z_1(t_0 - \tau)] = \begin{cases} 0, & \text{if } t_1 \neq t_0 - \tau \\ 0.5, & \text{if } t_1 = t_0 - \tau \end{cases}$$

$$E[Z_2(t_1 - \tau) Z_2(t_0)] = \begin{cases} 0, & \text{if } t_0 \neq t_1 - \tau \\ 0.5, & \text{if } t_0 = t_1 - \tau \end{cases}$$

Now, $E[Z_2(t_1 - \tau)\, Z_1(t_0 - \tau)] = 0$

Substituting above values in equation (vi), we get

$$R_{XY}(t_1, t_0) = 1.5\delta(t_1 - t_0 + \tau) + 1.5\,\delta(t_0 - t_1 + \tau) \quad ...(vii)$$

Because cross-correlation depends upon the difference of t_1 and t_0 therefore, random processes $X(t)$ and $Y(t)$ are jointly stationary or simply stationary. **Ans.**

Example 5.34. Stationary random process $X(t)$ has the following autocorrelation function

$$R_X(\tau) = \sigma^2 e^{-\mu|\tau|}$$

where μ and σ^2 are constants. It is passed through a filter whose impulse response is

$$h(\tau) = \alpha\, e^{-\alpha\tau}\, u(\tau)$$

here α is a contant and $u(\tau)$ is a step function.

(i) Find the power spectral density of random signal $X(t)$.

(ii) Find the power spectral density of the output random signal $Y(t)$.

(GATE Examination, 1998)

Solution: (i) Power spectral density (psd) of the input is expressed as

$$S_X(\omega) = \text{CTFT}\,\{R_X(t)\} = \int_{-\infty}^{\infty} R_X(\tau)\, e^{-j\omega\tau}\, d\tau$$

or
$$S_X(\omega) = \int_{-\infty}^{\infty} \sigma^2\, e^{-\mu|\tau|}\, e^{-j\omega\tau}\, \tau = \sigma^2 \int_{-\infty}^{0} e^{(\mu - j\omega)\tau}\, d\tau + \sigma^2 \int_{0}^{\infty} e^{-(\mu + j\omega)t}\, d\tau$$

or
$$S_X(\omega) = \sigma^2 \left[\frac{1}{\mu - j\omega} + \frac{1}{\mu + j\omega}\right] = \frac{2\mu\sigma^2}{\mu^2 + \omega^2}$$

(ii) Now, power spectral density of the output is related to the power spectral density of the input as

$$S_Y(\omega) = \{H(\omega)\}^2\, S_X(\omega) \quad ...(i)$$

and
$$H(e^{j\omega}) = \text{CTFT}\,\{h(\tau)\}$$

or
$$H(e^{j\omega}) = \int_{-\infty}^{\infty} h(\tau)\, e^{-j\omega\tau}\, d\tau = \int_{-\infty}^{\infty} \alpha\, e^{-\alpha\tau}\, u(\tau)\, e^{-j\omega\tau}\, d\tau$$

or
$$H(e^{j\omega}) = \alpha \int_{0}^{\infty} e^{-(\alpha + j\omega)\tau}\, d\tau = \frac{\alpha}{\alpha + \omega} \quad ...(ii)$$

Substituting equation (ii) in equation (i), we obtain

$$S_Y(jw) = \left|\frac{\alpha}{\alpha + j\omega}\right|^2 S_X(\omega) = \frac{\alpha}{\alpha^2 + \omega^2} S_X(\omega)$$

or
$$S_Y(jw) = \frac{\alpha}{\alpha^2 + \omega^2} \frac{2\mu\sigma^2}{\mu^2 + \omega^2} = \frac{2\mu\alpha^2\sigma^2}{(\alpha^2 + \omega^2)(\mu^2 + \omega^2)} \quad \textbf{Ans.}$$

Example 5.35. A deterministic signal $x(t) = \cos 2\pi t$ is passed through a differentiator as shown in figure 5.13.

(a) Determine the autocorrelation $R_{xx}(\tau)$ and the power spectral density $S_{xx}(f)$.

(b) Find the output power spectral density $s_{yy}(f)$.

(c) Evaluate $R_{xy}(0)$ and and $R_{xy}\left(\frac{1}{4}\right)$.

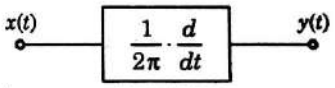


Fig. 5.13.

(GATE Examination, 2002)

Solution: We know that the autocorrelation is expressed as

$$R_{xx}(\tau) = E[X(t + \tau).X(t)]$$

or
$$R_{xx}(\tau) = \frac{1}{T} \int_{-T/2}^{T/2} \cos[2\pi(t + \tau)] \cos 2\pi t dt$$

or
$$R_{xx}(\tau) = \frac{1}{2T} \int_{-T/2}^{T/2} [\cos 2\pi(2t + \tau) + \cos 2\pi\tau]\, dt$$

or $$R_{xx}(\tau) = \frac{1}{2T}\left[\frac{\sin 2\pi(2t+\tau)}{4\pi}\Bigg|^{T/2} + (\cos 2\pi\tau).t\Big|_{-T/2}^{T/2}\right]$$

Simplifying, we get

$$R_{xx}(t) = \frac{\cos 2\pi\tau}{2}$$

Further, we know that

$$S_{xx}(f) = \int_{-\infty}^{\infty} R_{xx}(\tau).e^{-j\omega\tau}d\tau = \frac{1}{2}[\delta(f-1)+\delta(f+1)]$$

We have, using

$$\cos(2\pi f_c t) = \frac{1}{2}[\delta(f-f_c)+\delta(f+f_c)]$$

(b) Now, $$R_{yy}(\tau) = \frac{d}{dt}.R_{xx}(\tau) = \pi\sin(2\pi\tau)$$

Also,

$$S_{yy}(f) = H(j\omega)\,.\,S_{xx}(f) = j\omega\,.\,S_{xx}(f)$$

or $$S_{yy}(f) = (j.2\pi f).\frac{1}{2}[\delta(f-1)+\delta(f+1)]$$

$$= (j\pi f).\,[\delta(f-1)+\delta(f+1)] \qquad \textbf{Ans.}$$

(c) Again, we have

$$R_{xy}(t) = E[X(t).\,Y(t+\tau)]$$

or $$R_{xy}(t) = \frac{2\pi}{T}\int_{-T/2}^{T/2}[\sin 2\pi(2t+\tau)+\sin(2\pi\tau)].d\tau$$

or $$R_{xy}(\tau) = \frac{\pi}{T}\left[\frac{\cos 2\pi(2t+\tau)}{4\pi}\Bigg|_{-T/2}^{T/2} + \sin 2\pi\tau.T\right] = \pi.\sin 2\pi\tau$$

then $$R_{xy}(0) = 0$$

$$R_{xy}\left(\frac{1}{4}\right) = \pi.\sin\frac{\pi}{2} = \pi \qquad \textbf{Ans.}$$

SUMMARY

1. There is one other class of signals, the behaviour of which cannot be predicted. Such type of signals are called **random signals**.
2. These signals are called random signals because the precise value of these signals cannot be predicted in advance before they actually occur. The examples of random signals are the noise interferences in communication systems.
3. An experiment is defined as the process which is conducted to get some results. If the same experiment is performed repeatedly under the same conditions, similar results are expected.
4. The outcomes of an event are called **equally** likely, if any one of them cannot be expected to occur in preference to another. As an example, the tossing of a coin results in two outcomes, Head and Tail. Both have same possibility of 50%. Such type of outcomes are called **equally** likely outcomes.

5. Probability may be defined as the study of random experiments. In any random experiment, there is always an uncertainty that a particular event will occur *or* not.
6. The concept of conditional probability is used in conditional occurrences of the events.
7. The conditional probability of event B given that event A has already happened.

$$P(B/A) = \frac{P(AB)}{P(A)}$$

where $P(AB)$ is the joint probability of A and B.

Similarly, the conditional probability of event A given that event B has already happened.

$$P(A/B) = \frac{P(AB)}{P(B)}$$

The joint probability has commutative property which states that

$$P(AB) = P(BA)$$

8. If A and B are two events in an experiment, and possibility of occurrence of event B does not depend upon occurrence of event A, then these two events A and B are known as statistically independent events.
9. A function which can take on any value from the sample space and its range is some set of real numbers is called a **random variable** of the experiment.
10. Random variables may be classified as under:
    1. Discrete random variables
    2. Continuous random variables.
11. A discrete random variable may be defined as the random variable which can take on only finite number of values in a finite observation interval. This means that the discrete random variable has countable number of distinct values.
12. A random variable that takes on an infinite number of values is called a continuous random variable. Actually, there are several physical system (experiments) that generate continuous outputs *or* outcomes. Such systems generate infinte number of outputs or outcomes within the finite period.
13. The Comulative Distribution Function (CDF) of a random variable 'X' may be defined as the probability that a random variable 'X' takes a value less than or equal to x.
14. The derivative of cumulative distribution function (CDF) with respect to some dummy variable is known as Probability Density Function (PDF). Probability density function (PDF) is generally denoted by $f_X(x)$. Mathematically, PDF may be expressed as

$$PDF: f_X(x) = \frac{d}{dx}F_X(x)$$

where x is a dummy variable.

Probability density function (PDF) is the more convenient representation for **continuous random variable**.

15. The joint Cumulative Distribution Function may be defined systematically as:

The joint Cumulative Distribution Function $F_{XY}(x, y)$ may be defined as the probability that the outcome of an experiment will result in a sample point lying inside the range $(-\infty < X \le x, -\infty < Y \le y)$ of the joint sample space. The joint sample space is the combined sample space of X and Y. Mathematically,

$$F_{XY}(x, y) = P(X \le x,\ Y \le y)$$

16. The joint PDF of any two random variables X and Y may be defined as the partial derivative of the joint cumulative distribution function $F_{XY}(x, y)$ with respect to the dummy variables x and y.

Mathematically,

$$f_{XY}(x, y) = \frac{\partial^2 F_{XY}(x, y)}{\partial x \, \partial y}$$

17. The relationship between probability and PDF over a certain interval is expressed as

$$P(x_1 < X \le x_2) = \int_{x_1}^{x_2} f_X(x)\, dx$$

18. In case, when the probability density functions $f_X(x)$ and $f_Y(y)$ for any single random variable are obtained from joint PDF, then $f_X(x)$ and $f_Y(y)$ are called marginal PDF *or* simply marginal densities.
19. Out of the two random variables, one variable may take a fixed value. In this case, the PDF is called *conditional*.
20. There are some other measures or numbers which give more useful and quick information about the random variable. Collectively these characteristic numbers or measures are known as **statistical averages**.
21. The mean or average of any random variable is expressed by the summation of the values of random variables X weighted by their probabilities. Mean value of a random variable is denoted by m_x. Mean value is also known as expected value of random variable X.

$$m_x = E(X)$$

where $E[\ ]$ represents expection operator.

22. Also, in general term, mean or average is given by

$$\text{Mean or average of } X = \frac{\text{Arithmatic sum of all values of } X}{\text{Total number of values of } X}$$

23. Mean value of Discrete random variable

$$m_x = E(X) = \overline{X} = \sum_{i=1}^{n} x_i . P(x_i)$$

Here $\overline{X}$ is also the notation for mean values.

24. Continuous random variable X, the mean or average value is expressed as

$$m_x = \int_{-\infty}^{\infty} x f_X(x)\, dx$$

where $f_X(x)$ = Probability density function (PDF)

25. The nth moment of any random variable X may be defined as the mean value of X^n.

i.e. $\quad g(x) = X^n$

26. Variance is generally represented by σ_x^2 *i.e.*,

$$\sigma_x^2 = \text{Variance } (X)$$
$$\sigma_x^2 = \text{Var } (X)$$

or

$$\sigma_x^2 = E[(X - m_x)^2]$$

or

$$\sigma_x^2 = \int_{-\infty}^{\infty} (x - m_x)^2 f_X(x)\, dx.$$

27. The square root of variance is known as **standard deviation** of a random variable X. Standard deviation gives the measure of spread observed over the values of X related to mean value. Therefore,

$$\text{Standard deviation} = \sqrt{\text{Variance}} = \sqrt{\sigma_x^2} = \sigma_x .$$

28. The Probability Density Function (PDF) of an uniform distribution is expressed as

Uniform PDF: $f_X(x) = 0$ for $x < m - \frac{A}{2}$ and $x > m + \frac{A}{2}$

$$= \frac{1}{A} \quad \text{for} \left(m - \frac{A}{2}\right) \le x \le \left(m + \frac{A}{2}\right)$$

29. The value of PDF, $f_X(x)$ is same for all possible values of a random variable. Thus, this distribution is known as **Uniform Distribution**.
30. Gaussian Distribution is also known as Normal Distribution. It is defined for a continuous random variable. The *PDF* for a Gaussian random variable is expressed as

Gaussian PDF

$$f_X(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-m)^2/2\sigma^2}$$

In this equation $\quad m$ = mean value of the random variable

σ^2 = variance of the random variable.

31. Rayleigh Distribution is used for continuous random variables. It is produced from two Gaussian random variables. Let X and Y be independent Gaussian random variables having mean value zero and variance σ^2. Mathematically,

$$m_x = m_y = 0$$

and $\quad \sigma_x^2 = \sigma^2$

32. Rayleigh Distribution is always used for modelling of statistics of signals transmitted through radio channels such as cellular radio.
33. Let there be a random experiment E having outcome λ from the sample sapce S. This means that $\lambda \in S$. Thus every-time an experiment is conducted, the outcome λ will be one of the sample point in sample space. If this outcome λ is associated with time, then a function of λ and time t is formed *i.e.*, $X(\lambda, t)$. Then the function $X(\lambda, t)$ is known as **random process**.
34. When the statistical averages are taken along the time; they are known as time averages. As an example, we may define time mean value of a sample function $x(t)$ as

Time mean value: $$<m_x> = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} X(t)\, dt$$

35. The autocorrelation function may be expressed using time averaging as

$$<R_\tau(\varepsilon)> = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{T/2} x(t) \cdot x(t + \tau)\, dt$$

36. A random process $X(t)$ is called stationary if its statistics are not affected by any shift in the time origin.
37. The process may not be stationary in strict sense, still the mean and autocorrelation functions are independent of shift of time region. Such type of process is known as wide sense stationary process.
38. A random process is known as ergodic process if the time-averages are equal to ensemble averages. Hence for a ergodic process, we have

$$m_x = <m_x>$$
$$R_\tau(t_1, t_2) = <R_\tau(\varepsilon)>$$

39. It may be noted that a Gaussian process has two main advantages:
    (*i*) Firstly, the Gaussian process has several properties which make analytic results possible.
    (*ii*) Secondly, the random process produced by physical phenomena are often such that a Gaussian model may appropriate.
40. A Gaussian process has the following important properties:
    (*i*) If a Gaussian process $X(t)$ is made to apply to a stable linear filter, then the random process $Y(t)$ produced at the output of the filter is also Gaussian.
    (*ii*) Let us consider the samples of a random process $X(t_1)$, $X(t_2)$, ..., $X(t_n)$, obtained by observing a random process $X(t)$ at time $t = t_1$, $t = t_2$,....., $t = t_n$.
41. Also, if a Gaussian process is wide-sense stationary (WSS), then the process is also stationary in the strict sense.
42. The power spectral density of the sum of two uncorrelated WSS random processes is equal to the sum of their individual power spectral densities.
43. The correlation function provides a measure of similarity or coherence between a given signal and a replica of the same signal or other signal by a variable amount.
44. Autocorrelation function may be defined as a measure of similarity between a signal or process and its replica by a variable amount.
    The autocorrelation function of a stationary process $X(t)$ may be defined as
    $$R_{X(t_j - t_i)} = E[X(t_j)\, X(t_i)] \text{ for any } t_j \text{ and } t_i$$
    here $X(t_j)$ and $X(t_i)$ are the random variables obtained by ovserving the process $X(t)$ at times t_j and t_i, respectively.
45. The autocorrelation function of WSS process $X(t)$ is always an even function of time-lag. Autocorrelation function $R_X(\tau)$ satistifies the following mathematical relationship, *i.e.*
    $$R_X(\tau) = R_X(-\tau).$$
46. The mean square value of a WSS process is equal to the autocorrelation function of the random process for zero time-lag. It may be written as
    $$R_X(0) = R_X(\tau)\big|_{\tau=0} = E[X(t)\, X(t-\tau)]\big|_{\tau=0}$$
    or
    $$R_X(0) = E[X(t)\, X(t-0)] = E[X^2(t) = E[X^2(t)]$$
47. The autocorrelation function of a WSS random process has the maximummagnitude at zero time-lag ($\tau = 0$). This property states that
    $$|R_X(\tau)| = \text{magnitude of } R_X(\tau) = R_X(0)$$
48. Spectral densities are used to represent random process in frequency-domain. In this article, we will discuss frequency-domain description of random processes. Frequency-domain methods are very powerful tools for analysis of various types of signals.
49. Power spectral density $S_X(\omega)$ of a wide-sense stationary (WSS) random process $X(t)$ may be defined as
    $$S_X(\omega) = \text{Fourier transform } \{R_X(\tau)\} = \int_{-\infty}^{\infty} R_X(\tau)\, e^{-j\omega\tau} d\tau$$
    here $R_X(\tau)$ = Autocorrelation function of random process $X(t)$.
50. The cross power spectral density of two jointly WSS random processes $X(t)$ and $Y(t)$ may be defined as density of two jointly WSS random processes $X(t)$ and $Y(t)$ is defined as
    $$S_{XY}(\omega) = \text{Fourier transform } \{R_{XY}(\tau)\} = \int_{-\infty}^{\infty} R_{XY}(\tau)\, e^{-j\omega\tau} d\tau$$
    here $R_{XY}(t)$ = Cross-correlation function of random processes $X(t)$ and $Y(t)$.

51. The energy spectral density $\psi_X(\omega)$ may be defined as a measure of density of the energy contained in random process $X(t)$ in Joules per Hertz. It may be noted that since the amplitude spectrum of a real-valued random process $X(t)$ is an even function of ω, the energy spectral density of such a signal is symmetrical about the vertical axis passing through the origin.
    Thus, total energy of the random process $X(t)$ is defined as
    $$E = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Psi_X(\omega)\, d\omega$$

www.dahiyastech.blogspot.com

REVIEW QUESTIONS

1. Define probability. Explain conditional and joint probabilities.
2. State and prove different properties of probability. *(WBTU, Kolkata-2002)*
3. How the probabilities of statistically independent events are related with each other.
4. Define random variables and differentiate between discrete and continuous random variables.
5. Define CDF. Explain different properties of CDF.
6. Define PDF. Explain different properties of PDF. *(Kerla University, 1998)*
7. Derive the relationship between CDF and PDF.
8. Define joint distribution function. Explain its various properties. *(Anna University, Chennai-2003)*
9. Define joint PDF. Explain its various properties.
10. Explain the relationship between joint *PDF* and probability.
11. Define marginal densities.
12. Explain conditional PDF and its properties.
13. Define mean or average or expected value of a random variable. Explain how it is calculated for continuous and discrete random variables. *(A.M.I.E., Examination, 1999)*
14. Explain variance and standard deviation of a random variable.
15. Define random process and explain ensemble and sample function.
16. Explain ensemble and time averages.
17. Explain the stationary random process with the help of one example.
18. Define ergodic process. Explain the difference between ergodic process and stationary process. *(PTU, 1998)*

SHORT QUESTIONS WITH ANSWERS

Q.1. What are random signals? What is significance of random signals in probability theory?

Ans. There is one other class of signals, the behaviour of which cannot be predicted. Such type of signals are called **random signals**. These signals are called random signals because the precise value of these signals cannot be predicted in advance before they actually occur. The examples of random signals are the noise interferences in communication systems. This means that the noise interference during transmission is totally unpredictable. In the same way, the noise generated by the receiver itself is random. Even some other signals which are not noise signals are also random signals. These signals cannot be modelled mathematically. Actually the electro magnetic interference is the major source of random noise.

Q.2. What do you mean by Experiment in probability theory?

Ans. An experiment is defined as the process which is conducted to get some results. If the same experiment is performed repeatedly under the same conditions, similar results are expected. But there are few experiments which do not produce the same result as we have stated above. The theory of probability

mainly deals with such type of experiments. In fact, the theory of probability is applied to such type of experiments to predict the possibility of a particular output. An experiment is sometimes called trial. As an example, throw of a coin is an experiment or trial. This trial results in two outcomes namely Head and Tail.

Q.3. What is sample space?

Ans. A set of all possible outcomes of an experiment *or* trial is called the **sample space** of that experiment. It is generally denoted by 'S'. The total number of outcomes in a sample space is denoted by $n(s)$. As an example, the tossing of a coin has two outcomes. Hence, we may write its sample space as:

$$S = \{H, T\}$$

where, $H \rightarrow$ Head and $T \rightarrow$ Tail

Q.4. What is an in probability theory?

Ans. The expected subset of the sample space *or* happening is called an event. As an example, let us consider an experiment of throwing a cubic die. In this case, the sample space S will be as

$$S = \{1, 2, 3, 4, 5, 6\}$$

Now, if we want the number '3' to be an outcome or an even number, *i.e.*, $\{2, 4, 6\}$, then this subset is called an event. This is denoted by letter 'E'. Hence event E is a subset of the sample space 'S'. If event E has only one outcome, then it is called an elementary event. On the other hand, if event E does not contain any out come, then it is called a null event. If $E = S$, then an event contains all the outcomes. Such as event is called a certain event. It always occurs, no matter what so ever is the outcome.

Q.5. What do you mean by probability? Explain.

Ans. Probability may be defined as the study of random experiments. In any random experiment, there is always an uncertainty that a particular event will occur *or* not. As a measure of probability of occurrence of an event, a number between 0 to 1 is assigned. If it is sure that an event will occur, then we can say that its probability is 100% or 1. If it is not sure that an event will occur, then we can say that its probability is 0% or 0. If it is not sure whether the event will occur *or* not, then its probability is between 0 and 1.

Q.6. Write properties of probability? Explain.

Ans. ***Property* 1:** The probability of a certain event is unity *i.e.*,

$$P(A) = 1 \quad \text{...(5.1)}$$

***Property* 2:** The probability of any event is always less than *or* equal to 1 and non-negative. Mathematically,

$$0 \le P(A) \le 1 \quad \text{...(5.2)}$$

***Property* 3:** If A and B are two mutually exclusive events, then

$$P(A + B) = P(A) + P(B) \quad \text{...(5.3)}$$

***Property* 4:** If A is any event, then the probability of not happening of A is

$$P(\overline{A}) = 1 - P(A) \quad \text{...(5.4)}$$

where $\overline{A}$ represents the complement of event A.

***Property* 5:** If A and B are any two events (not mutually exclusive events), then

$$P(A + B) = P(A) + P(B) - P(AB) \quad \text{...(5.5)}$$

where $P(AB)$ is called the probability of events A and B both occurring simultaneously. Such an event is called joint event of A and B, and the probability $P(AB)$ is called the joint probability.

Now, if events A and B are mutually exclusive, then the joint probability, $P(AB) = 0$.

Q.7. What is conditional probability?

Ans. The concept of conditional probability is used in conditional occurrences of the events. Let us consider an experiment which involves two events A and B. Now the probability of event B, given that event A has occurred, is represented by $P(B/A)$. Similarly, $P(A/B)$ represents probability of event A given that event B has already occurred. Therefore, $P(B/A)$ and $P(A/B)$ are called conditional probabilities.

Q.8. Explain random variable?

Ans. A function which can take on any value from the sample space and its range is some set of real numbers is called a random variable of the experiment. Random variables are denoted by upper case letters such as X, Y etc. and the values taken by them are denoted by lower case letters with subscripts such as x_1, x_2, y_1, y_2 etc.

Random variables may be classified as under:

1. Discrete random variables
2. Continuous random variables.

Q.9. What is a continuous random variable?

Ans. A random variable that takes on an infinite number of values is called a continuous random variable. Actually, there are several physical system (experiments) that generate continuous outputs *or* outcomes. Such systems generate infinte number of outputs or outcomes within the finite period. Continuous random variables may be used to define the outputs of such systems.

As an example, the noise voltage generated by an electronic amplifier has a continuous amplitude. This means that sample space S of the noise voltage amplitude is continuous. Therefore, in this case, the random variable X has a continuous range of values.

Q.10. What do you mean by probability density function (PDF)?

Ans. The derivative of cumulative distribution function (CDF) with respect to some dummy variable is known as Probability Density Function (PDF). Probability density function (PDF) is generally denoted by $f_X(x)$. Mathematically, PDF may be expressed as

$$\text{PDF: } f_X(x) = \frac{d}{dx} F_X(x)$$

where x is a dummy variable.

Probability density function (PDF) is the more convenient representation for continuous random variable.

Q.11. What do you mean by conditional probability density function?

Ans. Out of the two random variables, one variable may take a fixed value. In this case, the PDF is called *conditional*. As an example, out of the two continuous random variables X and Y, let $X = x$.

Then we may find the conditional, *PDF* of Y given that $X = x$ as,

$$f_Y(y/x) = \frac{f_{XY}(x, y)}{f_X(x)}$$

where $f_X(x)$ is the marginal density of random variable X.

Similarly, we may find the conditional PDF of X given that $Y = y$. Then

$$f_X(x/y) = \frac{f_{XY}(x, y)}{f_Y(y)}$$

Q.12. What is mean or average?

Ans. The mean or average of any random variable is expressed by the summation of the values of random variables X weighted by their probabilities. Mean value of a random variable is denoted by m_x.

Mean value is also known as expected value of random variable X.

$$m_x = E(X)$$

where $E[\ ]$ represents expection operator.

Therefore both m_x and $E(X)$ have the same meaning.

Also, in general term, mean or average is given by

$$\text{Mean or average of } X = \frac{\text{Arithmatic sum of all values of } X}{\text{Total number of values of } X}$$

Q.13. Explain uniform distribution?

Ans. The Probability Density Function (PDF) of an uniform distribution is expressed as

$$\textit{Uniform PDF: } f_X(x) = 0 \quad \text{for } x < m - \frac{A}{2} \;\&\; x > m + \frac{A}{2}$$

$$= \frac{1}{A} \quad \text{for } \left(m - \frac{A}{2}\right) \le x \le \left(m + \frac{A}{2}\right) \qquad \text{...(5.84)}$$

Figure 5.14 shows the PDF of an uniformly distributed random variable.

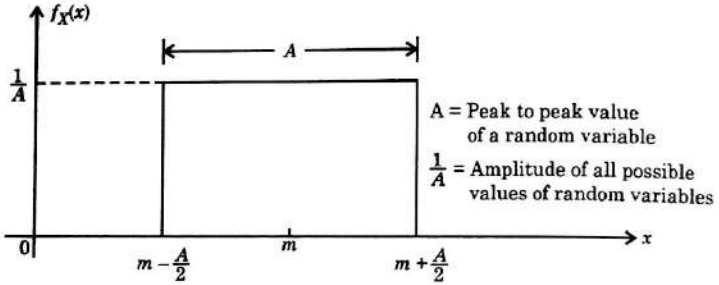


Fig. 5.14. *PDF of an uniformly distributed random variable*

In above figure, the peak to peak value is A and amplitude is uniform *i.e.* A. From figure 5.3, it is also clear that the random variable and its PDF are continuous. Therefore uniform distribution is utilised for continuous random variable.

Also, the value of PDF, $f_X(x)$ is same for all possible values of a random variable. Thus this distribution is known as Uniform Distribution.

Q.14. What is importance of Gaussian distribution?

Ans. The Gaussian distribution is used for continuous random variables. We know that the random motion of the thermally agitated electrons produces thermal noise. This thermal noise has Gaussian distribution. The random errors in the experimental measurements create the measured values to have Gaussian distribution about the true value. The Gaussian distribution is very important in the analysis of communication and statistical systems.

Q.15. What do you mean by random process? Explain.

Ans. Let there be a random experiment E having outcome λ from the sample sapce S. This means that $\lambda \in S$. Thus every time an experiment is conducted, the outcome λ will be one of the sample point in sample space. If this outcome λ is associated with time, then a function of λ and time t is formed *i.e.*, $X(\lambda, t)$. Then the function $X(\lambda, t)$ is known as random process. Hence when any random experiment E is given a time dimension, then each outcome appears at some certain time and the random experiment will be converted to Random Process. A random process is the function of two variables λ and t.

Q.16. What is Ergodic process? Explain.

Ans. A random process is known as ergodic process if the time-averages are equal to ensemble averages. Hence for a ergodic process, we have

$$m_x = \langle m_x \rangle$$
$$R_\tau(t_1, t_2) = \langle R_\tau(\varepsilon) \rangle$$

For any ergodic process, all ensemble averages will be equal to corresponding time averages for any particular sample function. The time averages are not a function of time. When time and ensemble averages are the same, it implies that ensemble averages also are not a function of time.

Thus an ergodic process is always stationary but converse in not true.

Ergodicity of the process may be defined in terms of some statistical averages like mean and autocorrelation.

The random process is ergodic in the mean if

$$m_x = \langle m_x \rangle$$

and variance of $\langle m_x \rangle \to 0$ as $T \to \infty$

Similarly, the random process is ergodic in the autocorrelation if

$$R_\tau(t_1, t_2) = \langle R_\tau(\varepsilon) \rangle$$

and variance of $\langle R_\tau(\varepsilon) \rangle \to 0$ as $T \to \infty$

Hence, the Ergodicity of any random process may be determined by evaluating statistical averages of single sample function. This means that a single sample function represents entire random process.

NUMERICAL PROBLEMS

1. A wheel of chance is divided into three equal segments coloured Green (G), Red (R) and Yellow (Y). The wheel is spinned twice and the outcome is taken as resulting colour sequence. That is GR, RG etc. Let A be the event which represents "neither colour is Yellow". Let B be the event which represents "matching colours". Calculate $P(A)$, $P(B)$, $P(AB)$ and $P(A + B)$.

   [Ans. $P(A) = \frac{4}{9}$, $P(B) = \frac{1}{3}$, $P(AB) = \frac{2}{9}$ and $P(A + B) = \frac{5}{9}$]

2. Three honest coins are tossed simultaneously and the resulting sequence of heads and tails is observed. Let A be the event in which "first two tosses match." Let B be the event representing "exactly two heads". Find out $P(A)$, $P(B)$, $P(B/A)$ and $P(A/B)$.

   [Ans. $P(A) = \frac{1}{2}$, $P(B) = \frac{3}{8}$, $P(B/A) = \frac{1}{4}$ and $P(A/B) = \frac{1}{3}$]

3. A long binary message contains 1428 binary 1's and 2668 binary 's. What is the probability of obtaining binary 1 in any received bit? [Ans. $P(A = 1) = 0.5486$]

4. Which of the following functions satisfy the properties of PDF and Why? Hence state whether the function is PDF or not. *(Kerala University, 1998)*

   (a) $f_X(x) = \frac{1}{\pi}\left(\frac{1}{1 + x^2}\right)$

   (b) $f_X(x) = |x|$ for $|x| < 1$
   $= 0$ otherwise

   (c) $f_X(x) = \frac{1}{6}(8 - x)$ for $4 \le x \le 10$
   $= 0$ otherwise

   [Hint: For all above function apply $\int_{-\infty}^{\infty} f_X(x)\,dx = 1$

   If this property holds, then it is PDF].

   [Ans. (a) This is valid PDF, (b) This is valid PDF, (c) This is not PDF]

5. The PDF of a random variable is given as

   $f_X(x) = Ke^{-bx}$ for $x \ge 0$
   $= 0$ for $x < 0$ and $K, b > 0$

   (i) Find the value of K in terms of b.
   (ii) Find m_x and σ^2_x. *(A.M.I.E., Examination, 1999)*

6. The PDF of a random variable 'X' is given as

   $f_X(x) = \frac{1}{A^2}(x + A)$ for $x < 0$

$$= \frac{1}{A^2}(-x + A) \quad \text{for } x \geq 0$$

Find out the mathematical expression for CDF.

[Ans. $f_X(x) = \frac{1}{2} + \frac{1}{A}x + \frac{1}{2A^2}x^2$ for $-A \geq x < 0$ $\frac{1}{2} + \frac{1}{A}x - \frac{1}{2A^2}x^2$ for $0 \leq x < A$]

7. A class of modulated signals are modelled by

$$x(t) = A\,X(t) \cos(\omega_c t + \theta)$$

Here $X(t)$ is the message signal and $A \cos(\omega_c t + \theta)$ is the carrier. The message signal $X(t)$ is a zero mean stationary random process. The carrier amplitude A and frequency ω_c are constants. The carrier phase θ is a random variable uniformly distributed in the interval $(-\pi, \pi)$ $X(t)$ and θ are independent find out mean and autocorrelation function of $x(t)$ and state whether the process is stationary or not.

[Ans. Mean of $x(t) = 0$ Auto correlation function $R_\tau(t_1, t_2) = \frac{A^2}{2} E[X(t_1) . X(t_2)] . \cos \omega_c(t_1 - t_2)$

$$= \frac{A^2}{2} R_\tau(t_1 - t_2) . \cos \omega_c(t_1 - t_2)]$$

Since mean and $R_\tau(t_1, t_2)$ does not depend on time 't' the process $x(t)$ is stationary in wide sense

□□□

EXCLUSIVE BY : www.dahiyastech.blogsopt.com

CHAPTER 6

Communcation System by Sanjay Sharma

Noise

Inside this Chapter

6.1. Introduction 362
6.2. Classification of Noise 362
6.3. External Noise 363
6.4. Atmospheric Noise 362
6.5. Extraterrestrial Noise 362
6.6. Industrial Noise 363
6.7. Internal Noise 363
6.8. Shot Noise 363
6.9. Partition Noise 365
6.10. Flicker Noise, or Low-Frequency Noise 365
6.11. Transit-time Noise or High Frequency Noise 365
6.12. Thermal Noise 366
6.13. Voltage and Current Models of a Noisy Resistor 366
6.14. Addition of Noise due to Several Sources in Series 368
6.15. Addition of Noise due to Several Sources in Parallel 368
6.16. Spectral Densities 370
6.17. Derivation of Thermal Noise Expressions Using Power Spectrum Density 372
6.18. Generalised Nyquist Theorem for Thermal Noise 373
6.19. Calculation of Thermal Noise for a Single Noise Source 375
6.20. Thermal Noise in RC Circuit 375
6.21. Calculation of Thermal Noise for Multiple Noise sources; Super position of Power Spectrum Density 376
6.22. White Noise 377
6.23. Equivalent Noise Bandwidth 378
6.24. Noise Equivalent Bandwidth for Low Pass Filter (LPF) 379
6.25. Nose in Reactive Circuits 381
6.26. Signal to Noise Ratio 383
6.27. Noise Figure 384
6.28. Noise Temperature 384
6.29. Calculation of Noise Figure 386
6.30. Noise Figure from Equivalent Noise Resistance 388
6.31. Experimental Determination of Noise Figure 390
6.32. Noise Figure Determination for Cascaded Stages of Amplifiers 392

6.1. Introduction

In electrical terms, noise may be defined as an unwanted form of energy which tend to interfere with the proper reception and reproduction of transmitted signals. For example, in receivers, several electrical disturbances produce noise and thus modifying the required signal in an unwanted form. In case of radio receivers, noise may produce hiss type sound in the output of loudspeaker. Similarly, in T.V. receivers, noise may produce 'snow' which becomes superimposed on the picture output. In addition to this, in pulse communications, noise may produce unwanted pulses or cancel the required pulses. In other words, we can say that noise may limit the performance of a communication system.

6.2. Classification of Noise

Although there are several ways of classifying the noise, but conveniently noise may be classified in two broad groups as under:

(*i*) External Noise
(*ii*) Internal Noise

6.3. External Noise

(U.P. Tech., Semester Exam., 2003-04)

External noise may be defined as that type of noise which is generated external to a communication system *i.e.* whose sources are external to the communication system.

External noise can not be analysed quantitatively. In addition to this, for a given geographical point or location, external noise cannot be controlled. Thus, to reduce the effect of external noise, the only way is to shift the communication system to other place or location which has comparatively smaller external noise. Thus due to this reason, the satellite earth stations are generally located in noise-free valleys.

External noise may be classified as under:

(*i*) Atmospheric Noise
(*ii*) Extraterrestrial Noise
(*iii*) Industrial Noise

6.4. Atmospheric Noise

Atmospheric noise, which is also called static, is produced by lightning discharges in thunderstorms and other natural electrical disturbances which occur in the atmosphere. These electric impulses are random in nature. Thus this energy is spread over the complete frequency spectrum used for radio communication. Atmospheric noise contains spurious radio signals which are distributed over a wide frequency range. In fact, these spurious radio waves constituting the noise, propagate over the earth in the manner as the required radio waves of same frequency. Thus, due to this reason, at any given receiving point, the receiving antenna picks up not only the required signal but also the static from all the thunderstorms.

It has been observed that the field strength of atmospheric noise varies approximately inversely with the frequency. This concludes that large atmospheric noise is produced in low and medium frequency bands where as very small noise is produced in the VHF and *UHF* bands.

Thus Atmospheric noise becomes less severe at frequencies above about 30 MHz.

6.5. Extraterrestrial Noise

There are several types of extraterrestrial noise or space noise depending upon their sources. Extraterrestrial noise may be divided into following two sub-groups as under:

(*i*) Solar Noise
(*ii*) Cosmic Noise

6.5.1. Solar Noise

Solar noise is the electrical noise eminating from the Sun.

Under steady conditions, there is a regular radiation of noise from the Sun. This radiation of noise from the sun is due to the fact that sun is a big body at an extremely high temperature and it radiates electric energy in the form of noise over a very wide frequency spectrum including also the frequency spectrum which is occupied by radio communication. However, the condition of the Sun varies and follows an eleven years cycle. Hence at the peak of this eleven years cycle, the electrical disturbances are caused. This form of noise from electrical disturbance is in addition to the quiet noise from the Sun. Although this additional noise is produced from a small portion of the Sun's surface, it can still be larger than the quiet solar noise. The solar cycle repeats this type of electrical disturbances nearly every eleven years.

6.5.2. Cosmic Noise

Distant stars can also be considered Suns. These distant stars have high temperatures and therefore radiate noise in the same manner as the Sun. The noise received from these distant stars is thermal noise and is distributed almost uniformly over the entire sky. The noise is also received from the centre of our own galaxy, for other distant galaxies and from other virtual point sources such as quasers and pulsers. This type of galactic noise is quite intense but since it comes from very distant sources, the angle subtended by the earth is quite small. Therefore, the strength of galactic noise received on the earth gets diminished.

As a conclusion, we can say that the space noise is quite, significant at frequencies in the 1.43 GHz. Also, in the frequency range of 20 to 12 MHz, the space noise becomes the strongest noise component next only to the industrial noise. But, below 20 MHz, the space noise does not penetrates much through the ionosphere to reach the earth.

6.6. Industrial Noise

The industrial noise or man-made noise is that type of noise which is produced by such sources as automobiles and aircraft ignition, electrical motors, switch gears and leakage from high voltage transmission lines and several other heavy electrical equipments. Such type of noise is produced by the arc discharge taking place during operation of all these machines or equipments. Industrial or man-made noise is quite intensive in industrial areas, dense populated urban areas. Industrial noise in such areas much stronger than all other sources of noise in the frequency range extending from about 1 MHz to 600 MHz .

Since industrial or man-made noise is highly variable and hence can be analyzed only statistically.

6.7. Internal Noise

(U.P. Tech., Semester, Exam., 2004-05)

Internal noise is that type of noise which is generated internally or within the communication system or receiver. Internal noise may be treated quantitatively and can also be reduced or minimized by proper system design. Since internal noise is randomly distributed over the entire frequency spectrum, the noise present in a given bandwidth B is the same at any frequency in the frequency spectrum. Hence, this random noise power is proportional to the bandwidth over which it is measured. Internal noise may be classified as under:

(*i*) Shot noise
(*ii*) Partition noise
(*iii*) Low frequency or Flicker noise
(*iv*) High-frequency or transit-time noise
(*v*) Thermal noise

6.8. Shot Noise

Shot noise arises in active devices due to random behaviour of charge carries. In electron tubes, shot noise is generated due to the random emission of electrons from cathodes, whereas in semiconductor devices shot noise is generated due to the random diffusion of minority carries or simply random generation and recombination of electron-hole pairs.

In fact, the current in electron devices (*i.e.* tubes or solid state device) flows in the form of discrete pulses, every time a charge carrier moves from one point to the other (*i.e.* cathode to plate). Hence, although the current appears to be continuous, it is still a discrete phenomena. Figure 6.1 shows this nature of current variation with time.

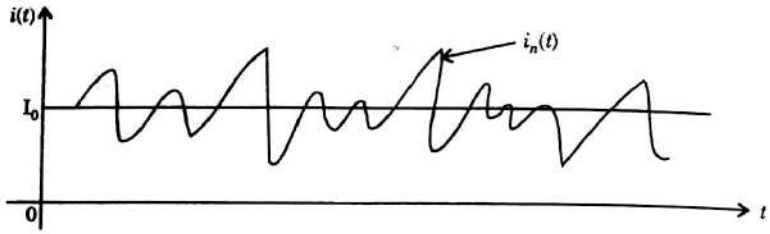


Fig. 6.1. *Current variation with time*

The current fluctuates about a mean value I_0. The current $i_n(t)$ wiggles around the mean value. This current is called as shot-noise current. The wiggling nature of this current is not visualized by normal instruments and normally it is assumed that the current is a constant equal to I_0. This wiggling nature of the current may be observed in a fast sweep oscilloscope.

Therefore, the total current $i(t)$ can be expressed as

$$i(t) = I_0 + i_n(t) \quad ...(6.1)$$

Here I_0 is the mean (constant) and $i_n(t)$ is the fluctuating (shot-noise) current.

6.8.1. Power Density Spectrum of Shot Noise in Diodes

We know that shot noise current is given by the expression

$$i(t) = I_0 + i_n(t) \quad ...(6.2)$$

The time-varying component $i_n(t)$ of the current $i(t)$ is always random in nature and hence it cannot be expressed as a function of time. This means that it is an indeterministic function. However, this indeterministic random function $i_n(t)$ may be specified by its power density spectrum.

In fact, the number of electrons, which contribute to the random current $i_n(t)$ is large.

Assuming that the electrons do not interact with each other during their movement or emission, this process (event) can be treated as statistically independent. Now, according to central limit theorem, such a process has a Gaussian distribution. Therefore, the shot-noise is Gaussian distributed with a zero mean value.

We can treat the total diode current as the sum of the current pulses and each pulse is formed by the transit of an electron from the cathode, to the anode. For all practical purposes, it may be observed that the power density spectrum of the statistically independent non-interacting random noise current $i_n(t)$ is expressed as

$$S_i(w) = qI_0 \quad ..(6.3)$$

Here q is the electronic charge and I_0 is the mean value of the current in amperes.

It may be noted that the power density spectrum in equation (6.3) is frequency independent. This type of frequency independence is only up to a frequency range determined by the transit time of an electron to reach from the anode to the cathode. Beyond this frequency range, the power density varies with frequency as shown in figure 6.2.

The transit time of an electron in a diode depends upon anode voltage V and may be expressed as

$$\tau = 3.36 \times \frac{d}{\sqrt{V}} \ \mu \text{ sec.} \quad ...(6.4)$$

Here d is the spacing between anode and the cathode.

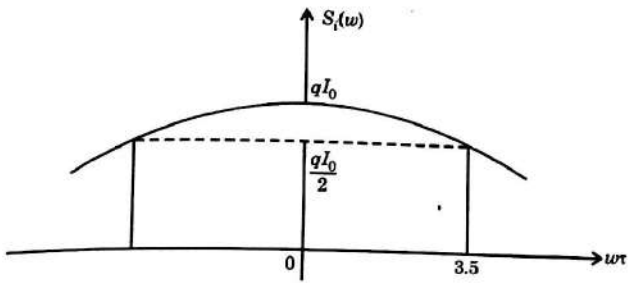


Fig. 6.2.

6.9. Partition Noise

Partition noise is generated in a circuit when a current has to divide between two or more paths. This means that the partition noise results from the random fluctuations in the division. Hence, it is expected that a diode must be less noisy than a transistor, all else being equal. Due to this reason, the inputs of microwave receivers are generally taken directly to the diode mixers. For partition noise, the spectrum is a flat spectrum. Also, the active three terminal components in which the control terminal draws less current are less noisy. In recent time, Metal-Semiconductor Field Effect Transistors have been developed which draws almost zero gate bias current. Due to this reason, these devices have low partition noise and therefore finds applications in low noise microwave amplification.

6.10. Flicker Noise, or Low-Frequency Noise

At low frequencies (below few kHz), a particular type of noise appears. The power spectral density of this noise increases as the frequency decreases. This noise is called as flicker noise or (1/f) noise. In case of vacuum tubes, the main causes of flicker noise are slow changes which take place in the oxide structure of oxide coated cathodes and migration of impurity ions.

In semiconductor devices, flicker noise is generated from the fluctuations in the carrier density and creates more problems in semiconductor amplifying device than in vacuum tubes at low frequencies.

Actually, the fluctuations in the carrier density generates fluctuations in the conductivity of the material. This produces a fluctuating voltage drop when a direct current flows. This fluctuating voltage drop is called as the flicker-noise voltage. The power density spectrum of the flicker noise is inversely proportional to frequency. Mathematically,

$$S(w) \propto \frac{1}{f} \quad ...(6.5)$$

Therefore, the flicker noise becomes significant at very low frequencies, generally below a few kHz.

6.11. Transit-Time Noise or High Frequency Noise

It is generally observed in semiconductor devices, when the transit-time of charge-carries crossing a junction is comparable with the time-period of the signal, some charge-carries diffuse back to the source or emitters. This process gives rise to an input admittance in which the conductance component increases with frequency. This conductance has a noise current source which is associated with it in parallel. Because this conductance increases with frequency, the power spectral density will also increase.

6.12. Thermal Noise

The thermal noise or white noise or Johnson noise is the random noise which is generated in a resistor or the resistive component of a complex impedance due to rapid and random motion of the molecules, atoms and electrons.

According to the kinetic theory of thermodynamics, the temperature of a particle denotes its internal kinetic energy. This means that the temperature of a body expresses the rms value of the velocity of motion of the particles in body. As per this kinetic theory, the kinetic energy of these particles becomes approximately zero (*i.e.* zero velocity) at absolute zero.

Therefore, the noise power produced in a resistor is proportional to its absolute temperature. Also the noise power is proportional to the bandwidth over which the noise is measured.

Therefore the expression for maximum noise power output of a resistor may be given as

$$P_n \propto T.B$$

or

$$P_n = k.T.B \qquad ...(6.6)$$

where k = Boltzmann's constant

$= 1.38 \times 10^{-23}$ Joule/deg. K

T = absolute temperature

B = Bandwidth of interest in Hz.

6.13. Voltage and Current Models of a Noisy Resistor

If we connect a d.c. voltmeter across any given resistor at room temperature *i.e.* 27°C (or 300 K), then no voltage is displayed by the dc voltmeter. However, if a very sensitive electronic voltmeter is used, then it displays some reading. This happens because of the fact that every resistor may be treated as a noise generator and so a sufficiently large voltage may develop across it.

However, the noise voltage so generated is always random and thus it has zero dc value but a definite rms value. Therefore only an a.c. meter across the resistor would register a value. This noise voltage is produced by the random movement of electrons within the resistor constituting a current. When measured over a long period of time, as many electrons arrive at one end of the resistor as at the other. But, at any particular instant of time, this principle does not hold good *i.e.* more electrons arrive at one end than at the other due to their random movement. This means that the rate of arrival of electrons at either end of the resistor will vary randomly and therefore the potential difference between the two ends will also vary randomly. This random voltage across the resistor can be both measured and calculated. The random noise voltage is always expressed in terms of rms value rather than instantaneous value.

Now, we can draw an equivalent circuit of a resistor as a noise voltage generator as shown in figure 6.3. This equivalent circuit is also called as **voltage model of a noisy resistor**.

From this equivalent circuit, we can compute the resistor's equivalent rms noise voltage V_n.

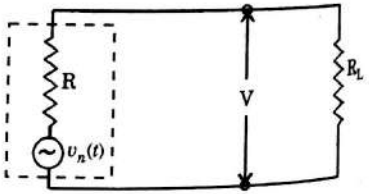


Fig. 6.3. *Voltage model of a noisy resistor*

Let us consider that a noiseless load resistor R_L is connected across the noise generator as shown in figure 6.3.

According to maximum power transfer theorem, for maximum transfer of power from noise voltage source V_n to load resistor R_L, we must have

$$R_L = R$$

Then the maximum noise power so transferred will be given as

$$P_n = \frac{V^2}{R_L}$$

But $R_L = R$

So, $P_n = \dfrac{V^2}{R}$

Applying voltage-divider method in figure 6.3, we get

$$V = \frac{V_n}{2}$$

So that $P_n = \dfrac{V^2}{R}$

$$P_n = \frac{(V_n/2)^2}{R} = \frac{V_n^2}{4R}$$

or $V_n^2 = 4RP_n$ ...(6.7)

But we know that $P_n = k\,T\,B$

Putting this value of P_n in equation (6.7), we get

$$V_n^2 = 4R(k\,T\,B) = 4R\,kTB$$

or $V_n = \sqrt{4RkTB}$ ...(6.8)

From equation (6.8), we may conclude that the square of the rms noise voltage associated with a resistor is proportional to the absolute temperature T of the resistor, value R of the resistor and the bandwidth B over which noise is being measured. At this point, it may be noted that the noise voltage is independent of the frequency at which it is measured.

This happens because of the fact that this noise is random and so on an average is evenly distributed over the frequency spectrum.

Norton's theorem may be used to find an equivalent current generator as shown in figure 6.4. This equivalent current generator is called as **current model of a noisy resistor.**

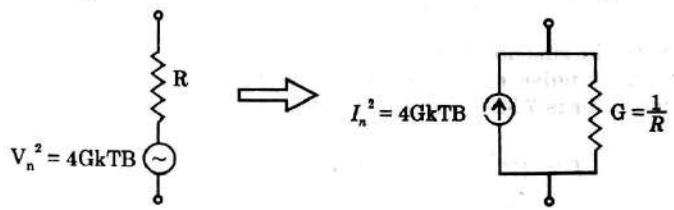


Fig. 6.4. *Voltage and current models of a noisy resistor.*

Using conductance $G = \dfrac{1}{R}$, the rms noise current I_n for current model of a noisy resistor will be expressed as

$$I_n^2 = 4\,GkTB \qquad ...(6.9)$$

Example 6.1. An amplifier operating over the frequency range from 18 to 20 MHz has a 10 KΩ input resistor. Calculate the rms noise voltage at the input to this amplifier if the ambient temperature is 27°C.

Solution: The rms noise voltage is given by the expression

$$V_n = \sqrt{4RkTB} \qquad ..(1)$$

Given that $R = 10\ \text{K}\Omega$

$T = 27 + 273 = 300°\ \text{K}$

Bandwidth $B = 20 - 18 = 2\ \text{MHz}$

$k = 1.38 \times 10^{-23}$ Joule/deg. K

Putting all these values in equation (*i*), we get

$$V_n = \sqrt{4\times10\times10^3\times1.38\times10^{-23}\times300\times2\times10^6}$$
$$= \sqrt{4\times1.38\times3\times2\times10^{-11}}$$
$$= 1.82\times10^{-5}\ \text{Volt} = 18.2\ \mu\text{V} \qquad \textbf{Ans.}$$

6.14. Addition of Noise Due to Several Sources in Series

Let us consider several thermal noise sources *i.e.*, resistors, R_1, R_2, R_3 etc. in series producing noise voltages V_{n1}, V_{n2}, V_{n3} etc. We know that the rms noise voltage produced by resistor R is given as

$$V_n = \sqrt{4kTBR}$$

Therefore,
$$V_{n1} = \sqrt{4kTBR_1}$$
$$V_{n2} = \sqrt{4kTBR_2}$$

and
$$V_{n3} = \sqrt{4kTBR_3} \quad \text{etc.}$$

Then the resultant noise voltage V_{nr} is given by the square root of the sum of the squares of individual *rms* noise voltages.

$$V_{nr} = \sqrt{V_{n_1}^2 + V_{n_2}^2 + V_{n_3}^2 + \ldots}$$

Putting all the values, we get

$$V_{nr} = \sqrt{4kTB(R_1 + R_2 + R_3 + \ldots)}$$

or
$$V_{nr} = \sqrt{4kTB.R_s} \qquad \ldots(6.10)$$

where R_s is the equivalent series resistance of the individual resistances and is given as

$$R_s = R_1 + R_2 + R_3 + \ldots$$

Example 6.2. Calculate the rms noise voltage at the input of a video amplifier using a device having 300 Ω equivalent noise resistance and 400 Ω input resistor. It is given that the bandwidth of the amplifier is 7 MHz and the ambient temperature is 27°C.

Solution: Given that

$$T = 273 + 27 = 300°\text{K}$$
$$R_s = R_1 + R_2 = 300 + 400$$
$$R_s = 700\ \Omega$$

Bandwidth $B = 7$ MHz

We know that the rms noise voltage is expressed as

$$V_{nr} = \sqrt{4kTBR_s}$$
$$V_{nr} = \sqrt{4\times1.38\times10^{-23}\times300\times7\times10^6\times700}$$
$$V_{nr} = 9\ \mu\text{V} \qquad \textbf{Ans.}$$

6.15. Addition of Noise due to Several Sources in Parallel

Let us consider several thermal noise sources *i.e.* resistors R_1, R_2, R_3 etc. in parallel producing rms noise voltages V_{n1}, V_{n2}, V_{n3} etc.

We know that the rms noise voltage produced by resistor R is given as

$$V_n = \sqrt{4kTBR}$$

Therefore, the resultant noise in this case will be

$$V_{nr} = \sqrt{4kTBR_{par}} \qquad \ldots(6.11)$$

where
$$\frac{1}{R_{par}} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \ldots$$

Example 6.3. Two resistors of 20 KΩ and 50 KΩ are at room temperature of 15°C or 290 K. For a given bandwidth of 100 kHz. Determine the thermal noise voltage generated by

(*i*) each resistor

(*ii*) the two resistors in parallel

Solution: This question is left as an exercise for the readers.

To solve this make use of equations (6.10) and (6.11).

6.15.1. Addition of Noise Due to Several Amplifiers in Series

Let us consider a two-stage amplifier. Let us assume that the gain of the first stage is A_1 and that of the second stage is A_2. The first stage has a total input-noise resistance R_1, the second R_2, and the output resistance is taken as R_3. The noise voltage at the output due to resistance R_3 is

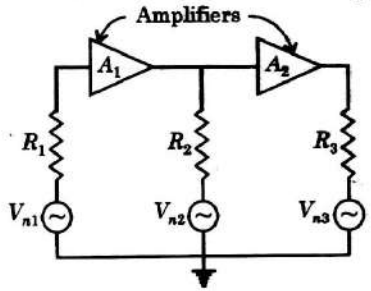


Fig. 6.5. *Calculation of noise of two amplifying stages in series.*

$$V_{n3} = \sqrt{4k\,TBR_3} \qquad \ldots(6.12)$$

The same noise voltage will be present at the output if there was no R_3, but instead R'_3 was present at the input of stage 2, such that

$$V'_{n_3} = \frac{V_{n_3}}{A_2} \qquad \ldots(6.13)$$

Substituting equation (6.12) in equation (6.13), we get

$$V'_{n_3} = \sqrt{\frac{4k\,TBR_3}{A_2^2}} = \sqrt{4kTB\frac{R_3}{A_2^2}} = \sqrt{4kTBR'_3} \qquad \ldots(6.14)$$

where
$$R'_3 = \frac{R_3}{A_2^2} \qquad \ldots(6.15)$$

(R'_3 is the resistance which if placed at the input of the second stage would produce the same noise voltage at the output as that produced by R_3) Equation (6.15) explains that when a noise resistance is "transferred" from the output of a stage to its input, it must be divided by the square of the voltage gain of that particular stage. Now the noise resistance actually present at the input of the second stage is R_2, so that the equivalent noise resistance at the input of the second stage, due to the second stage and the output resistance, will be

$$R'_{eq} = R_2 + R'_3 \qquad \ldots(6.16)$$

Substituting equation (6.15) in equation (6.16), we get

$$R'_{eq} = R_2 + \frac{R_3}{A_2^2} \qquad \ldots(6.17)$$

Similarly, a resistor R'_2 can be placed at the input of the first stage to replace R'_{eq}, both naturally producting the same noise voltage at the output.

$$R''_2 = \frac{R'_{eq}}{A_1^2} \qquad \ldots(6.18)$$

Substituting equation (6.17) in equation (6.18), we get

$$R''_2 = \frac{\left[R_2 + \frac{R_3}{A_2^2}\right]}{A_1^2} = \frac{R_2}{A_1^2} + \frac{R_3}{A_1^2 A_2^2} \quad ...(6.19)$$

The noise resistance actually present at the input of the first stage is R_1, so that the equivalent noise resistance of the whole cascaded amplifier at the input of the first stage, will be

$$R_{eq} = R_1 + R''_2 \quad ...(6.20)$$

Substituting equation (6.19) in equation (6.20), we get

$$R_{eq} = R_1 + R'_2 = R_1 + \frac{R_2}{A_1^2} + \frac{R_3}{A_1^2 A_2^2}$$

or

$$R_{eq} = R_1 + \frac{R_2}{A_1^2} + \frac{R_3}{A_1^2 A_2^2} \quad ...(6.21)$$

It may be noted that we can extend equation (6.21) to an n-stage caseaded amplifier.

Example 6.4. The first stage of a two-stage amplifier has a voltage gain of 10, a 600 Ω input resistor, a 1600 Ω equivalent noise resistance and a 27 KΩ output resistor. For the second stage, these values are 25, 81 KΩ, 10 KΩ and 1 MΩ, respectively. Find the equivalent input noise resistance of this two-stage amplifier. *(U.P. Technical University, 2001)*

Solution: Given for two-stage amplifier

The voltage gain for first stage is $A_1 = 10$

Voltage gain for second stage is $A_2 = 25$

Input resistor for first stage is $R_{i_1} = 600\ \Omega$

Equivalent noise resistance for first stage is $R_{eq_1} = 1600\ \Omega$

Output resistor for first stage is $R_{0_1} = 27\ \text{K}\Omega$

Input resistor for second stage is $R_{i_2} = 81\ \text{K}\Omega$

Equivalent noise resistance for second stage is $R_{eq_2} = 10\ \text{K}\Omega$

Output resistor for second stage is $R_{0_2} = 1\ \text{M}\Omega$

$$R_1 = R_{i_1} + R_{eq_1} = 600 + 1600 = 2200\ \Omega$$

$$R_2 = \frac{R_{0_1} R_{i_2}}{R_{0_1} + R_{i_2}} + R_{eq_2} = \frac{27 \times 81}{27 + 81} + 10 = 30.2\ \text{K}\Omega$$

$$R_3 = R_{02} = 1M\Omega$$

Equivalent input-noise resistance can be determined as

$$R_{eq} = R_1 + \frac{R_2}{A_1^2} + \frac{R_3}{A_1^2 A_2^2}$$

Substituting the values, we get

$$R_{eq} = 2200 + \frac{30.2 \times 10^3}{(10)^2} + \frac{1 \times 10^6}{(10)^2 (25)^2} = 2518\ \Omega \quad \textbf{Ans.}$$

6.16. Spectral Densities

We know that periodic waveforms are also known as deterministic waveforms since their values are known at all times. Periodic waveforms are also called as power waveforms. Another type of

power waveform encountered in communications engineering is the random waveform, for which the values may not be predicted. Noise waveforms are examples of random waveforms, like all information signals such as audio and video signals. These signals carry finite power, but unlike periodic waves they do not have a harmonic or line spectrum. Fourier methods applied to these random power signals result in a power spectral density function. This is a curve which shows the energy distribution as a continuous function of frequency. Such a typical curve has been illustrated in figure 6.6.

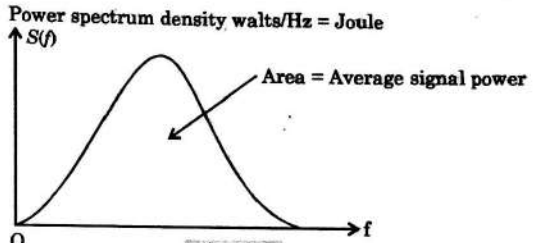


Fig. 6.6.

The units for power spectral density are watts per hertz, which are equivalent to joules. The area under the curve has units of joules × hertz, which is dimensionally equivalent to watts. But watts are the units of power. Therefore, the total area under the curve gives the average signal power. This means that the total average power is equal to the product of the power spectral density (watts/Hz) and bandwidth (Hz).

As discussed in last article, the thermal noise falls into the category of power signals and hence it has a power spectral density.

It may be shown that the power density spectrum of the current contributing the thermal noise is expressed as

$$S_i(w) = \frac{2kTG}{1 + \left(\frac{w}{\alpha}\right)^2} \quad ...(6.22)$$

Here, T is ambient temperature in kelvin.

G is the conductance of the resistor in mhos.

k is the Boltzmann constant

and α is the average number of collisions per second per electron.

Figure 6.6 shows the variation of power density spectrum with frequency.

From figure 6.6, it may be noted that the power density spectrum can be considered to be flat one for $\frac{w}{\alpha} \le 0.1$. This means that for this range of frequency, the power density spectrum $S_i(w)$ is almost constant and is expressed as

$$S_i(w) = 2kTG \quad ...(6.23)$$

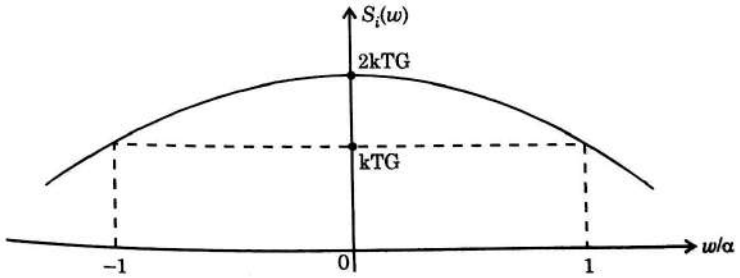


Fig. 6.7. *Power density spectrum of the resistor noise current.*

The value of α is generally of the order of 10^{14} and so the frequency range corresponding to $\frac{w}{\alpha} < 0.1$ is of the order of 10^{13} Hz. Thus, the frequency independent expression of $S_i(w)$ in equation

(6.23), is valid upto a frequency range of 10^{13} Hz. However, this range covers almost all the practical applications in communication systems.

Therefore, for all practical purposes, the power density spectrum $S_i(w)$ may be treated to be independent of frequency and is expressed by equation (6.23). This means that the contribution of any circuit is determined by the bandwidth of the circuit. Thermal noise can be considered to contain all frequency components in equal amount. Due to this reason, thermal noise is generally referred to as **white noise** (here white means the presence of all colours or frequencies).

Thermal noise is also known as **Johnson noise** after the scientist J.B. Johnson who presented a detailed investigation of it.

6.17. Derivation of Thermal Noise Expressions Using Power Spectrum Density

As discussed in article 6.13, noisy resistor may be represented as a voltage model and a current model. In voltage model of a noisy resistor, resistor R is connected with a thermal noise voltage source $v_n(t)$ in series.

In current model of a noisy resistor, the noise resistor R is represented by a conductance G in parallel with a thermal noise current source $i_n(t)$.

Figure 6.8 shows the voltage and current model of a noisy resistor R.

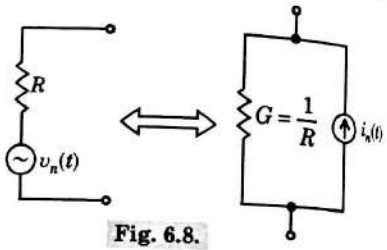


Fig. 6.8.

We know that the current $i_n(t)$ and voltage $v_n(t)$ are related as

$$v_n(t) = R.i_n(t) \qquad ...(6.24)$$

Now, since the power density spectrum is a function of the square of voltage or current, the relationship between the power density spectrum $S_i(w)$ of the current source $i_n(t)$ and the power density spectrum $S_v(w)$ of the voltage source $v_n(t)$ is expressed as

$$S_v(w) = R^2.S_i(w) \qquad ...(6.25)$$

But we know that the power density spectrum of the current contributing the thermal noise is given as

$$S_i(w) = 2kTG \qquad ...(6.26)$$

Therefore,

$$S_v(w) = R^2 S_i(w) = R^2.(2kTG)$$

Substituting $G = \frac{1}{R}$, we get

$$S_v(w) = R^2 \cdot \left(2kT.\frac{1}{R}\right) = 2kTR \qquad ...(6.27)$$

Thus, the power density spectrum $S_v(w)$ of a thermal noise voltage $v_n(t)$ is independent of frequency. Now, since the power density spectrum is defined as the power per unit bandwidth, therefore the noise power increases with increase in bandwidth.

We know that the average or total power P in terms of power density spectrum $S(w)$ is given as

$$P = \frac{1}{2\pi}\int_{-\infty}^{\infty} S(w)\,dw \qquad ...(6.28)$$

Hence, the noise power P_n in terms of power density spectrum is given as

$$P_n = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_v(w)\,dw \qquad ...(6.29)$$

But $\quad S_v(w) = 2kTR$

Therefore,

$$P_n = \frac{1}{2\pi}\int_{-\infty}^{\infty} 2kTR\,dw \qquad ...(6.30)$$

This integral would be infinite if it is integrated over an infinite bandwidth.

But for a finite bandwidth of $2\Delta f$, (from $-\Delta f$ to $+\Delta f$) the noise power is given as

$$P_n = S_v.(w).2\Delta f = 2kTR.2\Delta f$$

$$P_n = 4kTR(\Delta f) \qquad ...(6.31)$$

We know that the average power of a signal is the same as its mean square (ms) value

Thus, $\quad P_n = \overline{v_n^2} = 4kTR(\Delta f)$

Let the rms value of noise voltage v_n be V_n.

Then, the rms value is given as

$$V_n = \sqrt{\overline{v_n^2}} = \sqrt{4kTR(\Delta f)} \qquad ...(6.32)$$

6.18. Generalised Nyquist Theorem for Thermal Noise

With the help of generalised Nyquist theorem, the concept of power density spectrum of a single resistor is extended to a generalized circuit. Let us consider a generalized linear bilateral passive network (*i.e.* two port network) consisting of resistance, R capacitance C and inductance L as shown in figure 6.9.

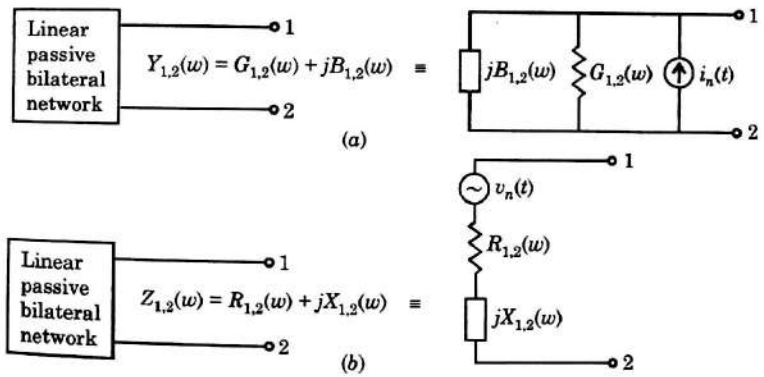


Fig. 6.9.

The Norton's equivalent circuit is representing the given network in terms of conductance $G_{1,2}(w)$ and susceptance $B_{1,2}(w)$ forming an equivalent admittance of the network expressed as

$$Y_{1,2}(w) = G_{1,2}(w) + jB_{1,2}(w) \qquad ...(6.33)$$

Here, the source of the thermal noise is conductance $G_{1,2}(w)$ and so the given network may be represented by a noiseless admittance $Y_{1,2}(w)$ in parallel with a noise current source with power density spectrum expressed as

$$S_i(w) = 2kT\,G_{1,2}(w) \qquad ...(6.34)$$

Similarly, the Thevenin's equivalent circuit is representing the given network in terms of resistance $R_{1,2}(w)$ and reactance $X_{1,2}(w)$ forming an equivalent impedance of the network expressed as

$$Z_{1,2}(w) = R_{1,2}(w) + jX_{1,2}(w) \quad ...(6.35)$$

Here, the source of the thermal noise is resistance $R_{1,2}(w)$ and so the given network may be represented by a noiseless impedance $Z_{1,2}(w)$ in series with a noise voltage source with the power density spectrum expressed as

$$S_v(w) = 2kT.R_{1,2}(w) \quad ...(6.36)$$

At this stage, following two points may be noted:

(i) In general, $R_{1,2}(w)$ and $G_{1,2}(w)$ are frequency dependent and therefore the power density spectrum is also frequency-dependent, *i.e.*, a function of frequency, unlike the white noise.

(ii) The major limitation of the theorem is that it is applicable to a linear bilateral passive circuit elements in which all of the noise sources are due to resistors only. This means that this theorem cannot be applied to a circuit containing active elements like tubes and transistors.

Example 6.5. Figure 6.10 shows a passive network. Determine the power density spectrum of the thermal noise voltage across terminals 1, 2 of the given network.

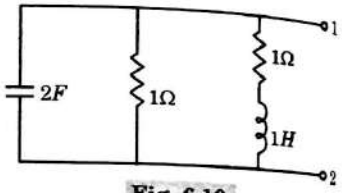


Fig. 6.10.

Solution: The Thevenin's equivalent circuit for the given network may be drawn as shown in figure 6.11.

The admittance of the network across terminals 1, 2 is expressed as

$$Y_{1,2}(w) = 1 + j.2w + \frac{1}{1+jw} = \frac{2-2w^2+3jw}{1+jw}$$

Thus, the impedance across the terminals 1, 2 is expressed as

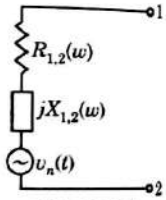


Fig. 6.11.

$$Z_{1,2}(w) = \frac{1}{Y_{1,2}(w)}$$

or
$$Z_{1,2}(w) = \frac{1+jw}{2-2w^2+3jw}$$

To separate the real and imaginary part, multiplying the above expression by conjugates, we get

$$Z_{1,2}(w) = \frac{1+jw}{(2-2w^2+3jw)} \times \frac{(2-2w^2-3jw)}{(2-2w^2-3jw)}$$

or
$$Z_{1,2}(w) = \frac{(2+w^2) - jw(1+2w^2)}{4+4w^4+w^2}$$

or
$$Z_{1,2}(w) = \frac{2+w^2}{4+4w^4+w^2} - j\frac{w(1+2w^2)}{4+4w^4+w^2}$$

or
$$Z_{1,2}(w) = R_{1,2}(w) + jX_{1,2}(w)$$

Hence, real part $R_{1,2}(w)$ is

$$R_{1,2}(w) = \frac{2+w^2}{4+w^2+4w^4}$$

This resistance $R_{1,2}(w)$ is the source of noise as shown in figure 6.10. Therefore, the noise power density spectrum of $v_n(t)$ is expressed as

$$S_v(w) = 2kTR_{1,2}(w) = \frac{2kT.(2+w^2)}{4+w^2+4w^4} \quad \textbf{Ans.}$$

Download From www.dahiyastech.blogspot.com

6.19. Calculation of Thermal Noise for a Single Noise Source

Let us consider a linear system or circuit consisting of noiseless elements. Let this system be excited by a random noise input voltage $v_{ni}(t)$. We have to determine the *rms* value of the output noise voltage $v_{no}(t)$.

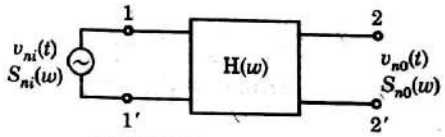


Fig. 6.12. *A linear system.*

Let $H(w)$ be the transfer function of the system.

Let us consider that $S_{ni}(w)$ and $S_{no}(w)$ be the power spectrum densities of the noise voltages $v_{ni}(t)$ and $v_{n0}(t)$ respectively.

We know that for a linear system, the relationship between input and output power spectrum densities is given as under:

$$S_{no}(w) = |H(w)|^2 S_{ni}(w) \quad ...(6.37)$$

Now, the output noise power P_0 may be calculated by integrating output power spectrum density $S_{no}(w)$ over the bandwidth under consideration.

$$P_o = \text{m.s. value of } v_{no}(t) = \overline{v^2_{no}}(t) \quad ...(6.38)$$

or
$$P_o = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_{n0}(w)\,dw$$

Substituting the values of $S_{no}(w)$ in above equation, we get

or
$$P_0 = \overline{v^2_{no}} = \frac{1}{2\pi}\int_{-\infty}^{\infty} |H(w)|^2 \cdot S_{ni}(w)\,dw \quad ...(6.39)$$

We know that the rms output noise voltage may be obtained by simply taking square root of mean square (ms) value of output noise voltage, *i.e.*

$$V_{rms} = \sqrt{\overline{v^2_{no}}} = \left[\frac{1}{2\pi}\int_{-\infty}^{\infty} |H(w)|^2 S_{ni}(w)\,dw\right]^{1/2} \quad ...(6.40)$$

This is the required expression.

6.20. Thermal Noise in RC Circuit

(U.P.S. C.I.E.S. Engg. Examination-1998)

Let us consider an RC circuit as shown in figure 6.13. Let us obtain an equivalent circuit by replacing the thermal resistor with its equivalent thevenin model as shown in figure 6.13. The equivalent circuit so obtained consists of a noiseless resistor and a series noise voltage source of $v_{ni}(t)$ having a power spectrum density $S_{ni}(w)$.

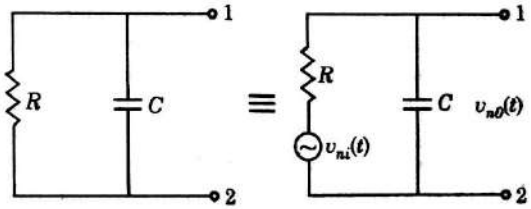


Fig. 6.13.

We know that the power spectrum density $S_{ni}(w)$ is given as

$$S_{ni}(w) = 2kTR \quad ...(6.41)$$

We can find a transfer function $H(w)$ for R-C circuit. This transfer function $H(w)$ relates the output voltage $v_{no}(t)$ with the input noise voltage $v_{ni}(t)$.

Thus,
$$H(w) = \frac{\frac{1}{jwC}}{R + \frac{1}{jwC}} = \frac{1}{1+jwRC}$$

or $$H(w) = \frac{1 - jwRC}{1 + w^2R^2C^2} = \frac{1}{1 + w^2R^2C^2} - j\frac{wRC}{1 + w^2R^2C^2}$$

or $$|H(w)| = \sqrt{\frac{1}{(1 + w^2R^2C^2)^2} + \frac{w^2R^2C^2}{1 + w^2R^2C^2}}$$

or $$|H(w)| = \frac{\sqrt{1 + w^2R^2C^2}}{1 + w^2R^2C^2} = \frac{1}{\sqrt{1 + w^2R^2C^2}} \quad ...(6.42)$$

The power density spectrum of the output noise voltage $v_{no}(t)$ is given as

$$S_{no}(w) = |H(w)|^2.S_{ni}(w) \quad ...(6.43)$$

Substituting the value of $|H(w)|$ and $S_{ni}(w)$ in above equation, we get

$$S_{no}(w) = \frac{1}{1 + w^2R^2C^2} \cdot 2kTR = \frac{2kTR}{1 + w^2R^2C^2} \quad ...(6.44)$$

The average power or mean square (ms) value of the output noise voltage $v_{no}(t)$ is given as

$$P_0 = \overline{v^2_{no}} = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_{no}(w)\,dw = \frac{1}{\pi}\int_0^{\infty} S_{no}(w)\,dw$$

Substituting the value of $S_{no}(w)$ in above equation, we get

$$\overline{v^2_{no}} = \frac{1}{\pi}\int_0^{\infty} \frac{2kTR}{1 + w^2R^2C^2}\,dw = \frac{2kTR}{\pi}\int_0^{\infty} \frac{1}{1 + R^2w^2C^2}\,dw$$

or $$\overline{v^2_{no}} = \frac{2kTR}{\pi RC}\left[\tan^{-1} RwC\right]_0^{\infty} = \frac{2kTR}{\pi RC} \cdot \frac{\pi}{2} = \frac{kT}{C} \quad ...(6.45)$$

Thus, the output rms noise voltage is expressed as

$$V_{rms} = \sqrt{\overline{v^2_{no}}} = \sqrt{\frac{kT}{C}} \quad ...(6.46)$$

6.21. Calculation of Thermal Noise for Multiple Noise Sources: Superposition of Power Spectrum Density

In any electrical circuit, there can be a large number of noise sources due to resistors and electronic elements. All these noise sources are independent sources of random signals.

This theorem states that for any system involving multiple independent sources of noise, the mean square (ms) value of response or output is equal to the sum of the mean square (ms) values of the responses calculated by individual sources at a time. In the same manner, the power spectrum density of the output or response may also be calculated by using the superposition theorem.

Let us consider a linear system with two independent input sources $x_1(t)$ and $x_2(t)$ at two different terminals generating an output or response $y(t)$ as shown in figure 6.14. Let us assume that $x_1(t)$ alone produces a response $y_1(t)$ and $x_2(t)$ alone produces a response $y_2(t)$ to this linear system.

Now since $x_1(t)$ and $x_2(t)$ are independent sources, then according to the superposition theorem,

$$\overline{y^2(t)} = \overline{y_1^2(t)} + \overline{y_2^2(t)} \quad ...(6.47)$$

and $$S_y(w) = S_{y1}(w) + S_{y2}(w) \quad ...(6.48)$$

We can extend this result for more than two sources as

$$\overline{y^2(t)} = \overline{y_1^2(t)} + \overline{y_2^2(t)} + \overline{y_3^2(t)} + ... + \overline{y_n^2(t)} \quad ...(6.49)$$

and $$S_y(w) = S_{y1}(w) + S_{y2}(w) + S_{y3}(w) + ... + S_{yn}(w) \quad ...(6.50)$$

6.22. White Noise

Noise in an idealized form is known as **white noise**. This mens that in a communication system, the noise analysis is based on an idealized form of noise, *i.e.*, white noise.

As we know that white light consists of all colour frequencies, in the same manner, white noise contains all frequencies in equal amount. Actually, the power density spectrum of a white noise is independent of frequency. This means that white noise consists of all the frequency components in equal amount. If the probability of occurrence of a white noise is specified by a Gaussian distribution function, it is called as white Gaussian noise. Since the power density spectrum of thermal and shot noise is independent of the operating frequencies, therefore, shot noise and thermal noise can be treated as white Gaussian noise for all practical purposes.

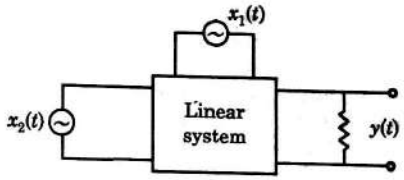


Fig. 6.14. *A linear system with two independent random noise sources*

The power spectrum density of white noise is expressed as

$$S_w(w) = \frac{N_0}{2} \quad ...(6.51)$$

Here, the factor 1/2 has been included to show that half of the power is associated with the positive frequencies and remaining half with the negative frequencies. This has been shown in figure 6.15. The power spectrum density of white noise shown in figure 6.15 reveals the fact that it has no dc power, *i.e.*, the mean or average value of white noise is zero. In addition to this, the autocorrelation function of the power spectrum density of white noise may be obtained by simply taking the inverse Fourier transform of both sides of expression

$$S_w(w) = \frac{N_0}{2} \quad ...(6.52)$$

Thus, the autocorrelation function for the white noise may be expressed as

$$R_w(\tau) = \frac{N_0}{2}\delta(\tau) \quad ...(6.53)$$

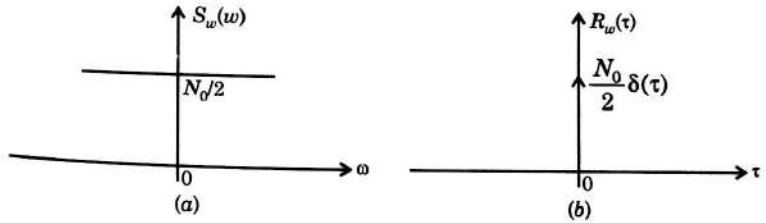


Fig. 6.15.

Equation (6.53) reveals the fact that the autocorrelation function of the white noise contains a delta function scaled by a factor $N_0/2$ and situated at $\tau = 0$ as shown in figure 6.15.

Also, $$R_w(\tau) = 0 \quad \text{for} \quad \tau \neq 0 \quad ...(6.54)$$

Strictly speaking, white noise has infinite average power and thus is not physically realizable. However, the concept of white noise provides convenient mathematical properties and this is a useful tool for system analysis.

Hence, we observe that the filtered white noise has a finite average power which is proportional to the bandwidth in both the cases. This is true for all kinds of low-pass filters.

For convenience, to generalize the above observation for all kinds of low-pass filters (LPFs), we define the noise equivalent bandwidth.

Now let us consider an arbitrary low-pass filter (LPF). The transfer function of this low-pass filter (LPF) is $H(f)$ at the input of which a white noise source is connected. This white noise has zero mean and $\frac{N_o}{2}$ power spectrum density.

The power spectrum density (psd) of the output noise is given as

$$S_{no}(f) = |H(f)|^2 \frac{N_o}{2} \qquad ...(6.64)$$

here $|H(f)|^2$ is the squared magnitude of the transfer function of the low-pass filter (LPF). Therefore, the resulting average output noise power is given as

$$P_{no}(f) = \int_{-\infty}^{\infty} S_{no}(f)\, df \qquad ...(6.65)$$

Substituting equation (6.64) in equation (6.65), we obtain

$$P_{no}(f) = \int_{-\infty}^{\infty} |H(f)|^2 \frac{N_o}{2}\, df$$

$$= \frac{N_o}{2} \int_{-\infty}^{\infty} |H(f)|^2\, df$$

or

$$P_{no}(f) = \frac{No}{2} \times 2\int_{0}^{\infty} |H(f)|^2\, df = No\int_{0}^{\infty} |H(f)|^2\, df \qquad ...(6.66)$$

The same white noise source is connected to the input of an ideal low-pass filter (LPF). This ideal low-pass filter (LPF) has zero frequency response $H(0)$ and bandwidth B_N.

The average output noise power is given as

$$P_{no}(f) = N_o.B_N |H(0)|^2 \qquad ...(6.67)$$

From above equation, it may be observed that the filtered noise power P_{no} is a finite value. It is also proportional to the bandwidth B_N. This bandwidth B_N is called **noise-equivalent bandwidth for a low-pass filter (LPF).**

Using equations (6.66) and (6.67), we obtain

$$B_N = \frac{\int_0^{\infty} |H(f)|^2\, df}{|H(0)|^2} \qquad ...(6.68)$$

Hence, the noise-equivalent bandwidth may be obtained by replacing the arbitrary low-pass filter of transfer function $H(f)$ by an equivalent ideal low-pass filter having zero-frequency response $H(0)$ and bandwidth B_N.

It has been shown in figure 6.18.

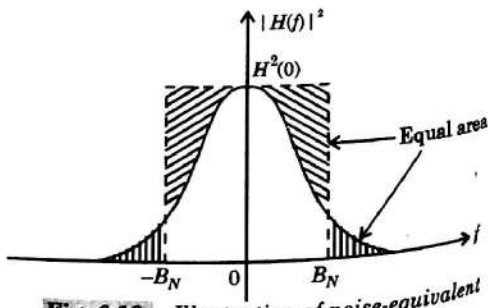


Fig. 6.18. *Illustration of noise-equivalent bandwidth of low-pass filter (LPF).*

Example 6.6. Determine the noise equivalent bandwidth of RC low-pass filter (LPF) whose transfer function or frequency response is given by

$$H(f) = \frac{1}{1 + j2\pi fRC}$$

(Kerala University-1999)

Solution: Given that the frequency response

$$H(f) = \frac{1}{1 + j2\pi fRC} \qquad ...(i)$$

or

$$H(f) = \frac{1}{1 + j\left(\frac{f}{B}\right)} \qquad ...(ii)$$

where

$$B = \frac{1}{2\pi RC} \qquad ...(iii)$$

Therefore

$$|H(f)|^2 = \frac{1}{1 = (f/B)^2} \qquad ...(iv)$$

For the centre frequency $f = 0$ of the filter, we have

$$|H(0)|^2 = \frac{1}{1 + (0/B)^2} = 1 \qquad ...(v)$$

We know that the noise-equivalent bandwidth for a low-pass filter is expressed as

$$B_N = \frac{\int_0^{\infty} |H(f)|^2\, df}{|H(0)|^2} \qquad ...(vi)$$

Substituting equations (*iv*) and (*v*) in equation (*vi*), we get

$$B_N = \frac{\int_0^{\infty} \frac{1}{1 + (f/B)^2}\, df}{1} = \frac{\pi}{2} \cdot B \qquad ...(vii)$$

Now substituting equation (*iii*) in equation (*vii*), we get

$$B_N = \frac{\pi}{2} \cdot \frac{1}{2\pi RC} = \frac{1}{4RC} \qquad ...(viii)$$

Hence, the noise-equivalent bandwidth decreases with increase in R.

The average noise output power of a low-pass filter (LPF) is given as

$$P_n = N_o B_N |H(0)|^2$$

Now, substituting equations (*v*) and (*viii*) in lost equation, we get

$$P_n = N_o \cdot \frac{1}{4RC} \times 1 = \frac{N_o}{4RC}$$

But for thermal noise source

$$N_o = 4\, kTR$$

Therefore

$$P_n = \frac{N_o}{4RC} = \frac{4kTR}{4RC} = \frac{kT}{C}$$

6.25. Noise in Reactive Circuits

Figure 6.19 shows a parallel resonant circuit. In this circuit, the inductor L is assumed to be an ideal one. The resistance R is the resistance associated with the inductor or coil. Thus resistance R is the only source of noise voltage in the circuit. Let v_{ni} be the noise voltage due to resistor R.

The mean square value of noise voltage is given as

$$\overline{v^2_{ni}} = 4kRT(\Delta f) \quad ...(6.69)$$

Figure 6.20 shows the equivalent circuit in which a noise voltage source v_{ni} is connected with the resistor R. Now, our purpose is to find the noise voltage v_{no} developed across the output terminals.

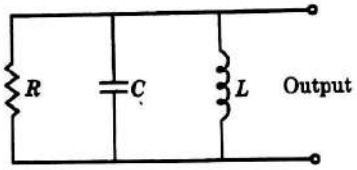


Fig. 6.19. *A parallel resonant circuit*

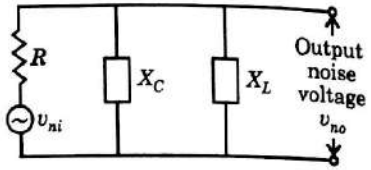


Fig. 6.20. *Equivalent circuit*

Let V_{ni} be the rms value of noise voltage v_{ni} and Z be the total impedance of the circuit.
Let I_{rms} be the rms value of the noise current generated due to noise voltage v_{ni}.

Thus, $$I_{rms} = \frac{V_{ni}}{Z}$$

But at resonance

$$Z = R$$

Hence, $$I_{rms} = \frac{V_{ni}}{R}$$

Let V_{no} be the rms value of the output noise voltage v_{no}. Thus, the rms value of output noise voltage is given by

$$V_{no} = I_{rms}.X_C$$

where X_C is the reactance of the capacitor C.

or $$V_{no} = \frac{V_{ni}}{R}.X_C$$

But $$X_C = Q.R \qquad \left[\because Q = \frac{X_C}{R}\right]$$

Therefore, $$V_{no} = \frac{V_{ni}}{R}.Q.R = Q.V_{ni} \quad ...(6.70)$$

Here V_{no} and V_{ni} are the rms values. Since we know that the rms value of a signal is equal to the square root of its mean square (ms) value, therefore,

$$\left(\overline{v^2_{no}}\right)^{1/2} = Q\left(\overline{v^2_{ni}}\right)^{1/2}$$

Squaring both sides, we get

$$\overline{v^2_{no}} = Q^2.\overline{v^2_{ni}}$$

Putting the value of v_{ni} in above equation from equation (6.69), we get

$$\overline{v^2_{no}} = Q^2.4kTR(\Delta f)$$

Substituting

$$Q^2R = R_p$$

where R_p is the equivalent parallel resistance, we get

$$\overline{v^2_{no}} = 4kTR_p(\Delta f) \quad ...(6.71)$$

Example 6.7. A parallel tuned circuit is made to resonate at a frequency of 100 MHz. The parallel tuned circuit uses a coil having quality factor Q of 10 and a capacitance of 10 pF. The temperature of the circuit is maintained at 17°C. Determine the output voltage across the circuit measured by a wideband voltmeter.

Solution: We know that the quality factor Q of the coil is given by

$$Q = \frac{f_r}{\Delta f}$$

where f_r is the resonance frequency and Δf is the bandwidth of the tuned circuit.

Thus, $$\Delta f = \frac{f_r}{Q} = \frac{100}{10} = 10 \text{ MHz}$$

The quality factor Q is also given by the following expression as

$$Q = \frac{1}{wCR}$$

or $$R = \frac{1}{QwC} = \frac{1}{10 \times w \times 10 \times 10^{-12}}$$

or $$R = \frac{1}{10 \times 2\pi \times 100 \times 10^6 \times 10 \times 10^{-12}} \cong 16 \text{ ohms}$$

The output voltage (rms value) is given as

$$V_{no} = \left(\overline{v^2_{no}}\right)^{1/2}$$

But mean square (ms) value $\overline{v^2_{no}}$ is given by

$$\overline{v^2_{no}} = 4kT(\Delta f)\, Q^2R$$

Therefore,

$$V_{no} = \sqrt{4kT(\Delta f)\,Q^2R} = \sqrt{4 \times 1.38 \times 10^{-23} \times (273+17) \times 10 \times 10^6 \times 10^2 \times 16}$$

Solving, we get

$$V_{no} = 160 \times 10^{-7} = 16\ \mu\text{V} \quad \textbf{Ans.}$$

6.26. Signal to Noise Ratio

(U.P. Tech., Semester, Exam., 2004-05)

The ratio of signal power to the associated noise power is defined as signal to noise ratio. In other words, signal to noise ratio is defined as the ratio of signal power to the noise power at the same point in the system. Signal to noise ratio is denoted by S/N.

Thus, if a signal voltage $v_s(t)$ is associated with a noise voltage source $v_n(t)$, then the ratio of signal power to the noise power will be

$$\frac{S}{N} = \frac{\overline{v_s^2}}{\overline{v_n^2}} \quad ...(6.72)$$

Because the power spectrum density is power per unit bandwidth, hence the above expression may be expressed as

$$\frac{S}{N} = \frac{S_s(w)}{S_n(w)}$$

or $$\frac{S}{N} = \frac{\text{Power spectrum density of signal voltage}}{\text{Power spectrum density of noise voltage}} \quad ...(6.73)$$

The S/N ratio is always a power ratio unless stated otherwise. In some cases, the ratio of a signal voltage (rms) to a noise voltage (rms) is also specified. However, in this case, it is specially stated that S/N voltage ratio is being considering.

The signal to noise ratio (S/N ratio) at the input and output of a noise free two-port network is the same since both signal and noise are treated identically. But, practical two-port networks introduce some additional noise. Thus for a practical two port network, the S/N ratio deteriorates as we move from input side to the output side. Infact, the comparison of S/N ratio at the input and the output of any two port network gives the noise indication of that network.

6.27. Noise Figure

(U.P. Tech., Semester, Exam., 2003-04)

It has been observed that for comparison of receivers or amplifiers working at various impedance levels the use of the equivalent noise resistance is quite misleading. However, instead of equivalent noise resistance, one another quantity known as **noise figure** is defined and used. Sometimes, this is also called **noise-factor**.

The noise figure F is defined as the ratio of the signal-to-noise power ratio supplied to the input terminals of a receiver or amplifier to the signal-to-noise power ratio supplied to the output or load resistor.

Hence, mathematically,

Noise figure $$F = \frac{\text{Input } SNR}{\text{Output } SNR} = \frac{(SNR)_i}{(SNR)_o}$$

At this point, it may be noted that the noise figure F is one for an ideal receiver which introduces no amount of noise of its own, due to which the SNR does not deteriorate as we pass from the input side to the output side of the receiver. The noise figure for practical receiver can be expressed as an actual ratio or in decibels (dBs).

Example 6.8. Determine receiver's noise figure in decibels and its equivalent noise temperature. A receiver is connected to an antenna whose resistance is 50 Ω. The equivalent noise resistance of this receiver is 30 ohms. *(U.P. Technical University, 2002)*

Solution: Given that: antenna resistance $R_a = 50\ \Omega$

Equivalent noise resistance of receiver $R_{eq} = 30\ \Omega$.

We know that the relationship between F and R_{eq} is expressed as

$$F = 1 + \frac{R_{eq}}{R_a} = 1 + \frac{30}{50} = 1 + 0.6 = 1.6$$

F in dB is given is

$$F_{dB} = 10 \log_{10}(F) = 10 \log_{10}(1.6) = 2.04\ dB$$

Also the relationship between F and T_{eq} is expressed as

$$T_{eq} = T_o(F - 1) = 290(1.6 - 1) = 174°\ \text{K} \qquad \textbf{Ans.}$$

6.28. Noise Temperature

(U.P. Tech., Semester, Exam., 2003-04)

According to thermodynamic considerations the power spectrum of thermal noise is constant over a very wide frequency range. Noise having such type of spectrum consists of all frequency components in equal proportion and is designated as white noise.

As we know that the available thermal noise power is expressed as ...(6.74)

$$P_n = k.T.B \text{ watts}$$

As obvious, this noise depends only upon the value of temperature and bandwidth. Therefore, a thermal resistor delivers a maximum of kT watts per unit bandwidth regardless of the value of R.

Thus, temperature is the fundamental parameter of thermal noise. However, there are other white noise sources which are non-thermal in the sense that noise power is not related to a physical temperature, still we can speak of the noise temperature T_n of any white-noise source, either thermal or non-thermal, by defining as under:

$$T_n = \frac{P_n}{kB} \quad ...(6.75)$$

Here P_n is the maximum noise power of the sources delivered over the bandwidth B.

For thermal sources, T_n is a physical temperature and for non-thermal sources, it is a measure of the available noise power. This implies that a non-thermal white-noise source having equivalent internal resistance R_n and noise temperature T_n may be replaced by a thermal resistor $R = R_n$ at physical temperature $T = T_n$ without altering the terminal properties.

Also it may be shown that the equivalent noise temperature T_n may be expressed in terms of noise figure as under:

$$T_n = T_o(F - 1) \quad ...(6.76)$$

where $T_o = 17°\text{C}$ *i.e.*, 290° K.

Thus, once the value of noise-figure is known, the equivalent-noise temperature can be calculated with the help of above equation.

At this stages, it may be noted that the concept of noise-figure, although frequently used, is not always the most convenient measure of noise particularly in dealing with UHF and microwave low-noise antennas, receivers or devices. Noise temperature is employed extensively for antennas and low-noise microwave amplifiers. Noise temperature is additive like noise power.

It may be observed as under:

We, know that total noise power is given as

$$P_t = k\,T_t.\,B \quad ...(6.77)$$

or $$P_t = P_1 + P_2 = kT_1B + kT_2B$$

or $$T_t = T_1 + T_2 \quad ...(6.77a)$$

where P_1 and P_2 are the two individual noise powers (*i.e.*, received by the antenna and generated by antenna respectively) and P_t is the sum of the two powers P_1 and P_2, T_1 and T_2 are the individual noise temperatures, T_t = the total noise temperature.

One more advantage of the use of noise temperature for low-noise levels is that it exhibits a greater variation for any given noise level. Hence, noise level changes are easier to grasp by using noise temperature concept.

It may also be noted that the concept of equivalent noise resistance is quite fictitious even then it is often employed due to its convenience. Equivalent noise temperature T_{eq} may also be utilized if it proves convenient. In defining the equivalent noise temperature of a receiver or amplifier we have assumed that

$$R'_{eq} = R_a. \quad ...(6.78)$$

Now, if this is to lead to the correct value of the noise output power then clearly R'_{eq} must be at a temperature other than the standard one at which all the components are assumed. Therefore, we can equate noise figure and equivalent noise temperature *i.e.*,

$$F = 1 + \frac{R'_{eq}}{R_a} \quad ...(6.79)$$

or $$F = 1 + \frac{kT_{eq}\,B\,R'_{eq}}{kT_o\,B.R_a} = 1 + \frac{T_{eq}}{T_o} \quad ...(6.80)$$

where $R'_{eq} = R_a$

$T_o = 17°C = 17 + 273 = 290°$ K

T_{eq} = Equivalent noise temperature of the receiver or amplifier whose noise figure is F.

It may be noted that here F is a ratio of powers and is not expressed in decibels. Equivalent noise temperature T_{eq} may be influenced by the actual ambient temperature of the receiver or amplifier. It is also to be noted that the equivalent noise temperature is just a convenient fiction. If all the noise of the receiver were generated by R_a, its temperature would have to be T_{eq}.

Lastly using equation (6.80), we have

$$F = 1 + \frac{T_{eq}}{T_o}$$

or $$T_o F = T_o + T_{eq}$$

or $$T_{eq} = T_o(F - 1) \quad ...(6.81)$$

Thus, once noise figure F is known, the equivalent noise temperature T_{eq} may be determined from equation (6.81).

Graphs of noise temperature of various sources verses frequency and the temperature verses frequency are also available.

6.29. Calculation of Noise Figure

To calculate the noise figure, let us consider a network (*i.e.*, receiver) This network has an input impedance R_t, an output impedance R_L and an overall voltage gain A. It is fed from a source *i.e.*, antenna of internal resistance R_a. The internal resistance R_a may or may not be equal to R_t as the circumstances warrant. Figure 6.21 shows a block diagram of such a four-terminals network (with the source feeding it).

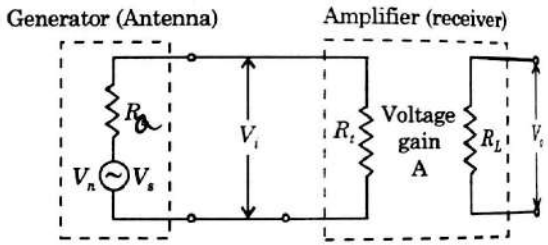


Fig. 6.21. *Block diagram for noise figure calculation.*

The calculation procedure can be divided into a number of steps. These steps are as under:

Step *I* : Determination of signal input power P_{si}.

Step *II* : Determination of noise input power P_{ni}.

Step *III* : Calculation of input signal-to-noise ratio $(SNR)_i$ with the help of P_{si} and P_{ni}.

Step *IV* : Determination of signal output power P_{so}.

Step *V* : Determination of noise output power P_{no}.

Step *VI* : Calculation of the output signal-to-noise ratio $(SNR)_o$ with the help of P_{so} and P_{no}.

Step *VII* : Calculation of the generalized form of noise figure with the help of step *III* and step *VI*.

Step *VIII* : Calculation of P_{no} from R_{eq} if possible and substitute into the general expression for noise figure F to get the actual formula or determination of P_{no} from measurement and substitute to obtain the formula for noise figure F.

Now, we shall discuss step-by-step as under:

Step *I*: From figure we can obtain signal input voltage V_{si} and Power P_{si} as under:

$$V_{si} = \frac{V_s . R_t}{R_a + R_t} \quad ...(6.82)$$

and $$P_{si} = \frac{V_{si}^2}{R_t} \quad ...(6.83)$$

Substituting the value of V_{si} from equation (6.82) in equation (6.83), we get

$$P_{si} = \frac{V_{si}^2}{R_t} = \frac{\left(\frac{V_s . R_t}{R_a + R_t}\right)^2}{R_t}$$

$$P_{si} = \frac{V_s^2 R_t}{(R_a + R_t)^2} \quad ...(6.84)$$

Step *II*: Similarly, we can obtain noise input voltage V_{ni} and power P_{ni} as under:

$$V_{ni} = \sqrt{4kTBR} = \sqrt{4kTB\left(\frac{R_a R_t}{R_a + R_t}\right)} \quad ...(6.85)$$

and $$P_{ni} = \frac{V_{ni}^2}{R_t}$$

Using equation (6.85), we get

$$P_{ni} = \frac{4kTB\left(\frac{R_a R_t}{R_a + R_t}\right)}{R_t} = \frac{4kTBR_a}{R_a + R_t} \quad ...(6.86)$$

Step *III*: The input signal-to-noise ratio will be found as

$$(SNR)_i = \frac{P_{si}}{P_{ni}}$$

Using equations (6.84) and (6.86), we get

$$(SNR)_i = \frac{V_s^2 R_t}{(R_a + R_t)^2} \div \frac{4kTBR_a}{R_a + R_t}$$

or $$(SNR)_i = \frac{V_s^2 R_t}{4kTBR_a(R_a + R_t)} \quad ...(6.87)$$

Step *IV*: The output signal power will be given as

$$P_{so} = \frac{V_{so}^2}{R_L} = \frac{(AV_{si})^2}{R_L} \quad ...(6.88)$$

or $$P_{so} = \frac{A^2 V_{si}^2}{R_L}$$

Substituting equation (6.82) in equation (6.88), we get

$$P_{so} = \frac{A^2\left(\frac{V_s R_t}{R_a + R_t}\right)^2}{R_L} = \frac{A^2 V_s^2 R_t^2}{(R_a + R_t)^2} \quad ...(6.89)$$

Step *V*: The noise output power P_{no} may be quite difficult to calculate. For instance, it can simply be written as

$$P_{no} = \text{Noise output power} \quad ...(6.90)$$

Step *VI*: Hence, the output signal-to-noise ratio $(SNR)_o$ will be found as

$$(SNR)_o = \frac{P_{so}}{P_{no}}$$

Using equations (6.89) and (6.90), we get

$$(SNR)_o = \frac{P_{so}}{P_{no}} = \frac{A^2 V_s^2 R_t^2}{(R_a + R_t)^2 R_L . P_{no}} \quad ...(6.91)$$

Step *VII*: Lastly, the general expression for the noise figure is

$$F = \frac{(SNR)_i}{(SNR)_o}$$

Using equations (6.87) and (6.91), we get

$$F = \frac{V_s^2 R_t}{4kTBR_a(R_a + R_t)} \div \frac{A^2 V_S^2 R_t^2}{(R_a + R_t)^2 R_L P_{no}} = \frac{R_L P_{no}(R_a + R_t)}{4kT\ BA^2 R_a R_t}$$

This is required expression.

6.30. Noise Figure from Equivalent Noise Resistance

As already discussed that the equivalent noise resistance R_{eq} of an amplifier or receiver is the sum of the input terminating resistance and the equivalent noise resistance of the first stage, together with the noise resistances of the previous stages referred to the input, *i.e*

$$R_{eq} = R_1 + \frac{R_2}{A_1^2} + \frac{R_3}{A_1^2 A_2^2} + \frac{R_4}{A_1^2 A_2^2 A_3^2} + ... \quad ...(6.92)$$

where R_{eq} = Equivalent noise resistance of a multistage amplifier or a receiver.

A_1, A_2, A_3 are gains of first stage, second stage and third stage, respectively, of a multistage amplifier or a receiver.

R_1, R_2, R_3, R_4 are the total input-noise resistances of first stage, second stage, third stage and fourth stage, respectively of an amplifier or a receiver.

Further, we may observe from another angle that all these resistances are added to R_t, giving a lumped resistance. This lumped resistance is then said to concentrate all the "noise making", of the receiver. The rest of it is now assumed to be noiseless. These noise resistances must now be added to the parallel combination of R_a and R_t.

Now, in order to correlate noise figure and equivalent noise resistance, let us define R'_{eq}.

R'_{eq} is a noise resistance. It does not incorporate R_t. It is given by

$$R'_{eq} = R_{eq} - R_t \quad ...(6.93)$$

The total equivalent noise resistance for the amplifier is given by

$$R = R'_{eq} + R_a \parallel R_t$$

$$R = R'_{eq} + \frac{R_a R_t}{R_a + R_t} \quad ...(6.94)$$

Hence, the equivalent noise voltage generated at the input of the receiver will be given as

$$V_{ni} = \sqrt{4kTBR} \quad ...(6.95)$$

But, since the amplifier has an overall voltage gain A, therefore this amplifier is treated as though it were noiseless.

The noise output now will be found as

$$P_{no} = \frac{V_{no}^2}{R_L} = \frac{(AV_{ni})^2}{R_L} = \frac{A^2 V_{ni}^2}{R_L} \quad ...(6.96)$$

Substituting equation (6.95) in equation (6.96), we get

$$P_{no} = \frac{A^2 V_{ni}^2}{R_L} = \frac{A^2 \left(\sqrt{4kTBR}\right)^2}{R_L}$$

or

$$P_{no} = \frac{A^2 4kTBR}{R_L} \quad ...(6.97)$$

It may be noted that we have already obtained the expression for noise figure in terms of P_{no} *i.e.*,

$$F = \frac{R_L P_{no}(R_a + R_t)}{4kTBA^2 R_a R_t} \quad ...(6.98)$$

Substituting equation (6.97) in equation (6.98), we get

$$F = \frac{R_L P_{no}(R_a + R_t)}{4kTBA^2 R_a R_t} = \frac{R_L \left[\frac{A^2 4kTBR}{R_L}\right](R_a + R_t)}{4kTBA^2 R_a R_t}$$

or

$$F = R\frac{(R_a + R_t)}{R_a R_t} \quad ...(6.99)$$

Substituting equation (6.94) in equation (6.99), we get

$$F = \left\{R'_{eq} + \frac{R_a R_t}{R_a + R_t}\right\}\left(\frac{R_a + R_t}{R_a R_t}\right)$$

@sTc

or

$$F = 1 + R'_{eq}\frac{(R_a + R_t)}{R_a R_t} \quad ...(6.100)$$

This equation is an expression for the noise figure in terms of the equivalent noise resistance.

It may be observed from equation (6.100) that if the noise has to be minimum for any given value of antenna resistance R_a, the ratio $(R_a + R_t)/R_t$ must also be a minimum, so that R_t is much larger than R_a. This situation is very much used in practice. Under these mismatched conditions $(R_a + R_t)/R_t$ approaches unity. Thus expression for the noise figure, now, will reduce to

$$F = 1 + \frac{R'_{eq}}{R_a} \quad ...(6.101)$$

This is the most important relationship between F and R'_{eq}.

Example 6.9. Determine the noise figure of an amplifier for which noise equivalent resistance R_{eq} = 2518 Ω and R_t = 600 Ω. It is driven by a generator whose output impedance is 50Ω. It is given that this constitutes a large enough mismatch.

Solution: Given that:

$$R_{eq} = 2518\ \Omega$$
$$R_t = 600\ \Omega$$
$$R_a = 50\ \Omega$$

We know that $R'_{eq} = R_{eq} - R_t$

or $R'_{eq} = 2518 - 600 = 1918\ \Omega$

Now, noise figure, $F = 1 + \frac{R'_{eq}}{R_a} = 1 + \frac{1918}{50}$

$$F = 1 + 38.4 = 39.4$$

$$F \text{ in (dB)} = 10 \log_{10}(39.4) = 15.84 \text{ dB}$$

It may be noted that if an "equivalent noise resistance" is given without any other comment in connection with noise figure calculations, it may be assumed to be R'_{eq}.

6.31. Experimental Determination of Noise Figure *(U.P. Technical University, 2003-04)*

In this article, let us describe a method used for measuring the noise figure of an amplifier or a receiver.

Let us assume that amplifier has a source impedance R_a. A diode is connected to the amplifier circuit as shown in figure 6.22. The diode is operated in the temperature-limited region so that the effect of space charge can be eliminated and the shot-noise current of the diode is expressed as

$$I_n = \sqrt{2qI_pB} \qquad ...(6.103)$$

where I_n = rms shot noise current

q = electronic charge = $e = 1.6 \times 10^{-19}$ coulomb

I_p = direct diode current

B = bandwidth of the interest

Equation (6.103) is an expression for the exact plate noise current of a vacuum-tube diode, and this can now be used. Anode current is controlled by means of the potentiometer which varies filament voltage, and that is how shot-noise current is adjusted.

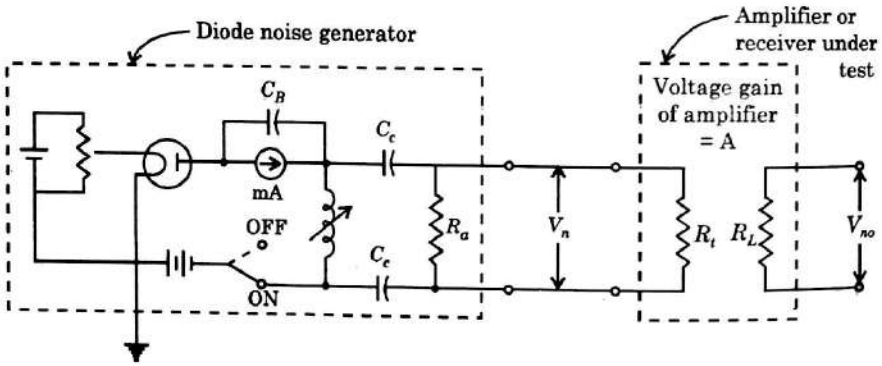


Fig. 6.22. *Experimental determination of noise figure (F).*

The output capacitance of the diode and its associated circuit is resonated at the operating frequency of the receiver with the help of the variable inductance, and thus it can be ignored. The output impedance of the noise generator or source will now simply be R_a. Hence, the noise voltage supplied to the input of the receiver by the diode is given by

$$V_n = I_nZ_n \qquad ...(6.104)$$

where

$$Z_n = R_a \parallel R_t = \frac{R_aR_t}{R_a + R_t} \qquad ...(6.105)$$

Substituting equations (6.103) and (6105) in equation (6.104), we get

$$V_n = I_nZ_n = \sqrt{2qI_pB}\left(\frac{R_aR_t}{R_a + R_t}\right)$$

or

$$V_n = \frac{R_aR_t}{R_a + R_t}\sqrt{2qI_pB} \qquad ...(6.106)$$

The noise generator or source is connected to the receiver (or amplifier) under test, and the noise output power of the receiver is measured with zero diode plate current, *i.e.*, with the diode plate voltage supply switched off. The diode plate voltage supply is now switched on and the filament potentiometer is adjusted so that diode plate current begins to flow. It is further adjusted until the noise power developed in R_L is twice as large as the noise power in the absence of diode plate current. The diode plate current at which this happens, I_p, is measured with the milliammeter and noted. The additional noise power output is now equal to the normal noise power output, so that the latter can be expressed in terms of the diode plate current.

Additional noise power output can be written as

$$P_{no} = \frac{V_{no}^2}{R_L} \qquad (\because V_{no} = A\,V_n)$$

or

$$P_{no} = \frac{(AV_n)^2}{R_L} = \frac{A^2V_n^2}{R_L} \qquad ...(6.107)$$

Substituting equation (6.106) in equation (6.107), we get

$$P_{no} = \frac{A^2V_n^2}{R_L} = \frac{A^2\left\{\frac{R_aR_t}{R_a + R_t}\sqrt{2qI_pB}\right\}^2}{R_L}$$

or

$$P_{no} = \frac{A^2R_a^2R_t^2\,2qI_pB}{R_L(R_a + R_t)^2} \qquad ...(6.108)$$

As already discussed, the general expression for noise figure is given as

$$F = \frac{R_LP_{no}(R_a + R_t)}{4kTBA^2R_aR_t} \qquad ...(6.109)$$

Substituting equation (6.108) in equation (6.109), we get

$$F = \frac{R_LP_{no}(R_a + R_t)}{4kTBA^2R_aR_t} = \frac{R_L\left\{\frac{A^2R_a^2R_t^2\,2qI_p}{R_L(R_a + R_t)^2}\right\}(R_a + R_t)}{4kTBA^2R_aR_t}$$

Simplifying, we get

$$F = \frac{qI_pR_aR_t}{2kT(R_a + R_t)} \qquad ...(6.110)$$

If we assume once again that the system is mismatched and $R_t >> R_a$, then the equation (6.110) can be simplified as

$$F = \frac{R_aqI_p}{2kT} \qquad ...(6.111)$$

If the above procedure is repeated right from the beginning for a system under matched conditions then equation (6.111) applies exactly to such a system instead of being merely a good approximation.

Putting the values of constants in the equation (6.111), we get

$$F = \frac{R_aqI_p}{2kT} = \frac{1.6\ 10^{-19}\ R_aI_p}{2 \times 1.38 \times 10^{-23} \times 290} = 20\ R_aI_p \qquad ...(6.112)$$

We have assumed that the normal temperature is 17°C or 290° K, where R_a is measured in ohms and I_p in amperes.

6.32. Noise Figure Determination for Cascaded Stages of Amplifiers

In this article we shall determine the noise figure (F) of a cascaded amplifier in terms of noise figure of individual stages of amplifiers. It is evident that the noise generated in the first stage is amplified by the subsequent stages. Therefore, the noise figure (F_1) of first stage is much more important compared to the later stages in a multistage amplifier to determine the overall noise figure of the amplifier. Let the two stages of the amplifiers have available power gains A_1 and A_2, respectively. Now, the overall noise figure F of the cascaded amplifier can be determined as under:

The total available noise power density S_{no} consists of the total noise power density S_{n1} available at the output due to the first stage. It also consists of the total noise power density S_{n2} at the output due to second stage of the amplifier.

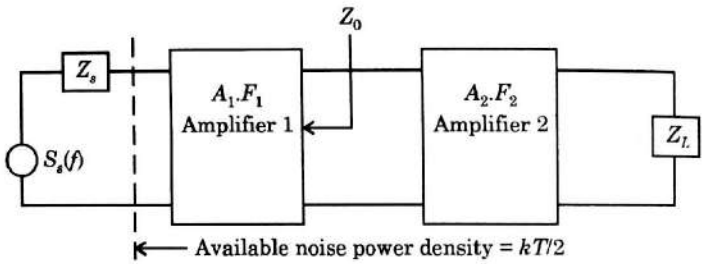


Fig. 6.23. *Determination of noise figure of cascaded amplifier or a multistage cascaded amplifier.*

Let F_1 and F_2 be the noise figures of the first stage and second stage of the amplifier, respectively. Total noise power density S_{n1} available at the output due to the first stage is given as

$$S_{n1} = \frac{kTF_1}{2} A_1A_2 \quad ...(6.113)$$

Let S_{n2} is the noise power density available at the output due to the second stage only. The source impedance for the second stage is the impedance seen at the output terminals of the first stage. Impedance seen at the output terminals of the first stage is denoted by Z_0. The noise component S_{n2} is due to noise sources inside the second stage amplifier only and it is immaterial whether Z_o is thermal or not. For convenience, we are assuming here that Z_o is a thermal.

Hence, the available power density is $kT/2$.

Let the noise figure of amplifier stage 2 be F_2. The available noise power density at the output due to amplifier in second stage will be given by

$$S_{n2} = \frac{kT}{2}(F_2 - 1)A_2 \quad ...(6.114)$$

Hence, the total noise power density delivered to the load by the cascaded amplifier is given as

$$S_n = S_{n1} + S_{n2} \quad ...(6.115)$$

Substituting equations (6.113) and (6.114) in equation (6.115), we get

$$S_n = S_{n1} + S_{n2} = \frac{kT}{2}F_1A_1A_2 + \frac{kT}{2}(F_2 - 1)A_2$$

or

$$S_n = \frac{kT}{2}[F_1A_1A_2 + (F_2 - 1)A_2] \quad ...(6.116)$$

Now, we are considering the overall amplifier for which noise power density S_n is expressed as

$$S_n = \frac{kT}{2}FA \quad ...(6.117)$$

where A = available gain of the cascaded amplifier

F = overall spectral noise figure

Because the overall available gain of a cascaded amplifier is the multiplication of the available gain of each stage, therefore,

we have

$$A = A_1A_2 \quad ...(6.118)$$

Substituting equation (6.118) in equation (6.117), we obtain

$$S_n = \frac{kT}{2}FA_1A_2 \quad ...(6.119)$$

On comparing equations (6.116) and (6.119), we get

$$S_n = \frac{kT}{2}[F_1A_1A_2 + (F_1 - 1)A_2] \quad ...(6.120)$$

$$S_n = \frac{kT}{2}FA_1A_2 \quad ...(6.121)$$

$$F = F_1 + \frac{(F_2 - 1)}{A_1} \quad ...(6.122)$$

It must be remembered that Noise Figure F_2 is the noise figure of the amplifier in second stage under the condition that the source is a positive impedance Z_0 equal to the output impedance of the amplifier in first stage.

In general, for a multistage amplifier, the overall spectral noise figure is given by

$$F = F_1 + \frac{F_2 - 1}{A_1} + \frac{F_3 - 1}{A_1A_2} + \frac{F_4 - 1}{A_1A_2A_3} + ... \quad ...(6.123)$$

From equation (6.123), we conclude that the first stage is most significant in determining the overall noise figure of the amplifier. Hence, in low noise amplifiers, the primary consideration is the design of the first stage which must have a low noise figure. This is sometimes obtained even at the cost of gain of the first stage of the multistage amplifier.

Example 6.10. A cascaded two-stage amplifier has been shown in figure 6.24. The first stage has a noise figure of 2 dB and a power gain of 12 dB. The second stage has a noise figure of 6 dB and a power gain of 10 dB. Find the overall Noise Figure in dB.

(U.P. Tech., Semester, Exam., 2004-05)

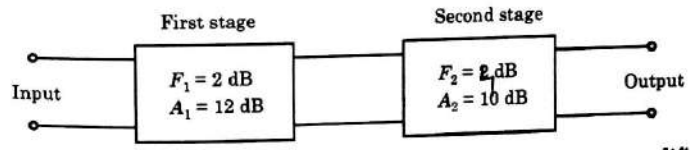


Fig. 6.24. *Overall noise figure F determination of a two-stages amplifier*

Solution: Firstly, we must express noise figures in terms of ratio, as under:

$$F_{dB} = 10 \log_{10}(F_{ratio})$$

or

$$2 = 10 \log_{10}(F_{1\ ratio})$$

or

$$F_{1,\ ratio} = \text{Antilog}_{10}\left(\frac{2}{10}\right)$$

$$F_{1,\ ratio} = \text{Antilog}_{10}(0.2) = 1.59$$

In the same way, we can express F_2 in ratio form as

$$F_{2,\ ratio} = 4$$

Thus, by using the above relationship between F_{dB} and F_{ratio} we can easily convert gain into ratio.

Therefore, $A_{1,\ ratio} = 15.9$

$A_{2,\ ratio} = 10$

We know that, the overall noise figure of a two-stage amplifier can be found by

$$F = F_1 + \frac{F_2 - 1}{A_1}$$

Here F, F_1, F_2 and A_1 are in ratio form only.

Therefore, $$F = 1.59 + \frac{4-1}{15.9} = 1.59 + \frac{3}{15.9} = 1.779$$

The overall noise figure of the two-stage amplifier in dB is found as

$$F_{dB} = 10 \log_{10} (F_{ratio})$$

or $$F_{dB} = 10 \log_{10} (1.779)$$

or $$F_{dB} = 2.5 \text{ dB} \quad \textbf{Ans.}$$

Example 6.11. A mixer stage has a noise figure of 20 dB. This mixer stage is preceded by an amplifier which has a noise figure of 9 dB and an available power gain of 15 dB. Find the overall noise figure referred to the input.

Solution: Given that, Noise figure of an Amplifier, $F_1 = 10$ dB

Noise figure of Mixer, $F_2 = 20$ dB

Available power gain of an amplifier, $A_1 = 15$

This is a two-stage system where both stages (Amplifier and Mixer) are cascaded.

Thus, $$F_1 \text{ in ratio} = \text{Antilog}_{10}\left(\frac{F_1 \text{ in dB}}{10}\right) = \text{Antilog}_{10}\left(\frac{9}{10}\right) = 7.94$$

$$F_2 \text{ in ratio} = \text{Antilog}_{10}\left(\frac{F_2 \text{ in dB}}{10}\right) = \text{Antilog}_{10}\left(\frac{20}{10}\right)$$

or $$F_2 \text{ in ratio} = \text{Antilog}_{10}(2) = 100$$

$$A_1 \text{ in ratio} = \text{Antilog}_{10}\left(\frac{A_1 \text{ in dB}}{10}\right) = \text{Antilog}_{10}\left(\frac{15}{10}\right)$$

or $$A_1 \text{ in ratio} = \text{Antilog}_{10}(1.5) = 31.62$$

Overall noise figure of the cascaded two-stage system referred to input can be found as under:

$$F = F_1 + \frac{F_2 - 1}{A_1}$$

Substituting all the values, we get

$$F = 7.94 + \frac{100-1}{31.62} = 7.94 + \frac{99}{31.62} = 11.07$$

Now, $$F \text{ in dB} = 10 \log_{10} (F_{ratio}) = 10 \log_{10} (11.07)$$

$$= 10.44 \text{ dB} \quad \textbf{Ans.}$$

Example 6.12. An amplifier operating on a frequency range from 18 to 20 MHz has a 10 kΩ input resistance. Find the rms noise voltage at the input to this amplifier if the ambient temperature is 27°C.

Solution: We know that the rms noise voltage of an amplifier is given by

$$V_n = \sqrt{4kTBR} \quad ...(i)$$

Given that the operating bandwidth of the amplifier is

$$B = (20 - 18) \text{ MHz} = 2 \text{ MHz} = 2 \times 10^6 \text{ Hz}$$

Temperature $T = (273°C + 17)$ K = 290 K

Input resistor $R = 10$ KΩ $= 10 \times 10^3\ \Omega = 10^4\ \Omega$

Substituting all the values in equation (*i*), we get

$$V_n = \sqrt{4kTBR} = \sqrt{4 \times 1.38 \times 10^{-23} \times 290 \times 2 \times 10^6 \times 10^4}$$

$$V_n = 1.78 \times 10^{-5} \text{ volt} = 17.8 \times 10^{-6} \text{ volt}$$

or $$V_n = 17.8\ \mu \text{ volt} \quad \textbf{Ans.}$$

or

Example 6.13. Two resistors R_1 and R_2 at absolute temperatures T_1 and T_2 are connected in series to form a white noise source. Find the equivalent noise temperature T_{eq}.

Solution: The equivalent resistance will be

$$R = R_1 + R_2$$

Hence, the available noise power will be

$$P_n = \frac{V_n^2}{4R} = \frac{V_{n1}^2 + V_{n2}^2}{4R}$$

Substituting for V_{n1} and V_{n2}, we get

or $$P_n = \frac{4kT_1BR_1 + 4kT_2BR_2}{4R} = (R_1T_1 + R_2T_2)\frac{kB}{R} \quad ...(i)$$

Therefore, the equivalent noise resistance R_{eq} is determined by

$$T_{eq} = \frac{P_n}{kB} \quad ...(ii)$$

Substituting equation (*i*) in equation (*ii*), we obtain

$$T_{eq} = \frac{(R_1T_1 + R_2T_2)\,kB/R}{kB}$$

or $$T_{eq} = \frac{(R_1T_1 + R_2T_2)}{R} \quad [\because R = R_1 + R_2]$$

Thus, $$T_{eq} = \frac{(R_1T_1 + R_2T_2)}{R_1 + R_2} \quad ...(iii)$$

Note that if $R_1 = R_2$, the noise resistance is the arithmetic average of the physical temperature, *i.e.*, $T_{eq} = (T_1 + T_2)/2$.

Because T_{eq} is not a physical temperature, the series connection is a non-thermal source, as should be expected since the resistors are not in thermal equilibrium. However, viewed from the terminals, this non-thermal source looks identical to thermal resistance R at temperature T_{eq}.

Example 6.14. The noise output of a resistor is amplified by a noiseless amplifier having a gain of 60 and a bandwidth of 20 kHz. A meter is connected to the output of the amplifier reads 1 m volt r.m.s.

(*i*) The bandwidth of the amplifier is reduced to 5 kHz, its gain remaining constant. What does the meter read now?

(*ii*) If the resistor is operated at 80°C, what is its resistance?

Solution: (*i*) Since the bandwidth is reduced to 1/4th of its value, therefore the noise voltage will be

$$V_n \propto \sqrt{B}$$

Hence, the noise voltage at 5 kHz will become half its value at 20 kHz bandwidth *i.e.*,

$$V_n = 0.5 \text{ m Volt}$$

(*ii*) Now $$V_n = \sqrt{4kTBR} \quad ...(i)$$

Here we have been given the following values of parameters.

$$V_{no} = \frac{10^{-3}}{60} \text{ volt}$$

$$B = 20 \text{ kHz} = 20 \times 10^3 \text{ Hz}$$

$$T = 80°\text{C}$$

$$T = 80 + 273 = 353° \text{ K}$$

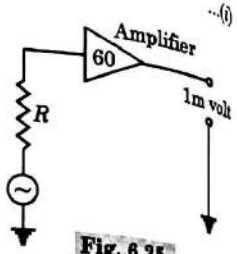


Fig. 6.25.

Substituting the above values in equation (*i*), we get

$$V_n = \sqrt{4kTBR}$$

$$\frac{10^{-3}}{60} = \sqrt{1 \times 1.38 \times 10^{-23} \times 353 \times 20 \times 10^3 \times R}$$

Solving, we get $R = 715 \text{ k}\Omega$ **Ans.**

Example 6.15. A coil having an inductance of 2 henry and an internal resistance of 1 Ohm is shunted by a capacitor of 2 farad. Determine the power density spectrum of the thermal noise across network terminals. *(GATE Examination-1999)*

Solution: The given network has been shown in figure 6.26. Because the network is a parallel circuit, its admittance can be found easily.

Admittance across terminal *AB* can be found as

$$Y_{AB} = 2jw + \frac{1}{1+2j\omega} = \frac{j2\omega - 4\omega^2 + 1}{1 + j2\omega}$$

or $$Y_{AB} = \frac{1 + j8\omega^3}{1 + 4\omega^2} \quad ...(i)$$

But $$Y_{AB} = G_{AB} + jH_{AB} \quad ...(ii)$$

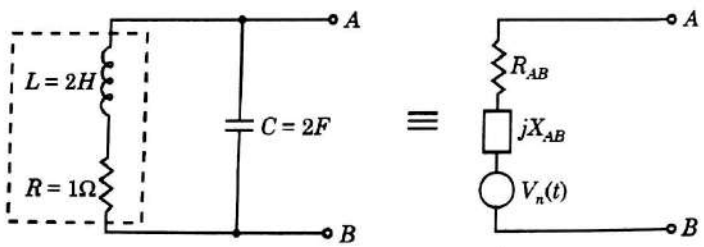


Fig. 6.26. *Network for example 6.15.*

On comparing equations (*i*) and (*ii*), we get

$$G_{AB} = \text{Conductance across terminal } AB$$

or $$G_{AB} = \frac{1}{1 + 4\omega^2} \quad ...(iii)$$

and $$jH_{AB} = \frac{j8\omega^2}{1 + 4\omega^2} \quad ...(iv)$$

Power density spectrum $S_1(\omega)$ of noise current will be

$$S_1(\omega) = 2kTG_{AB} \quad ...(v)$$

Substituting value of G_{AB} from equation (*iii*) in equation (*v*), we get

$$S_1(\omega) = 2kT \frac{1}{1 + 4\omega^2} = \frac{2kT}{1 + 4\omega^2} \quad ...(vi)$$

Impedance across terminal *AB* will be

$$Z_{AB} = \frac{1}{Y_{AB}} \quad ...(vii)$$

Substituting equation (*i*) in equation (*vii*), we get

$$Z_{AB} = \frac{1 + j2\omega}{1 - 4\omega^2 + j2\omega} = \frac{1 - j8\omega^2}{1 + 16\omega^4 - 4\omega^2} \quad ...(viii)$$

But $$Z_{AB} = R_{AB} + jX_{AB} \quad ...(ix)$$

On comparing equations (*viii*) and (*ix*), we obtain

$$R_{AB} = \frac{1}{1 + 16\omega^4 - 4\omega^2} \quad ...(x)$$

$$jX_{AB} = \frac{-j8\omega^3}{1 + 16\omega^4 - 4\omega^2} \quad ...(xi)$$

or $$X_{AB} = \frac{-8\omega^3}{1 + 16\omega^4 - 4\omega^2} \quad ...(xii)$$

Power density spectrum $S_v(\omega)$ of the noise voltage is given by

$$S_v(\omega) = 2kTR_{AB} \quad ...(xiii)$$

Substituting equation (*x*) in equation (*xiii*), we get

$$S_v(\omega) = 2kT \frac{1}{1 + 16\omega^4 - 4\omega^2} = \frac{2kT}{1 + 16\omega^4 - 4\omega^2} \quad \textbf{Ans.}$$

Example 6.16. Find the overall noise figure of a three stage cascaded amplifier, each stage having a power gain of 10 dB and noise figure of 6 dB.

Solution: Gain of each stage is

$$A_1 = A_2 = A_3 = 10 \text{ dB}$$

Noise figure of each stage is

$$F_1 = F_2 = F_3 = 6 \text{ dB}$$

Gain of each stage in ratio is found as

$$A_1 = A_2 = A_3 = \text{Antilog}_{10}\left(\frac{10}{10}\right) = 10$$

Noise figure of each stage will be given by

$$F_1 = F_2 = F_3 = \text{Antilog}_{10}\left(\frac{6}{10}\right) = 4$$

Overall noise figure of three stage cascaded amplifier is determined as

$$F = F_1 + \frac{F_2 - 1}{A_1} + \frac{F_3 - 1}{A_1 A_2}$$

Substituting all the values, we get

$$F = 4 + \frac{4-1}{10} + \frac{4-1}{10 \times 10}$$

or $$F = 4 + 0.3 + 0.03 = 4.33 \quad \textbf{Ans.}$$

Example 6.17. Obtain the equivalent noise temperature of the following low noise receiving system. Assume room temperature to be 17°C.

It is given that:

Low noise amplifier gain A_1 = 25 dB

Low noise amplifier noise temperature T_{e1} = 4 K

Pre-amplifier gain A_2 = 17 dB

Pre-amplifier noise figure F_2 = 6 dB

Converter and IF amplifier noise figure F_3 = 12 dB *(Delhi University-1999)*

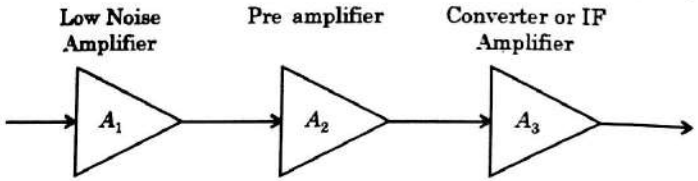


Fig. 6.27.

Solution: Firstly, let us convert all noise figures and gains in ratio with the help of following relation:

$$F_{dB} = 10 \log_{10} (F_{ratio})$$

Therefore,

$$A_{dB} = 10 \log_{10} (A_{ratio})$$

$$A_1 = 25 \text{ dB} = 316$$

$$T_{e1} = 4 \text{ K}$$

$$A_2 = 17 \text{ dB} = 50$$

$$F_2 = 6 \text{ dB} = 4$$

$$F_3 = 12 \text{ dB} = 16$$

@sTc

Equivalent noise temperature is given by

$$T_e = T_{e1} + \frac{T_{e2}}{A_1} + \frac{T_{e3}}{A_1 A_2} = T_{e1} + \frac{T_o(F_2 - 1)}{A_1} + \frac{T_o(F_3 - 1)}{A_1 A_2}$$

Substituting all the values, we get

$$= 4 + \frac{290(4-1)}{316} + \frac{(16-1)\,290}{316 \times 50}$$

or $$T_e = 7.025 \text{ K}$$

SUMMARY

1. In electrical terms, noise may be defined as an unwanted form of energy which tend to interfere with the proper reception and reproduction of transmitted signals.
2. Although there are several ways of classifying the noise, but conveniently noise may be classified in two broad groups as under:
   (*i*) External Noise, (*ii*) Internal Noise

3. External noise may be defined as that type of noise which is generated external to a communication system *i.e.* whose sources are external to the communication system.
4. External noise may be classified as under:
   (*i*) Atmospheric Noise
   (*ii*) Extraterrestrial Noise
   (*iii*) Industrial Noise
5. Atmospheric noise, which is also called static, is produced by lightning discharges in thunderstorms and other natural electrical disturbances which occur in the atmosphere.
6. There are several types of extraterrestrial noise or space noise depending upon their sources. Extraterrestrial noise may be divided into following two sub-groups as under:
   (*i*) Solar Noise
   (*ii*) Cosmic Noise
7. Solar noise is the electrical noise eminating from the Sun.
8. Distant stars can also be considered Sun. These distant stars have high temperatures and therefore radiate noise in the same manner as the Sun. The noise received from these distant stars is thermal noise and is distributed almost uniformly over the entire sky.
9. The industrial noise or man-made noise is that type of noise which is produced by such sources as automobiles and aircraft ignition, electrical motors, switch gears and leakage from high voltage transmission lines and several other heavy electrical equipments.
10. Internal noise is that type of noise which is generated internally or within the communication system or receiver.
11. Shot noise arises in active devices due to random behaviour of charge carries. In electron tubes, shot noise is generated due to the random emission of electrons from cathodes, whereas in semiconductor devices shot noise is generated due to the random diffusion of minority carries or simply random generation and recombination of electron-hole pairs.
12. The transit time of an electron in a diode depends upon anode voltage V and may be expressed as

$$\tau = 3.36 \times \frac{d}{\sqrt{V}} \ \mu \text{ sec.}$$

Here d is the spacing between anode and the cathode.

13. Partition noise is generated in a circuit when a current has to divide between two or more paths. This means that the partition noise results from the random fluctuations in the division.
14. At low frequencies (below few kHz), a particular type of noise appears. The power spectral density of this noise increases as the frequency decreases. This noise is called as flicker noise or (1/*f*) noise. In case of vacuum tubes, the main causes of flicker noise are slow changes which take place in the oxide structure of oxide coated cathodes and migration of impurity ions.
15. In semiconductor devices, flicker noise is generated from the fluctuations in the carrier density and creates more problems in semiconductor amplifying device than in vacuum tubes at low frequencies.
16. It is generally observed in semiconductor devices, when the transit-time of charge-carries crossing a junction is comparable with the time-period of the signal, some charge-carries diffuse back to the source or emitters.
17. The thermal noise or white noise or Johnson noise is the random noise which is generated in a resistor or the resistive component of a complex impedance due to rapid and random motion of the molecules, atoms and electrons.
18. The expression for maximum noise power output of a resistor may be given as

$$P_n \propto T.B$$

or $$P_n = k.T.B$$

where k = Baltzmann's constant
$= 1.38 \times 10^{-23}$ Joule/deg. K

T = absolute temperature
B = Bandwidth of interest in Hz.

19. The resultant noise in this case will be

$$V_{nr} = \sqrt{4kTBR_{par}}.$$

where $$\frac{1}{R_{par}} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \ldots$$

20. The power density spectrum of the current contributing the thermal noise is expressed as

$$S_i(\omega) = \frac{2kTG}{1+\left(\frac{\omega}{\alpha}\right)^2}$$

Here, T is ambient temperature in kelvin.
G is the conductance of the resistor in mhos.
k is the Boltzmann constant.
and α is the average number of collisions per second per electron.

21. Thermal noise can be considered to contain all frequency components in equal amount. Due to this reason, thermal noise is generally referred to as white noise (here white means the presence of all colours or frequencies).
22. Thermal noise is also known as Johnson noise after the scientist J.B. Johnson who presented a detailed investigation of it.
23. Noise in an idealized form is known as white noise. This means that in a communication system, the noise analysis is based on an idealized form of noise, *i.e.*, white noise.
24. The power spectrum density of white noise is expressed as

$$S_w(j\omega) = \frac{N_0}{2}$$

Here, the factor 1/2 has been included to show that half of the power is associated with the positive frequencies and remaining half with the negative frequencies.

25. Strictly speaking, white noise has infinite average power and thus is not physically realizable. However, the concept of white noise provides convenient mathematical properties and this is a useful tool for system analysis.
26. "Equivalent noise bandwidth may be defined as the bandwidth of an ideal bandpass system which produces the same noise power as the actual system does."
27. The ratio of signal power to the associated noise power is defined as signal to noise ratio. In other words, signal to noise ratio is defined as the ratio of signal power to the noise power at the same point in the system. Signal to noise ratio is denoted by S/N.
28. Because the power spectrum density is power per unit bandwidth, hence the above expression may be expressed as

$$\frac{S}{N} = \frac{S_s(j\omega)}{S_n(j\omega)}$$

or $$\frac{S}{N} = \frac{\text{Power spectrum density of signal voltage}}{\text{Power spectrum density of noise voltage}}$$

29. The noise figure F is defined as the ratio of the signal-to-noise power ratio supplied to the input terminals of a receiver or amplifier to the signal-to-noise power ratio supplied to the output or load resistor.
Hence, mathematically,

$$\text{Noise Figure, } F = \frac{\text{Input } SNR}{\text{Output } SNR} = \frac{(SNR)_i}{(SNR)_o}$$

SHORT QUESTIONS WITH ANSWERS

Q.1. Define noise.

Ans. In electrical terms, noise may be defined as an unwanted form of energy which tend to interfere with the proper reception and reproduction of transmitted signals. For example, in receivers, several electrical disturbances produce noise and thus modifying the required signal in an unwanted form. In case of radio receivers, noise may produce hiss type sound in the output of loudspeaker. Similarly, in T.V. receivers, noise may produce 'snow' which becomes superimposed on the picture output. In addition to this, in pulse communications, noise may produce unwanted pulses or cancel the required pulses. In other words, we can say that noise may limit the performance of a communication system.

Q.2. How can you classify noise?

Ans. Noise may be classified in two broad groups as under:
(*i*) External Noise
(*ii*) Internal Noise.

Q.3. What is an external noise?

Ans. External noise may be defined as that type of noise which is generated external to a communication system *i.e.* whose sources are external to the communication system. External noise cannot be analysed quantitatively.

Q.4. How can you classify external noise.

Ans. External Noise may be classified as under:
(*i*) Atmospheric Noise
(*ii*) Extraterrestrial Noise
(*iii*) Industrial Noise.

Q.5. What is an internal noise?

Ans. Internal noise is that type of noise which is generated internally or within the communication system or receiver. Internal noise may be treated quantitatively and can also be reduced or minimized by proper system design.

Q.6. How can you classify internal noise.

Ans. Internal noise may be classified as under:
(*i*) Shot noise
(*ii*) Partition noise
(*iii*) Low frequency or Flicker noise
(*iv*) High-frequency or transit-time noise
(*v*) Thermal noise.

Q.7. Write the expression for thermal noise generated in a resistor. *(Madras University-1998)*

Ans. the expression for maximum noise power output of a resistor may be given as

$$P_n \ \alpha \ T.B = k.T.B$$

where k = Baltzmann's constant = 1.38×10^{-23} Joule/deg. K
T = absolute temperature
B = Bandwidth of interest in Hz.

Q.8. What is white noise?

Ans. Noise in an idealized form is known as white noise. This means that in a communication system, the noise analysis is based on an idealized form of noise, *i.e.*, white noise.

Q.9. Write the expression for PSD of white noise.

Ans. The power spectrum density of white noise is expressed as

$$S_w(w) = \frac{N_0}{2}$$

Here, the factor 1/2 has been included to show that half of the power is associated with the positive frequencies and remaining half with the negative frequencies.

Q.10. Define equivalent noise bandwidth of an ideal bandpass system.

Ans. "Equivalent noise bandwidth may be defined as the bandwidth of an ideal bandpass system which produces the same noise power as the actual system does."

REVIEW QUESTIONS

1. What do you mean by the term 'noise' used in communication system?
2. What is meant by the term 'atmospheric noise'?
3. What are the causes and effects of solar noise?
4. What is meant by the term 'cosmic noise'? Give its salient features.
5. Name the different noises which may be created within a receiver or amplifier. Discuss the effects of these noises on the performance of a receiver.
6. Explain thermal noise.
7. What is shot noise? Write expression for the shot noise current in a diode.
8. What is meant by the transit time noise in a transistor?
9. Define transistor signal-to-noise ratio and noise figure of a receiver.

NUMERICAL PROBLEMS

1. An amplifier operating over the frequency range of 470 to 480 kHz has 300 kW input resistor. Calculate the rms noise voltage at the input of this amplifier of the ambient temperature is 27°C? [Ans. 7.05 μV]
2. A parallel tuned circuit having $Q = 20$ resonates at 5 MHz with 20 pF capacitor. If the circuit is maintained at 17°C, what noise voltage will a wideband voltmeter measure when placed across it?

□□□

Download From www.dahiyastech.blogspot.com

For More Books Comment Now

CHAPTER 7

Performance of Communication Systems

Inside this Chapter

7.1. Introduction 403
7.2. Bandpass Noise Model 404
7.3. Noise in Communication System 405
7.4. Noise Calculation in Various Amplitude Modulation (AM) Systems 406
7.5. Double Sideband Suppressed Carrier (DSB-SC) System 407
7.6. Single Sideband Suppressed Carrier (SSB-SC) System 409
7.7. Amplitude-Modulation System (Envelope Detector Method) 411
7.8. Evaluation of Noise in Angle Modulated Systems 413
7.9. Evaluation of Noise in Frequency Modulation (FM) Systems 414
7.10. Noise in Phase Modulation (PM) 421
7.11. Noise in Pulse-Modulated Systems 425
7.12. Noise in Pulse Amplitude Modulation (PAM) 425
7.13. Noise in Pulse Position Modulation (PPM) 426
7.14. Coded Systems 428
7.15. Noise in Binary PCM 428
7.16. Comparison of Coded and Uncoded Systems 431

7.1. Introduction

In this chapter, we shall study continuous wave modulation systems from noise point of view, *i.e.*, we shall evaluate the performance of continuous wave modulation systems in the presence of noise. In our noise analysis, we shall assume noise to be white and Gaussian just to simplify the noise analysis. Noise analysis in continuous wave (CW) modulation is carried out in the form of a parameter known as **figure of merit** denoted by γ. This parameter figure of mert γ is the ratio of output singal-to-noise ratio to the input signal-to-noise ratio of the receiver. This means that we compare output signal-to-noise ratio (S_0/N_0) with the input signal-to-noise ratio (S_i/N_i) of the receiver to evaluate the noise performance.

In this chapter, we shall consider two forms of continuous-wave (CW) modulation systems namely Amplitude Modulation (AM) and Angle Modulation from noise point of view. Further, in AM case,

we shall consider the double-sideband suppressed carrier (DSB-SC), single-sideband suppressed carrier (SSB-SC) and vestigial sideband (VSB). In case of angle modulation, we shall consider two cases of angle modulation namely frequency modulation and phase modulation.

7.2. Bandpass Noise Model

We know that in any practical communication system, the front end of the receiver consists of a tuned circuit also known as bandpass filter. This bandpass filter processes the incoming desired signal and the noise. Infact, this tuned circuit is so designed as to have a bandwidth large enough to pass the desired signal without any distortion and hence is also called bandpass filter. This filter is usually narrowband in the sense that its bandwidth or passband is quite small compared to the centre or midband frequency. With the desired signal, the noise signal is also passed through this bandpass filter. The noise signal appearing at the output of this bandpass filter is known as bandpass noise or simply narrowband noise.

It may be shown that a bandpass or narrowband noise concentrated around a midband or carrier frequency $\pm w_c$ may be represented in the time-domain form approximately by a sinusoidal wave of frequency w_c which varies slowly in both amplitude and phase as shown in figure 7.1(a).

Let us consider that a bandpass noise $n(t)$ which is produced at the output of a bandpass filter in response to the white and Gaussian noise at the input of bandpass filter with zero mean and unity power spectrum density (PSD).

Let $H(w)$ be the transfer function of the bandpass filter.

The power spectrum density $S_n(w)$ of bandpass noise $n(t)$ may be expressed in terms of $H(w)$ as

$$S_n(w) = |H_n(w)|^2 \quad ...(7.1)$$

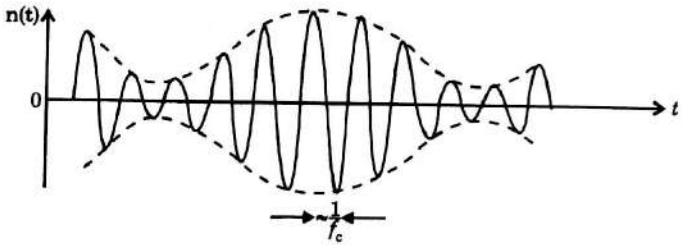


Fig. 7.1. *(a) Bandpass noise representation in time-domain.*

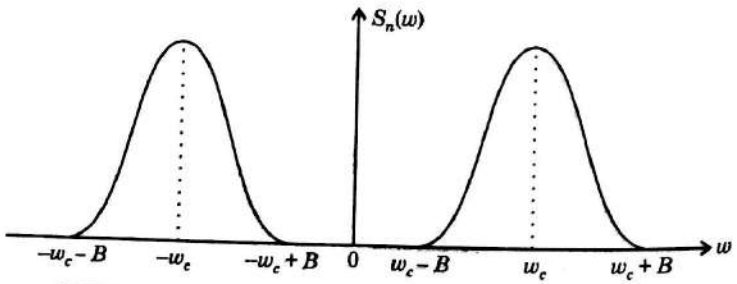


Fig. 7.1. *(b) Power spectrum density of bandpass noise.*

Now, we shall represent the bandpass noise in terms of in-phase and quadrature components. This may be derived by utilizing the concept of pre-envelope and complex envelope.

Let $n_p(t)$ and $n_n(t)$ be the pre-envelope and the complex envelope of the bandpass noise respectively. Let us assume that the power spectrum density of bandpass noise $n(t)$ is centred about the carrier frequency w_c.

Then we can write

$$n_p(t) = n(t) + j\hat{n}(t) \quad ...(7.2)$$

and

$$n_h(t) = n_p(t)\, e^{-jw_c t} \quad ...(7.3)$$

where $\hat{n}(t)$ is the Hilbert transform of bandpass noise $n(t)$.

Further, the complex envelope $n_h(t)$ may itself be expressed as

$$n_h(t) = n_c(t) + jn_s(t) \quad ...(7.4)$$

Now, using equations (7.2), (7.3) and (7.4) we obtain the in-phase and quadrative components of bandpass noise $n(t)$ as:

inphase noise component

$$n_c(t) = n(t)\cos w_c t + \hat{n}(t)\sin w_c t \quad ...(7.5)$$

and quadrature noise component

$$n_s(t) = \hat{n}(t)\cos w_c t - n(t)\sin w_c t \quad ...(7.6)$$

Now, eliminating $\hat{n}(t)$ from equations (7.5) and (7.6), we have bandpass noise as

$$n(t) = n_c(t)\cos w_c t - n_s(t)\sin w_c t \quad ...(7.7)$$

which is the desired bandpass noise model.

7.3. Noise in Communication System

We know that in a communication system, the message signal travels from the transmitter to the receiver through a medium called channel (*i.e.* radio or line). Now, noise is present in every communication system. Infact, the channel introduces an additive noise in the message signal and thus the message which is received at the receiver is distorted. Since, the receiver detects both message signal and the noise signal, it will reproduce a message signal which contains noise. The noise calculation in a communication system is carried out in the form of a parameter which is known as figure of merit. This parameter, figure of merit is denoted by lettet 'γ'.

The figure of merit 'γ' may be defined as the ratio of output signal to noise ratio to input signal to noise ratio of a receiver.

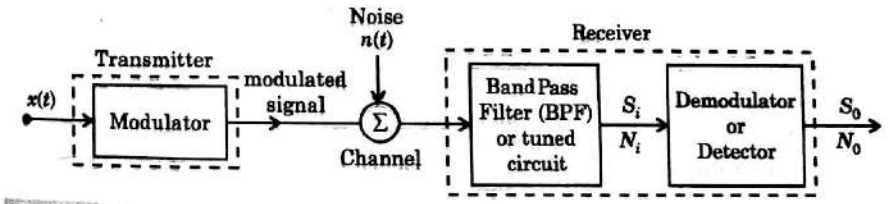


Fig. 7.2. *Generalised equivalent model of a communication system for noise calculation.*

Mathematically,

$$\gamma = \frac{\text{Output signal to noise ratio}}{\text{Input signal to noise ratio}}$$

$$\gamma = \frac{S_0 / N_0}{S_i / N_i} \quad ...(7.8)$$

From this expression, it is evident that a communication system with higher γ will have a better noise performance, *i.e.*, the adverse effect of noise is small.

7.3.1. Few Assumptions to Calculate the Figure of Merit for Various Communication Systems

We can evaluate the noise performance of a communication system by calculating the figure of merit for that particular communication system. Now in calculating the figure of merit, we make few assumptions as under:

(i) Channel Noise is Always White and Gaussian. We assume that the channel noise $n(t)$ is always a white noise. This means that it is uniformly distributed over the entire band of frequencies

under consideration. Hence, the power spectrum density of channel noise will be uniform over the frequency range under consideration. Thus, the total noise power N may be obtained by taking the product of noise power spectrum density $n/2$ with the bandwidth under consideration, *i.e.*,

Total noise power N = white noise power spectrum density × Bandwidth

or $$N = \frac{n}{2} \times \text{Bandwidth} \quad ...(7.9)$$

Thus, the noise has a Gaussian distribution.

(ii) Channel Noise is Always Additive. We assume that the disturbing effect of channel noise is always additive. This means that the effect of channel noise may be obtained by simple addition of signal $x(t)$ and noise $n(t)$.

(iii) The Noise at the Input of Demodulator is a Bandpass Noise. We know that the first range of each receiver is a tuned circuit which works as a bandpass filter (BPF). The function of this tuned circuit or bandpass filter is to allow only a narrowband signal centred about carrier frequency $\pm f_c$ and reject all other frequencies. This means that the noise signal lying outside this range is also rejected and thus the bandwidth of the noise signal at the input of detector will be same as that of the incoming modulated signal. But in assumption (*i*), we have already assumed that the channel noise is white in nature. Therefore the noise (*i.e.* white noise) at the input of the demodulator will have a constant or uniform power spectrum density over the passband. This power spectrum density of the white noise at the input of demodulator is given by

$$S_{ni} = \frac{\eta}{2} \quad ...(7.10)$$

For example, the passband for amplitude modulation is twice the maximum modulating frequency f_m around the carrier frequency f_c as shown in figure 7.3.

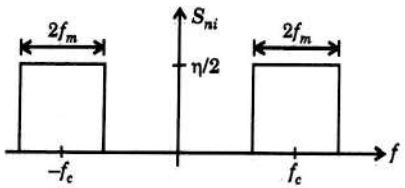


Fig. 7.3. *Power spectrum density of the bandpass white noise at the input of demodulator*

(iv) The Noise Input Power N_i. To calculate the figure of merit γ, we evaluate the noise input power N_i for passband or for bandwidth of incoming modulated signal ($2f_m$ in case of amplitude modulation).

Using assumption (*i*), the total noise power N may be obtained as

$$N = \frac{\eta}{2} \times \text{bandwidth}$$

Therefore, the input noise power N_i at the input of demodulator for AM is given as

$$N_i = \frac{\eta}{2} \times 2f_m = \eta f_m \quad ...(7.11)$$

We shall treat above equation as the reference input noise power to calculate figure of merit γ for various AM and FM receiver.

7.4 Noise Calculation in Various Amplitude Modulation (AM) Systems

Now, we shall find figure of merit for various AM systems to analyze their performance from noise point of view.

7.5 Double Sideband Suppressed Carrier (DSB-SC) System

(U.P. Tech., Semester, Exam., 2004-05) (06 marks)

We know that the DSB-SC system uses coherent or synchronous detection scheme at the receiver. Figure 7.4 shows the block diagram of DSB-SC system from noise point of view.

To judge the performance of a DSB-SC system, we shall calculate its figure of merit γ which is given as

$$\gamma = \frac{S_0 / N_0}{S_i / N_i} \quad ...(7.12)$$

where
S_0 = output signal power
N_0 = output noise power
S_i = input signal power at the detector
N_i = input noise power at the detector

Now, we shall find all the above four factors separately and then substitute their values in equation (7.12) to calculate the figure of merit γ.

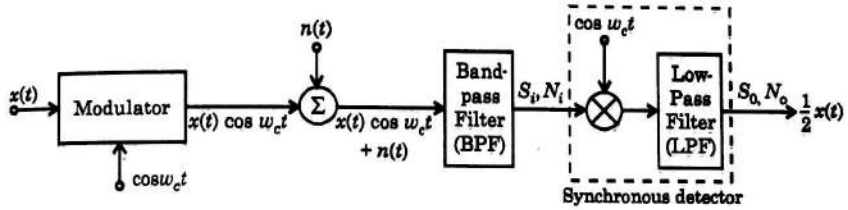


Fig. 7.4. *A DSB-SC system from noise point of view.*

(i) Input Signal Power: The modulated (DSB-SC) signal at the input of the demodulator is given as

$$s(t) = x(t) \cos w_c t \quad ...(7.13)$$

where $x(t)$ is the modulating signal and $\cos w_c t$ is the carrier signal.

Thus the input signal power S_i is given as the mean square (ms) value of signal $s(t) = x(t) \cos w_c t$

Hence, S_i = ms value of $s(t)$

$$S_i = \overline{[x(t) \cos w_c t]^2} = \frac{1}{2}\overline{x^2(t)} \quad ...(7.14)$$

(ii) Output Signal Power: We know that in synchronous detection method, the incoming DSB-SC signal at the detector input is multiplied by a synchronous carrier $\cos w_c t$. Thus, the signal at the multiplier output is

$$e(t) = x(t) \cos w_c t \cdot \cos w_c t \quad ...(7.15)$$

$$e(t) = x(t) \cos^2 w_c t = \frac{1}{2} x(t)[2 \cos^2 w_c t]$$

or $$e(t) = \frac{1}{2} x(t)[1 + \cos 2w_c t]$$

or $$e(t) = \frac{1}{2} x(t) + \frac{1}{2} x(t) \cos 2w_c t \quad ...(7.16)$$

This signal is finally passed through the low pass filter (LPF). Thus the signal at the demodulator output is

$$s_0(t) = \frac{1}{2} x(t) \quad ...(7.17)$$

Therefore, the output signal power at the demodulator output is given as

$$S_0 = \text{ms value of } \frac{1}{2}x(t) = \overline{\left[\frac{1}{2}x(t)\right]^2} = \frac{1}{4}\overline{x^2(t)} \quad ...(7.18)$$

But according to equation (7.14), we have

$$S_i = \frac{1}{2}\overline{x^2(t)}$$

Hence, $$S_0 = \frac{1}{2}\cdot S_i$$

or $$\frac{S_0}{S_i} = \frac{1}{2} \quad ...(7.19)$$

***(iii)* Input Noise Power:** The noise at the input of the receiver is white and Gaussian in nature. This noise after passing through the bandpass filter (BPF) is converted into a bandpass noise. The input bandpass noise is given as

$$n_i(t) = n_c(t)\cos w_c t - n_s(t)\sin w_c t \quad ...(7.20)$$

The input noise power at the input of demodulator denoted by N_i is given as

$$N_i = \overline{n_i^2(t)} = \overline{n_c^2(t)} = \overline{n_s^2(t)} \quad ...(7.21)$$

***(iv)* Output Noise Power:** Now, we shall find the noise power N_0 at the output of demodulator.

Since the noise signal at the input of the demodulator is a bandpass noise, therefore, it is expressed as

$$n(t) = n_c(t)\cos w_c t - n_s(t)\sin w_c t \quad ...(7.22)$$

This noise $n(t)$ is multiplied by synchronous carrier signal $\cos w_c t$ in synchronous demodulator. Thus the multiplied signal n_d is expressed as

$$n_d(t) = n(t)\cos w_c t = [n_c(t)\cos w_c t - n_s(t)\sin w_c t)]\cos w_c t$$
$$= n_c(t)\cos^2 w_c t - n_s(t)\cos w_c t \sin w_c t$$

or $$n_d(t) = \frac{1}{2}n_c(t)\left[2\cos^2 w_c t\right] - \frac{1}{2}n_s(t)\left[2\sin w_c t\cos w_c t\right]$$

or $$n_d(t) = \frac{1}{2}n_c(t)\left[1+\cos 2w_c t\right] - \frac{1}{2}n_s(t)\sin 2w_c t$$

or $$n_d(t) = \frac{1}{2}n_c(t) + \frac{1}{2}n_c(t)\cos 2w_c t - \frac{1}{2}n_s(t)\sin 2w_c t \quad ...(7.23)$$

This noise signal $n_d(t)$ is passed through a low pass filter (LPF). Since the terms $\frac{1}{2}n_c(t)\cos 2w_c t$ and $\frac{1}{2}n_s(t)\sin 2w_c t$ are centered around $\pm 2w_c$, therefore, they are filtered out by the low pass filter.

This means that the final noise $n_0(t)$ at the output of the demodulator is only $\frac{1}{2}n_c(t)$.

Thus, $$n_0(t) = \frac{1}{2}n_c(t) \quad ...(7.24)$$

Thus the output noise power N_0 will be

$$N_0 = \text{m.s. value of } n_0(t)$$
$$N_0 = \overline{n_0^2(t)}$$

or $$N_0 = \frac{1}{4}\overline{n_c^2(t)} = \frac{1}{4}N_i$$

Hence, $$\frac{N_0}{N_i} = \frac{1}{4} \quad ...(7.25)$$

Using equations, we get the figure of merit γ for DSB-SC system as

$$\gamma = \frac{S_0/N_0}{S_i/N_i} = \frac{S_0/S_i}{N_0/N_i} = 2 \quad ...(7.26)$$

7.6 Single Sideband Suppressed Carrier (SSB-SC) System

We know that the receiver of an SSB-SC system is similar to that for DSB-SC system for synchronous detection except for the fact that the bandwidth of the bandpass filter of SSB-SC receiver is one-half of that required for DSB-SC. Hence, the procedure for evaluating the figure of merit is similar to that of DSB-SC system.

Thus, to judge the performance of a SSB-SC system, we shall calculate its figure of merit γ which is given as

$$\gamma = \frac{S_0/N_0}{S_i/N_i} \quad ...(7.27)$$

where
S_0 = output singal power
N_0 = output noise power
S_i = input signal power at the detector
N_i = input noise power at the detector

Now, we shall find all the above four factors separately and then substitute their values in equation (7.27) to calculate the figure of merit γ.

(i) Input Signal Power: We know that the incoming signal to the SSB-SC receiver may be expressed as

$$s(t)_{SSB} = x(t)\cos w_c t \pm x_h(t)\sin w_c t \quad ...(7.28)$$

Here the plus sign is for upper sideband and minus sign is for lower sideband, $x(t)$ is the message signal, $x_h(t)$ is its Hilbert transform.

The signal power available at the input of the demodulator is

$$S_i = \frac{1}{2}\overline{x^2(t)} + \frac{1}{2}\overline{x^2{}_h(t)} \quad ...(7.29)$$

$$S_i = \overline{x^2(t)} \quad ...(7.30)$$

At this point, it may be noted that the average power of the baseband signal $x(t)$ is same as that of its Hilbert transform. This is due to the fact that $x_h(t)$ is basically the signal $x(t)$ in which all the frequency components undergo a phase shift of $\pi/2$. Hence, $x(t)$ and $x_h(t)$ occupy the same spectral range *i.e.*, w_m.

Moreover, the power spectrum density of $x(t)$ and $x_h(t)$ are the same. This means that the average power of the message or baseband signal and its Hilbert transform are the same.

***(ii)* Output Signal Power S_0:** Since, the incoming signal is multiplied by a synchronous local carrier $\cos w_c t$, therefore the output will be

$$e(t) = [x(t)\cos w_c t \pm x_h(t)\sin w_c t]\cos w_c t \quad ...(7.31)$$

or $$e(t) = x(t)\cos^2 w_c t \pm x_h(t)\sin w_c t\cos w_c t$$

or $$e(t) = \frac{1}{2}x(t)[1+\cos 2w_c t] + \frac{1}{2}x_h(t)[\sin 2w_c t] \quad ...(7.32)$$

This signal is finally passed through the low-pass filter (LPF) which rejects all the terms excepting the term which occupies a bandwidth w_m.

Hence, the final demodulated output signal is

$$s_0(t) = \frac{1}{2}x(t) \quad ...(7.33)$$

Hence, the output signal power of the demodulator is

$$S_0 = \frac{1}{4}\overline{x^2(t)} = \frac{1}{4}S_i$$

or
$$\frac{S_0}{S_i} = \frac{1}{4} \quad ...(7.34)$$

(iii) Input Noise Power: The noise at the input of the receiver is white and Gaussian in nature. This noise after passing through the bandpass filter (BPF) is converted into a bandpass noise. The input bandpass noise is given as

$$n_i(t) = n_c(t)\cos w_c t - n_s(t)\sin w_c t \quad ...(7.35)$$

The input noise power at the input of demodulator denoted by N_i is given as

$$N_i = \overline{n_i^2(t)} = \overline{n_c^2(t)} = \overline{n_s^2(t)} \quad ...(7.36)$$

(iv) Output Noise Power: Now, we shall find the noise power N_0 at the output of demodulator.

Since the noise signal at the input of the demodulator is a bandpass noise, therefore, it is expressed as

$$n(t) = n_c(t)\cos w_c t - n_s(t)\sin w_c t \quad ...(7.37)$$

This noise $n(t)$ is multiplied by synchronous carrier signal $\cos w_c t$ in synchronous demodulator. Thus the multiplied signal n_d is expressed as

$$n_d(t) = n(t)\cos w_c t = [n_c(t)\cos w_c t - n_s(t)\sin w_c t)]\cos w_c t$$
$$= n_c(t)\cos^2 w_c t - n_s(t)\cos w_c t \sin w_c t$$

or
$$n_d(t) = \frac{1}{2}n_c(t)\left[2\cos^2 w_c t\right] - \frac{1}{2}n_s(t)\left[2\sin w_c t\cos w_c t\right]$$

or
$$n_d(t) = \frac{1}{2}n_c(t)\left[1+\cos 2w_c t\right] - \frac{1}{2}n_s(t)\sin 2w_c t$$

or
$$n_d(t) = \frac{1}{2}n_c(t) + \frac{1}{2}n_c(t)\cos 2w_c t - \frac{1}{2}n_s(t)\sin 2w_c t \quad ...(7.38)$$

This noise signal $n_d(t)$ is passed through a low pass filter (LPF). Since the terms $\frac{1}{2}n_c(t)\cos 2w_c t$ and $\frac{1}{2}n_s(t)\sin 2w_c t$ are centred around $\pm 2w_c$, therefore, they are filtered out by the low pass filter.

This means that the final noise $n_0(t)$ at the output of the demodulator is only $\frac{1}{2}n_c(t)$.

Thus,
$$n_0(t) = \frac{1}{2}n_c(t) \quad ...(7.39)$$

Thus the output noise power N_0 will be

$$N_0 = \text{m.s. value of } n_0(t) = \overline{n_0^2(t)}$$
$$= \frac{1}{4}\overline{n_c^2(t)} = \frac{1}{4}N_i$$

Hence,
$$\frac{N_0}{N_i} = \frac{1}{4} \quad ...(7.40)$$

Using equations (7.34) and (7.40), we get the figure of merit γ for SSB-SC system as

$$\gamma = \frac{S_0/N_0}{S_i/N_i}$$

or
$$\gamma = \frac{S_0/S_i}{N_0/N_i} = \frac{1/4}{1/4} = 1 \quad ...(7.41)$$

Amplitude-Modulation System (Envelope Detector Method)

(i) Input Singal Power: We know that in amplitude modulation system, a large carrier is accompanied with the two upper and lower sidebands. Now, if $n_i(t)$ is the additive noise signal, then the input to the detector is expressed as

$$s(t)_{AM} = [A + x(t)]\cos w_c t + n_i(t) \quad ...(7.42)$$

The input signal power S_i is given as,

$$S_i = \text{ms value of carrier + ms value of sidebands}$$
$$= \frac{1}{2}\left[A^2 + \overline{x^2(t)}\right] \quad ...(7.43)$$

(ii) Output Power: The output of the envelope detector (popularly used in AM receivers) will be the envelope of the AM signal $s(t)_{AM}$ given in equation (7.42). Thus, to find the detected output, we are required to find the envelope of $s(t)_{AM}$. Substituting bandpass noise model for $n_i(t)$ in equation (7.42), we get

$$s(t)_{AM} = [A + x(t)]\cos w_c t + n_c(t)\cos w_c t - n_s(t)\sin w_c t$$
$$= [A + x(t) + n_c(t)]\cos w_c t - n_s(t)\sin w_c t \quad ...(7.44)$$

After trigonometric manipulation, equation (7.44) becomes

$$s(t)_{AM} = A(t)\cos[w_c t + \phi(t)] \quad ...(7.45)$$

where $A(t)$, and $\phi(t)$, respectively, are randomly time varying amplitude and phase angle of $s(t)_{AM}$. These are given as

$$A(t) = \sqrt{[A + x(t) + n_c(t)]^2 + n_s^2(t)} \quad ...(7.46)$$

$$\phi(t) = \tan^{-1}\left[\frac{n_s(t)}{A + x(t) + n_c(t)}\right] \quad ...(7.47)$$

The time varying amplitude $A(t)$ is the envelope of $s(t)_{AM}$ and therefore, the output of the envelope detector will be envelope $A(t)$. The envelope $A(t)$ has both signal and noise components. The noise performance depends on the relative magnitudes of the signal and noise. The analysis will be carried out for two cases, as under

(i) Small noise case,
(ii) Large noise case

(i) Small Noise Case: In small noise case, noise is taken to be much smaller than signal, i.e.,

$$n_i(t) << [A + x(t)] \quad ...(7.48)$$

A phasor-representation of envelope $A(t)$ is shown in figure 7.5. The noise component $n_c(t)$ is shown to be in phase with signal $[A + x(t)]$, whereas, $n_s(t)$ is in phase quadrature. Now, since $n_i(t) << [A + x(t)]$, the noise component $n_s(t)$ is also much smaller than $[A + x(t)]$.

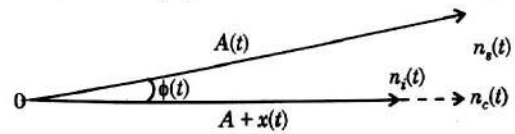


Fig. 7.5. *Phasor diagram of the envelope A(t)*

It is clear from the phasor diagram that if $n_s(t)$ is much smaller, then $\phi(t)$ is also much smaller and may be assumed to be zero. When $(t) = 0$, the envelope $A(t)$ becomes,

$$A(t) = A + x(t) + n_c(t) \quad ...(7.49)$$

Therefore, the output of the envelope detector contains a useful (modulating) signal $x(t)$ and a noise component $n_c(t)$. The carrier amplitude A carriers no useful information. The signal power S_o and the noise power N_o, at the output of the detector, may be calculated as follows :

Output Signal Power S_0: It is the ms value of useful signal $x(t)$,
Mathematically,

$$S_0 = \overline{X^2(t)} \quad ...(7.50)$$

Output Noise Power: The noise signal $n_0(t)$ at the output of the detector is $n_c(t)$ with the power spectrum density given as

$$S_{n0}(\omega) = S_{nc}(\omega) \quad ...(7.51)$$

As the AM system contains both sidebands, $S_{nc}(\omega)$ will be identical to that of DSB-SC system

Thus, $S_{no}(\omega) = S_{no}(\omega) = \eta$

The output noise power is given by

$$N_o = S_{no}(\omega)(BW) = \eta\,(2f_m) = 2\eta\, f_m \quad ...(7.52)$$

The figure of merit γ is obtained as.

$$\gamma = \frac{S_o/N_o}{S_i/N_i} = \frac{\dfrac{\overline{x^2(t)}}{(2\eta f_m)}}{\dfrac{1}{2}[A^2\overline{x^2(t)}]/\eta f_m}$$

or

$$\gamma = \frac{\overline{x^2(t)}}{A^2+\overline{x^2(t)}} \quad ...(7.53)$$

From equation (7.53), it is clear that the noise performance improves with a reduction in carrier amplitude A and it is maximum when $A = 0$, which is equivalent to the suppressed carrier AM system. Hence, the AM system remains always inferior to the suppressed system from the point of view of noise performance. The greatest value of γ that may be achieved depend on the minimum possible value of carrier amplitude A. The minimum value of A is equal to the maximum value of $x(t)$ to avoid over modulation in the envelope detector. Thus, the best noise performance is achieved when carrier amplitude A is equal to the maximum value of $x(t)$ *i.e.*, for 100% modulation.

At this point it may be noted that when we use the synchronous detection method for AM demodulation or detection, the noise performance is identical to the envelope detector method. In addition, the synchronous detection is complex and quite costly. Thus, it is never used for AM detection.

(ii) Large Noise Case: In large-noise this case, $n_i(t) >> [A + x(t)]$

Therefore quadrature components of $n_i(t)$ are also much larger than $[A+x(t)]$. Here, the noise terms dominate, and hence the performance of the envelope detector is completely different from the low noise case. The envelope of the modulated signal at the input of the detector is expressed

$$A(t) = \sqrt{[A+x(t)+n_c(t)]^2+n_s(t)^2}$$

$$= \sqrt{[A+x(t)]^2+2n_c(t)[A+x(t)]+n_c^2(t)+n_s^2(t)} \quad ...(7.54)$$

Now, since the noise dominates over the signal, the independent signal $[A + x(t)]^2$ is very small as compared to the other terms. Thus, $[A + x(t)]^2$ can be neglected, and the expression for $A(t)$ becomes

$$A(t) = \sqrt{n_c^2(t)+n_s^2(t)+2n_c(t)[A+x(t)]}$$

or

$$A(t) = \sqrt{[n_c^2(t)+n_s^2(t)]\left[1+\frac{2n_c(t)[A+x(t)]}{n_c^2(t)+n_s^2(t)}\right]}$$

Putting $R(t) = \sqrt{n_c^2(t)+n_s^2(t)}$

and $\theta(t) = \tan^{-1}\left[\dfrac{n_s(t)}{n_c(t)}\right]$

We get

$$A(t) = \sqrt{R^2(t)\left[1+2\{A+x(t)\}.\frac{n_c(t)}{R^2(t)}\right]} \quad ...(7.55)$$

$$= \sqrt{R^2(t)\left[1+\frac{2\{A+x(t)\}}{R(t)}\cos\theta(t)\right]} \quad ...(7.56)$$

$$= R(t)\left[1+\frac{2\{A+x(t)\}}{R(t)}\cos\theta(t)\right]^{\frac{1}{2}} \quad ...(7.57)$$

Since the noise component $R(t) >> 2[A + x(t)]$, equation (7.57) can be expressing in approximate form as

$$A(t) = R(t)\left[1+\frac{1}{2}\frac{2[A+x(t)]}{R(t)}\cos\theta(t)\right]$$

$$= R(t) + A\cos\theta(t) + x(t)\cos\theta(t) \quad ...(7.58)$$

It clear from equation (7.58). that envelope $A(t)$ which appears at the output of the envelope detector has no exclusive term proportional to the baseband (modulating) signal $x(t)$. The term $x(t)$ is multiplied with a large noise term $\cos\theta(t)$ and hence $x(t)$ cannot be separated from the noise. As the modulating term $x(t)$ is completely mingled with noise, it carries no useful information. The loss of the message $x(t)$ in an envelope detector due to the presence of the large noise is known as the ***threshold effect.***

Threshold Effect in an Envelope Detector

We have observed that in case, when a noise is large compared to the signal at the input of the envelope detector, the detected output has a message signal completely mingled with noise. It means that if the input signal-to-noise ratio (S_i/N_i) is below a certain level, called *threshold level*, the noise dominates over the message signal, *Threshold* is defined as value of the input *signal to noise ratio (S_o/N_o) below which the output signal to noise ratio (S_i/N_i) deteriorates much more rapidly than the input* signal to noise ratio. The threshold effect starts in an envelope detector whenever the carrier power-to-noise power ratio approaches unity or less. It is important to state that the threshold effect is the property of an envelope detector, on the other hand, *such an effect is not observed in a coherent detector.*

7.8. Evaluation of Noise in Angle Modulated Systems

In this article, let us discuss abcut quantitative calculations to determine the noise performance of angle modulated systems *i.e.*, FM and PM systems.

Figure 7.6 shows the block diagram of an angle modulated system.

We know that the generalised form of Angle Modulated carrier is represented as

$$s(t) = A\cos[\omega_c t + \phi(t)]$$

$$= A\cos[2\pi f_c t + \phi(t)] \quad ...(7.59)$$

where A_c is the amplitude of unmodulated carrier.

f_c is the frequency of unmodulated carrier and $\phi(t)$ is the instantaneous phase angle.

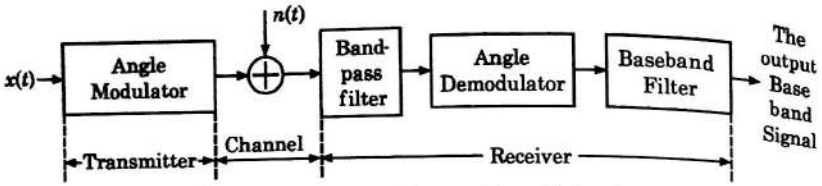


Fig. 7.6. *Block diagram of an angle modulated system.*

Further, instantaneous phase angle is given by

$$\phi(t) = k_p x(t) \text{ for phase modulation (PM)} \quad ...(7.60)$$

$$\phi(t) = 2\pi k_f \int_0^t x(t)\,dt \text{ for frequency modulation (FM)} \quad ...(7.61)$$

where $x(t)$ is the modulating or message signal band limited to f_m hertz.

k_p and k_f are the sensitivities of the phase and frequency modulators respectively.

Now channel noise $n_i(t)$ at the input of the demodulator is a bandpass noise with power spectral density (psd) $S_n(f)$ and band with $2(\Delta f + f_m)$, Hertz.

Thus, noise $n_i(t)$ may be expressed as

$$n_i(t) = n_c(t)\cos(\omega_c t) - n_s(t)\sin(\omega_c t)$$
$$= n_c(t)\cos(2\pi f_c t) - n_s(t)\sin(2\pi f_c t) \quad ...(7.62)$$

Equation (7.62) can be modified as

$$n_i(t) = r(t)\cos[\omega_c t + \theta(t)] = r(t)\cos[2\pi f_c t + \theta(t)] \quad ...(7.63)$$

where $$r(t) = \sqrt{n_c^2(t) + n_s^2(t)} \quad ...(7.64)$$

and $$\theta(t) = \tan^{-1}\left[\frac{n_s(t)}{n_c(t)}\right] \quad ...(7.65)$$

Hence, the in-phase and quadrature components of bandpass channel noise at demodulator input can be expressed as under:

$$n_c(t) = r(t)\cos[\theta(t)] \quad ...(7.66)$$

and $$n_s(t) = r(t)\sin[\theta(t)] \quad ...(7.67)$$

Now, let us obtain the Figure of Merit (FOM) for FM and PM systems in the subsections to follow:

7.9. Evaluation of Noise in Frequency Modulation (FM) Systems

The signal present at the FM demodulator input is given by

$$s_i(t) = s(t) + n_i(t) \quad ...(7.68)$$

Substituting equations (7.59) and (7.63) in equation (7.68), we obtain

$$s_i(t) = A\cos[2\pi f_c t + \phi(t)] + r(t)\cos[2\pi f_c t + \theta(t)] \quad \left[\text{Because } \phi(t) = 2\pi k_f \int_0^t x(t)\,dt \text{ for FM}\right]$$

or $$s_i(t) = A\cos\left[2\pi f_c t + 2\pi k_f \int_0^t x(t)\,dt\right] + r(t)\cos[2\pi f_c t + \theta(t)] \quad ...(7.69)$$

where $r(t)$ and $\theta(t)$ are the amplitude and phase of the bandpass channel noise $r(t)$ and $\theta(t)$ are given by equations (7.64) and (7.65), respectively. This may be represented by a phasor diagram shown in figure 7.7. In this phasor diagram, the carrier signal term is used as reference.

The relative phase $\psi(t)$ of the resulting phasor representing $s_i(t)$ may be obtained from the figure 7.7 as under:

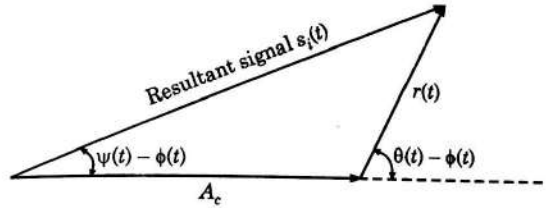


Fig. 7.7. *Illustration of phasor diagram of FM wave.*

$$\psi(t) = \phi(t) + \tan^{-1}\left\{\frac{r(t)\sin[\theta(t) - \phi(t)]}{A + r(t)\cos[\theta(t) - \phi(t)]}\right\} \quad ...(7.70)$$

where $$\phi(t) = 2\pi k_f \int_0^t x(t)\,dt \quad ...(7.71)$$

The envelope of resultant signal $s_i(t)$ is useful because FM demodulators use the instantaneous phase variation to determine the modulating signal $x(t)$.

For simplicity, let us assume that the amplitude of the unmodulated carrier is very large so that the **Carrier to Noise Ratio (CNR)** measured at the discriminator input is large compared to unity.

Using this condition, the relative phase $\psi(t)$ of the resultant phasor $s_i(t)$ may be simplified as under:

$$\psi(t) \cong \phi(t) + \frac{r(t)}{A}\sin[\theta(t) - \phi(t)]$$

Now, substituting $\phi(t) = 2\pi k_f \int_0^t x(t)\,dt$ in last equation, we have

$$\psi(t) \cong \phi(t) + \frac{r(t)}{A}\sin[\theta(t) - \phi(t)]$$

or $$\psi(t) \cong 2\pi k_f \int_0^t [x(t)\,dt] + \frac{r(t)}{A}\sin[\theta(t) - \phi(t)] \quad ...(7.72)$$

The discriminator output may be written as under:

$$s_d(t) = \frac{1}{2\pi}\frac{d\psi(t)}{dt} \quad ...(7.73)$$

Substituting equation (7.72) in equation (7.73), we have

$$s_d(t) \cong \frac{1}{2\pi}\frac{d}{dt}\left\{2\pi k_f \int_0^t [x(t)dt] + \frac{r(t)}{A}\sin[\theta(t) - \phi t)]\right\}$$

or $$s_d(t) \cong \frac{1}{2\pi}\frac{d}{dt}\left\{2\pi k_f \int_0^t [x(t)dt] + \frac{1}{2\pi}\frac{d}{dt}\left\{\frac{r(t)}{A}\sin[\theta(t)-\phi(t)]\right\}\right\}$$

or $$s_d(t) \cong k_f x(t) + \frac{1}{2\pi A}\frac{d}{dt}\{r(t)\sin[\theta(t)-\phi(t)]\}$$

or $$s_d(t) \cong k_f x(t) + n_d(t) \quad ...(7.74)$$

where $n_d(t)$ is the noise term and it is given by

$$n_d(t) = \frac{1}{2\pi A}\frac{d}{dt}\{r(t)\sin[\theta(t)-\phi(t)]\} \quad ...(7.75)$$

From equation $s_d(t) \cong k_f x(t) + n_d(t)$, it may be observed that the Carrier-to-Noise ratio [CNR] is high, the discriminator output consists of the original message signal $x(t)$ multiplied by the frequency sensitivity k_f plus additive noise component $n_d(t)$. There is further need of simplification of $n_d(t)$ before we actually calculate the output signal-to-noise ratio (SNR).

Now, from the phasor diagram of figure 7.7, it may be noted that the effect of variation in the phasor $\theta(t)$ of the narrowband noise appears known as the signal term $\phi(t)$. The phase $\theta(t)$ is uniformly distributed over 2π radians. Therefore, we can assume that the phase difference $[\theta(t) - \psi(t)]$ is also uniformly distributed over 2π radians. If phase difference $[\theta(t) - \psi(t)]$ is uniformly distributed over 2π radians, then the noise $n_d(t)$ at the discriminator output will be independent of the modulating signal $x(t)$. However, it will only depend upon the characteristics of the carrier and narrowband noise.

Above assumption is true only if the **Carrier-to-Noise ratio (CNR)** is high.

However, above assumption further simplifies equation (7.75) to the following expression:

$$n_d(t) = \frac{1}{2\pi A}\frac{d}{dt}\{r(t)\sin[\theta(t)]\} \quad ...(7.76)$$

Also from equation (7.67), we have

$$r(t)\sin[\theta(t)] = n_s(t) \quad ...(7.77)$$

Now, putting equation (7.77) in equation (7.76), we have

$$n_d(t) \cong \frac{1}{2\pi A}\frac{d}{dt}\{n_s(t)\}$$

Hence, $n_d(t)$ may be determined by A and $n_s(t)$. Here, $n_d(t)$ is the additive noise which appears at the discriminator output and $n_s(t)$ is the quadrature component of the narrowband noise $n(t)$.

When the discriminator output signal $s_d(t) \cong k_f x(t) + n_d(t)$ is made to pass through the low pass filter (LPF), the final demodulator signal output is given by

$$s_0(t) = k_f x(t) \quad ...(7.78)$$

Hence, mean square signal power is found as

$$P_{so} = k_f^2 \overline{x^2(t)} \quad ...(7.79)$$

Now, to determine the average output noise power, we note that the noise $n_d(t)$ at the discriminator output is directly proportional to the first derivation of the quadrature component $n_s(t)$. It means that $n_d(t)$ may be obtained by passing $n_s(t)$ through a differentiator. We know from the property of Fourier transform that the differentiation of a function w.r.t. time "t" corresponds to multiplication of its Fourier transform by $j2\pi f$. Hence, the noise $n_d(t)$ can be obtained by passing $n_s(t)$ through a linear filter whose transfer function is given as under

$$H(f) = \frac{j2\pi f}{2\pi A} = \frac{if}{A} \quad ...(7.80)$$

Further, let us consider that $S_{ns}(f)$ and $S_{nd}(f)$ denote power spectral density (psd) of the quadrature component of bandpass noise and the noise at the output of the demodulator, respectively.

Then, we get $$S_{nd}(f) = |H(f)|^2 S_{ns}(f) \quad ...(7.81)$$

Putting equation (7.80) in equation (7.81), we have

$$S_{nd}(f) = \left|\frac{jf}{A}\right|^2 S_{ns}(f)$$

or $$S_{nd}(f) = \left|\frac{j^2 f^2}{A^2}\right| S_{ns}(f) = \left|\frac{-f^2}{A^2}\right| S_{ns}(f)$$

or $$S_{nd}(f) = \frac{f^2}{A^2} S_{ns}(f) \quad ...(7.82)$$

If the bandpass filter at the input of the FM receiver is assumed to have an ideal bandpass characteristic of bandwidth B_T and midband frequency f_c, then the narrowband noise $n(t)$ would have a similar power spectral characteristic. In other words, we can say that the quadrature component $n_s(t)$ of the narrow and noise $n(t)$ will have ideal low pass characteristic. These characteristic have been shown in figure 7.8(a). The corresponding power spectral density (psd) of the noise $n_d(t)$ can be obtained from equation (7.82) as below:

$$S_{nd}(f) = \begin{cases} \dfrac{\eta f^2}{A^2} & \text{for } |f| \le \dfrac{B_T}{2} \\ 0, & \text{elsewhere} \end{cases} \quad ...(7.83)$$

where η is the value of psd of $n_s(t)$. It has been assumed to be constant over the range $-\frac{B_T}{2} \le f \le \frac{B_T}{2}$.

The variation of $S_{nd}(f)$ with frequency f has been shown in figure 7.8b.

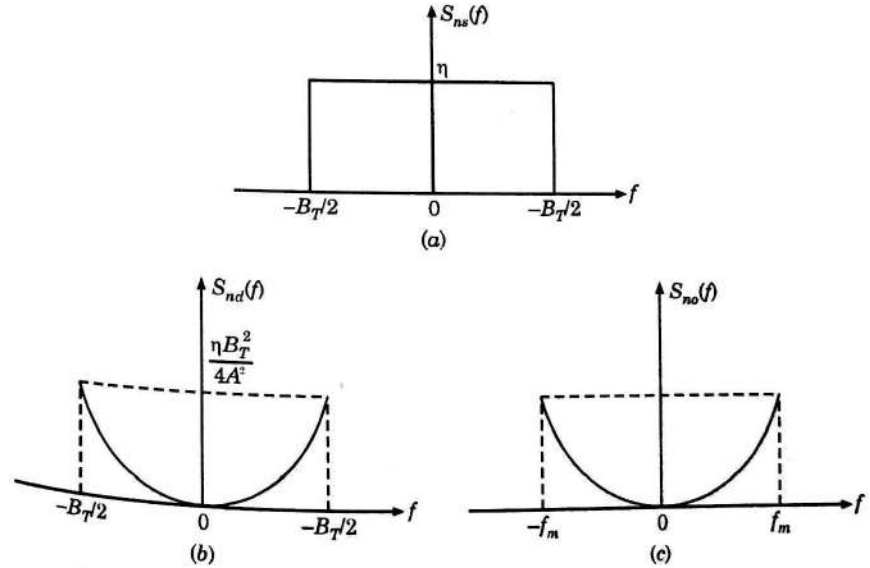


Fig. 7.8. *Power spectral density (psd) of (a) $n_s(t)$, (b) $n_d(t)$, (c) $n_0(t)$*

The discriminator output is followed by a low pass filter (LPF) with bandwidth equal to the message bandwidth f_m.

We know that for wideband FM, f_m is smaller than $B_T/2$, where B_T is transmission bandwidth of FM. This means that the output-of-band components of noise $n_d(t)$ will be rejected by using low pass filter (LPF). Therefore, the PSD $S_{no}(f)$ of the noise $n_o(t)$ appearing at the receiver output is given as

$$S_{no}(f) = \begin{cases} \dfrac{\eta f^2}{A^2} & \text{for } |f| \le f_m \\ 0, & \text{elswhere} \end{cases} \qquad ...(7.84)$$

Plot of psd of $n_o(t)$, *i.e.*, $S_{no}(f)$ versus frequency has been shown in figure 7.8(c).

The average output noise power is obtained by integrating the psd $S_{no}(f)$

$$P_{no} = \int_{-\infty}^{\infty} S_{no}(f)\, df \qquad ...(7.85)$$

Putting equation (7.84) in equation (7.85), we have

$$P_{no} = \int_{-f_m}^{f_m} \frac{\eta f^2}{A^2}\, df = \frac{\eta}{A^2}\int_{-f_m}^{f_m} f^2 df = \frac{\eta}{A^2}\left[\frac{f^3}{3}\right]_{-f_m}^{f_m}$$

or

$$P_{no} = \frac{2\eta f_m^{\,3}}{3A^2} = \frac{2\eta f_m^3}{3A^2} \qquad ...(7.86)$$

Now, it may be observed from equation (7.86) that the output noise power P_{no} is inversely proportional to the average carrier power $P_c = A^2/2$.

Hence, in an FM system, increasing the carrier power has a **noise-quite effect.**

The average signal power at demodulator input is given by

$$P_{si} = \frac{A^2}{2} \qquad ...(7.87)$$

Now, the average noise power at the demodulator input is given by

$$P_{ni} = 2\eta B_T \qquad ...(7.88)$$

Therefoe, Figure of Merit (FOM) of FM system is given by

$$[\text{FOM}]_{\text{FM}} = \frac{(SNR)_o}{(SNR)_i} \qquad ...(7.89)$$

The Signal-to-Noise Ratio (SNR) at the FM demodulator or receiver input is the ratio of signal power to noise power at the demodulator input, *i.e.*,

$$(SNR)_i = \frac{\text{Average signal power}}{\text{Average noise power}} = \frac{P_{si}}{P_{ni}} \qquad ...(7.90)$$

Putting equations (7.87) and (7.88) in equation (7.90), we get

$$(SNR)_i = \frac{P_{si}}{P_{ni}} = \frac{(A^2/2)}{2\eta B_T} = \frac{A^2}{4\eta B_T} \qquad ...(7.91)$$

The Signal-to-Noise Ratio (SNR) at the FM demodulator output is the ratio of signal power to noise power at demodulator output *i.e.*,

$$(SNR)_0 = \frac{P_{so}}{P_{no}} \qquad ...(7.92)$$

Putting equations (7.79) and (7.85) in equation (7.92), we obtain

$$(SNR)_o = \frac{P_{so}}{P_{no}} = \frac{\left[k_f^2\, \overline{x^2(t)}\right]}{\left[\dfrac{2\eta f_m^{\,3}}{3A^2}\right]}$$

Simplifying, we have

$$(SNR)_o = \frac{3k_f^2 A^2\, \overline{x^2(t)}}{2\eta f_m^3} \qquad ...(7.93)$$

or

Substituting equations (7.91) and (7.93) in equation (7.89), we get

$$[FOM]_{FM} = \frac{(SNR)_o}{(SNR)_i} = \frac{\left[\dfrac{3k_f^2 A^2\, \overline{x^2(t)}}{2\eta f_m^3}\right]}{\left[\dfrac{A^2}{4\eta B_T}\right]}$$

Again, simplifying, we get

$$[FOM]_{FM} = \frac{6kf^2\, \overline{x^2(t)}\, B_T}{f_m^3} \qquad ...(7.94)$$

We know that the Signal-to-Noise Ratio (SNR) at the FM demodulator output is given by

$$[(SNR)_o]_{FM} = \frac{3k_f^2 A^2\, \overline{x^2(t)}}{2\eta f_m^3} \qquad ...(7.95)$$

The carrier frequency deviation in FM is directly proportional to frequency sensitivity k_f. Therefore, the bandwidth of FM is directly proportional to K_f.

Now let us compare signal-to-noise ratio (SNR) at the demodulator output for FM and AM. If the signal $x(t)$ were transmitted using AM, the signal output power of the envelope detector may be given by

$$P_{so} = k_a^2 A^2\, \overline{x^2(t)} \qquad ...(7.96)$$

and the noise output power of the envelope detector may be given by

$$P_{no} = \overline{n_c^2(t)} = 2\eta f_m \qquad ...(7.97)$$

Note that here we have assumed that the input noise is white. Its uniform power spectral density (psd) is $\eta/2$ and $2f_m$ is the bandwidth of transmission of AM signal.

Therefore, $(SNR)_0$ for AM is determined by using equations (7.96) and (7.97) as under

$$[(SNR)_o]_{AM} = \frac{P_{so}}{P_{no}} = \frac{k_a^2 A^2\, \overline{x^2(t)}}{2\eta f_m} \qquad ...(7.98)$$

Further, the ratio of $[(SNR)_o]_{FM}$ to $[(SNR)_o]_{AM}$ can be obtained as follows:

$$\frac{[(SNR)_o]_{FM}}{[(SNR)_o]_{AM}} = \frac{\left(\dfrac{3k_f^2 A^2\, \overline{x^2(t)}}{2\eta f_m^3}\right)}{\left(\dfrac{k_a^2 A^2\, \overline{x^2(t)}}{2\eta f_m}\right)} = \frac{3k_f^2}{f_m^2 k_a^2} \qquad ...(7.99)$$

For comparison, we also consider AM under most favourable condition *i.e.*, AM with 100% modulation. In this case, the amplitude of $x(t)$ becomes same as that of the carrier *i.e.*, A.

Therefore, we have

$$k_a = \frac{1}{A} = \frac{1}{A_m}$$

where A_m is the amplitude of the message or baseband signal.

In such a condition, equation (7.99) may be written as

$$\frac{[(SNR)_o]_{FM}}{[(SNR)_o]_{AM}} = \frac{3k_f^2}{f_m^2 k_a^2} = \frac{3k_f^2}{f_m^2\left(\frac{1}{A}\right)^2} = \frac{3k_f^2 A^2}{f_m^2}$$

or $$\frac{[(SNR)_o]_{FM}}{[(SNR)_o]_{AM}} = \frac{3k_f^2 A_m^2}{f_m^2} = 3\left(\frac{k_f A_m}{f_m}\right)^2$$

or $$\frac{[(SNR)_o]_{FM}}{[(SNR)_o]_{AM}} = 3\left(\frac{k_f A_m}{f_m}\right)^2 \quad ...(7.100)$$

where f_m has been assumed to be the frequency of the sinusoidal modulating signal. For single-tone sinusoidal modulation, the frequency deviation Δf of the FM signal caused by modulating signal $x(t)$ of amplitude $A_m = A$.

In this case, frequency deviation is given as

$$\Delta f = k_f A_m = k_f A \quad ...(7.101)$$

Hence, equation (7.100) may be written as

$$\frac{[(SNR)_o]_{FM}}{[(SNR)_o]_{AM}} = 3\left(\frac{\Delta f}{f_m}\right)^2 = 3m_f^2 \quad ...(7.102)$$

where m_f is the modulating index of FM. It is defined as under

$$m_f = \frac{\Delta f}{f_m} = \frac{k_f A_m}{f_m} = \frac{k_f A}{f_m} \quad ...(7.103)$$

Further, if we consider the root mean square (rms) value of Signal-to-Noise voltage ratio, rather power ratio, we get

$$\frac{[(SNR)_o]_{FM,V}}{[(SNR)_o]_{AM,V}} = \sqrt{3m_f^2} = \sqrt{3}\, m_f \quad ...(7.103a)$$

It may be observed form equation (7.103*a*) that the SNR can be made much higher in FM than in AM by simply increasing the modulation index m_f. We also know that an increase in m_f value results in an increase in the bandwidth required for transmission.

For example: If $m_f = 5$ then, we have

$$\frac{[(SNR)_o]_{FM}}{[(SNR)_o]_{AM}} = 3m_f^2 = 3 \times (5) = 75$$

or $$[(SNR)_o]FM = 75[(SNR)_o]_{AM} \quad ...(7.104)$$

For modulation index $m_f = 5$, the transmission bandwidth of FM would be equal to

$$B_T \cong 3.3\Delta f \text{ (from universal bandwidth curve)}$$

$$B_T \cong 3.3 m_f \cdot f_m \cong 3.3 \times 5\, f_m \cong 16.5\, f_m \quad ...(7.105)$$

when compared with the AM bandwidth ($= 2f_m$), to achieve an improvement of SNR in FM by a factor of 75 in comparison with AM, the bandwidth of FM has to be increased approximately by a factor of 7.

For large value of modulation index m_f ($m_f > 10$), the bandwidth B_T of the FM signal is given by

$$B_T \cong 2\Delta f \quad ...(7.106)$$

Hence, Modulation Index of FM can be written as

$$m_f = \frac{\text{Frequency deviation}}{\text{Modulating frequency}} = \frac{\Delta f}{f_m} = \frac{B_T}{2f_m} \quad ...(7.107)$$

Putting equation (7.107) in equation (7.102), we have

$$\frac{[(SNR)_o]_{FM}}{[(SNR)_o]_{AM}} = 3m_f^2 = 3\left(\frac{B_T}{2f_m}\right)^2$$

$$\frac{[(SNR)_o]_{FM}}{[(SNR)_o]_{AM}} = 3\frac{B_T^2}{4f_m^2} = \frac{3}{4}\frac{B_T^2}{4f_m^2} = \frac{3}{4}\left(\frac{B_T}{f_m}\right)^2 \quad ...(7.108)$$

Equation (7.108) may be written as under

$$[(SNR)_o]_{FM} = \frac{3}{4}\left(\frac{B_T}{f_m}\right)^2 [(SNR)_o]_{AM} \quad ...(7.109)$$

Note: *It may be concluded from equation (7.109) that if we want to improve $(SNR)_o$ in FM system by a factor of 4 (i.e., equal to 6 dB), we should increase the bandwidth by a factor of 2. Hence, the signal-to-noise power ratio improves by 6 dB for each two to one increase in the bandwidth of FM transmission. This result is valid for a large value of modulation index m_f and under small noise condition. Here, we have observed that an FM system allows us to exchange the SNR for the bandwidth of transmission. However, such type of exchange is not possible in AM system where the bandwidth of transmission is fixed ($= 2f_m$).*

It may be noted from equation (7.103*a*) that in order to realise an improvement in SNR ratio of FM system over AM system, the following condition must be satisfied:

$$m_f > \frac{1}{\sqrt{3}} \cong 0.6 \quad ...(7.110)$$

For a value of modulation index m_f equal to or less than above given value, there will be no improvement in SNR ratio of FM over AM. The value of $m_f = 0.6$ is actually the transition point between the Narrowband Frequency Modulation (NBFM) and Wideband Frequency modulation (WBFM). In such a way, we can say that NBFM, therefore, provides no improvement over AM. The result, however, is obvious since the NBFM and AM have the same bandwidth.

One more important point may be noted about FM that the SNR in FM cannot be improved indefinitely by increasing bandwidth. As the bandwidth increase the input noise also increases and ultimately a point is reached where the carrier power P_c will become of the order of the input noise power.

Therefore, the results derived above under the assumption of small noise condition do not hold true any longer. Under low-noise condition, the phenomenon of threshold occurs.

7.10, Noise in Phase Modulation (PM)

In this article, let us discuss about the noise performance of PM system. For wideband modulation, the modulating signal $x(t)$ changes very slowly relative to the noise $n_i(t)$. The signal bandwidth is

f_m and the noise bandwidth (or transmission bandwidth is $B_T = 2(\Delta f + f_m)$, with $\Delta f >> f_m$. Therefore, the phase and frequency variations of the modulated carrier appears to have constant frequency and phase over several cycles, and hence the carrier appears to be unmodulated. We may therefore, determine the output noise by assuming $x(t)$ to be zero.

To determine the output signal power and the noise power, we first draw the phasor diagram of the signal $s_i(t)$ at the demodulator input.

It is given by

$$s_i(t) = A\cos[\omega_c t + \phi(t)] + n_i(t)$$

Because $n_i(t) = r(t)\cos[\omega_c t + \phi(t)]$ and $\omega_c = 2\pi f_c$

Therefore, we write

$$s_i(t) = A\cos[2\pi f_c t + \phi(t)] + r(t)\cos[2\pi f_c t + \phi(t)]$$

or

$$s_i(t) = e(t)\cos[2\pi f_c t + \phi(t) + \Delta\phi(t)] \quad ...(7.111)$$

where

$$\phi(t) = k_p x(t) \quad ...(7.112)$$

Figure 7.9 shows the phasor diagram of $s_i(t)$ given in equation (7.11).

We already know that for small noise case where

$$r(t) << A \text{ for all } t$$

and

$$\Delta\phi(t) << \frac{\pi}{2} \text{ for all } t.$$

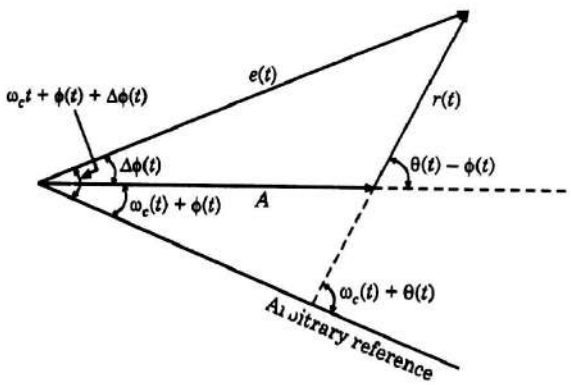


Fig. 7.9. *Illustration of phasor diagram of phase modulated (PM) signal for small signal case.*

We have

$$\Delta\phi(t) \cong \frac{r(t)}{A}\sin[\theta(t) - \phi(t)] \quad ...(7.113)$$

The PM demodulator detects the phase of the input $s_i(t)$. Therefore, output of the phase demodulator is given by

$$s_0(t) = \phi(t) + \Delta\phi(t) \quad ...(7.114)$$

Putting equation (7.113) in equation (7.114), we obtain

$$s_o(t) = \phi(t) + \frac{r(t)}{A}\sin[\theta(t) - \phi(t)] \quad ...(7.115)$$

From equation (7.113), we know that the noise term $\Delta\phi(t)$ involves the $\phi(t)$ because of non-linear nature of angle modulation. Since $\Delta\phi(t)$ varies much more slowly than $\theta(t)$, so we can approximate $\phi(t)$, by a constant ϕ.

Thus, equation (7.113) may be written as

$$\Delta\phi(t) \cong \frac{r(t)}{A}\sin[\theta(t) - \phi]$$

or

$$\Delta\phi(t) \cong \frac{r(t)}{A}\{\sin[\theta(t)]\cos\phi - \cos[\theta(t)]\sin\phi\}$$

or

$$\Delta\phi(t) \cong \frac{r(t)}{A}\sin[\theta(t)]\cos\phi + \frac{r(t)}{A}\cos[\theta(t)]\sin\phi$$

or

$$\Delta\phi(t) = \frac{n_s(t)}{A}\cos\phi + \frac{n_c(t)}{A}\sin\phi \quad ...(7.116)$$

Because $n_c(t)$ and $n_s(t)$ are incoherent for white noise, therefore, we have

$$S_{\Delta\phi}(f) = \frac{\cos^2\phi}{A^2}S_{ns}(f) + \frac{\sin^2\phi}{A^2}S_{nc}(f) = \frac{S_{ns}(f)}{A^2} \quad ...(7.117)$$

For white channel noise with power density spectrum $\eta/2$, equation (7.117) can be written as under:

$$S_{\Delta\phi}(f) = \begin{cases} \dfrac{\eta}{A^2} & \text{for } |f| < B_T \\ 0, & \text{elsewhere} \end{cases} \quad ...(7.118)$$

Because, we know that the useful signal bandwidth is only f_m Hertz, therefore, the demodulator output is passed through a low pass filter (LPF) of bandwidth f_m. It removes the output of band noise. Therefore, the power spectral density (psd) at the output of low pass filter is given by

$$S_{no}(f) = \begin{cases} \dfrac{\eta}{A^2} \text{ for } & |f| < f_m \\ 0 \text{ for } & |f| > f_m \end{cases} \quad ...(7.119)$$

The mean square noise power at the demodulator output is given by

$$P_{no} = 2f_m\left(\frac{\eta}{A^2}\right) = \frac{2\eta f_m}{A^2} \quad ...(7.120)$$

The mean square signal power at the demodulator output is given by

$$P_{so} = k_p^2\,\overline{x^2(t)} \quad ...(7.121)$$

The signal-to-noise ratio (SNR) at demodulator output is given by

$$(SNR)_o = \frac{\text{Mean square signal power at the demodulator output}}{\text{Mean square noise power at the demodulator output}}$$

or

$$(SNR)_o = \frac{P_{so}}{P_{no}} \quad ...(7.122)$$

Putting equations (7.120) and (7.121) in equation (7.122), we get

$$(SNR)_o = \frac{P_{so}}{P_{no}} = \frac{k_p^2\,\overline{x^2(t)}}{\dfrac{2\eta f_m}{A^2}} = \frac{k_p^2\,\overline{x^2(t)A^2}}{2\eta f_m}$$

or

$$(SNR)_o = \frac{(A\,k_p)^2\,\overline{x^2(t)}}{2\eta f_m} \quad ...(7.123)$$

Also, we have already determined the signal power at the demodulator input which is given by

$$P_{si} = \frac{A^2}{2}$$

and noise power at the demodulator input is given by ...(7.124)

$$P_{ni} = 2\eta B_T$$

The signal-to-noise ratio (SNR) at demodulator input is given by ...(7.125)

$$(SNR)_i = \frac{P_{si}}{P_{ni}}$$ ...(7.126)

Putting equations (7.124) and (7.125) in equation (7.126), we have

$$(SNR)_i = \frac{P_{si}}{P_{ni}} = \frac{\frac{A^2}{2}}{2\eta B_T}$$

or $$(SNR)_i = \frac{A^2}{4\eta B_T}$$ ...(7.127)

The figure of Merit (FOM) of PM demodulator is defined as below:

$$[FOM]_{PM} = \frac{(SNR)_o}{(SNR)_i}$$ ...(7.128)

Substituting equations (7.123) and (7.127) in equation (7.128), we have

$$[FOM]_{PM} = \frac{(SNR)_o}{(SNR)_i} = \frac{\left[\frac{(A\,k_p)^2\,\overline{x^2(t)}}{2\eta f_m}\right]}{\left[\frac{A^2}{4\eta B_T}\right]}$$

Simplifying, we have

$$= \frac{2k_p^2\,\overline{x^2(t)}\,B_T}{f_m}$$ ...(7.129)

But we know that

$$[(SNR)_o]_{FM} = \frac{3k_f^2 A^2\,\overline{x^2(t)}}{2\eta f_m^{\,3}}$$ ...(7.130)

From equation (7.123), we know that

$$[(SNR)_o]_{PM} = \frac{A^2 k_p^2\,\overline{x^2(t)}}{2\eta f_m}$$ ...(7.131)

Dividing equation (7.131) by equation (7.130), we get

$$\frac{[(SNR)_o]_{PM}}{[(SNR)_i]_{FM}} = \frac{\left(\frac{A^2 k_p^2\,\overline{x^2(t)}}{2\eta f_m}\right)}{\left(\frac{3k_f^2 A^2\,\overline{x^2(t)}}{2\eta f_m^3}\right)}$$

Simplifying, we get

$$\frac{[(SNR)_o]_{PM}}{[(SNR)_i]_{FM}} = \frac{k_p^2\,f_m^{\,2}}{6k_f^2}$$ ...(7.132)

Also, we also know that for 100% amplitude modulation, we have

$$[(SNR)_o]_{AM} = \frac{k_a^2 A^2\,\overline{x^2(t)}}{2\eta f_m}$$ ...(7.133)

Now, dividing equation (7.131) by equation (7.133), we obtain

$$\frac{[(SNR)_o]_{PM}}{[(SNR)_i]_{AM}} = \frac{\left(\frac{A^2 k_p^2\,\overline{x^2(t)}}{2\eta f_m}\right)}{\left(\frac{k_f^2 A^2\,\overline{x^2(t)}}{2\eta f_m}\right)} = \frac{(k_p)^2}{(k_a)^2}$$ ...(7.134)

For single-tone sinusoidal case of 100% modulation, we have

$$x(t) = A\cos(\omega_c t) = A\cos(2\pi f_c t)$$

Phase modulated (PM) carrier is defined as

$$s_i(t) = A\,[\omega_c t + A\,k_p\cos(\omega_m t)]$$
$$= A\,[2\pi f_c t + A\,k_p\cos(2\pi f_m t)]$$ ...(7.135)

Hence, we have $$f_i = f_c - \frac{A\,k_p}{2\pi}\,2\pi f_m\sin(2\pi f_m t)$$

$$f_i = f_c - A\,k_p f_m\sin(2\pi f_m t) = f_c - \Delta f\sin(2\pi f_m t)$$ ...(7.136)

where $$\Delta f = A\,k_p f_m$$ ...(7.137)

In the case of 100% modulation, the amplitude of modulating signal $x(t)$ becomes same as that of the carrier *i.e.*, A.

Thus, $$k_a = \frac{m}{A_m} = \frac{1}{A_m} = \frac{1}{A}$$ ...(7.138)

7.11. Noise in Pulse-Modulated Systems

In pulse-modulated systems, we shall discuss pulse amplitude modulation (PAM), pulse position modulation (PPM), and pulse code modulation (PCM) from noise-analysis point of view.

7.12. Noise in Pulse Amplitude Modulation (PAM)

As a matter of fact, the Signal-to-noise behaviour of PAM systems is identical to that of AM/SC systems. The S/N improvement in PAM is unity, the same as that of SSB-SC. Similarly, the S/N improvement ratio in PAM/AM is 2, the same as that of DSB-SC. The result is not surprising since the bandwidth requirements of PAM and PAM/AM are identical to SSB and DSB, respectively.

In PAM, the signal $x(t)$ is sampled. The sampled signal $x_s(t)$ is then transmitted through a low-pass filter* which yields the original continuous signal. So, in effect, we are transmitting $x(t)$ directly. In the process of transmission, it is corrupted by the noise. At the receiver, a low-pass

* Actually, there are M message whose samples are interleaved for the purpose of time division multiplexing. These combined samples are then transmitted through a low-pass filter of cut-off frequency Mf_m. Here, for convenience, we shall consider the case of a single signal only.

filter is used. It is evident that the signal-to-noise ratio in this case remains unchanged.

Thus, we have

$$\frac{S_o/N_o}{S_i/N_i} = 1 \qquad ...(7.139)$$

It can be seen that

$$S_o = S_i = \overline{x^2(t)}$$

and $$N_o = N_i = Nf_m \quad \text{(for white noise)}$$

If PAM is transmitted by PAM-AM, it can be readily shown that the signal-to-noise improvement ration is 2 (same as DSB-SC).

7.13. Noise in Pulse Position Modulation (PPM)

In PPM, a message signal $x(t)$ is sampled and the sample values are transmitted in the form of pulse positions. The kth pulse is shifted from its quiescent position by an amount proportional of $x(kT)$, the kth sample of $x(t)$. The pulses are transmitted over a channel of bandwidth B. This causes the dispersion of the pulse. The received pulse can be approximated with fair accuracy by a trapezoidal pulse. This has been shown in figure 7.10

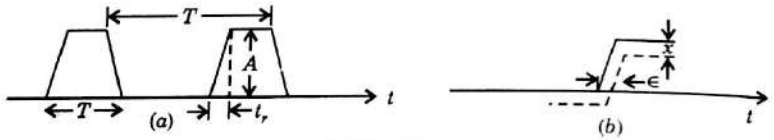


Fig. 7.10.

The rise time t_r of a pulse is given by

$$t_r \cong \frac{1}{B} \qquad ...(7.140)$$

where B is the bandwidth (in Hz) of the channel. The position of the trapezoidal pulse is sensitive to the additive channel noise. This can be seen from figure 7.10.

An addition of a constant signal of magnitude x will shift the pulse position by ϵ, where

$$\frac{\epsilon}{x} = \frac{t_r}{A} \qquad ...(7.141)$$

If x varies randomly, then ϵ will also vary randomly. Since ϵ is proportional to x, the root mean square value of ϵ will be proportional to the root mean square value of x.

From equation (7.141) it follows that

$$\frac{\overline{\epsilon^2}}{\overline{x^2}} = \left(\frac{t_r}{A}\right)^2$$

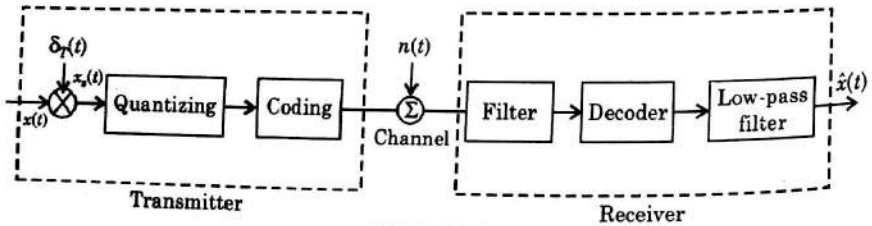


Fig. 7.11.

Thus if a noise signal $n(t)$ with mean square value $\overline{n^2(t)}$ is added to PPM signal, the mean square value of ϵ is given by

$$\overline{\epsilon^2} = \left(\frac{t_r}{A}\right)^2 \overline{n^2(t)} \qquad ...(7.142)$$

In the absence of noise, the pulse position of the kth pulse varies in proportion of $x(kT)$, the kth sample of $x(t)$. If we denote the change in the position of the kth pulse by y_k, then, we have

$$y_k = k_1 x(kT)$$

where k_1 is a constant of proportionality of the modulator.

However, because of the channel noise, the change in the pulse position is not y_k but $y_k + \epsilon_k$, where ϵ_k is random and has a mean square value given by equation (7.142).

Let $$\hat{y}_k = y_k + \epsilon_k = k_1 x(kT) + \epsilon_k \qquad ...(7.143)$$

At the receiver, the pulse positions are converted back into samples. The kth sample at the receiver is $\hat{y}_k$ (equation 7.143).

To obtain the original message $x(t)$, we transmit the samples through a low-pass filter of cut-off frequency ω_m.

The output of the low-pass filter is $\hat{y}(t)$, given by

$$\hat{y}(t) = \sum_k \hat{y}_k \, Sa(\omega_m t - k\pi) = \sum_k [k_1 x(kT) + \epsilon_k] \, Sa(\omega_m t - k\pi)$$

or $$= \sum_k k_1 x(kT) \, Sa(\omega_m t - k\pi) + \sum_k \epsilon_k Sa(\omega_m t - k\pi)$$

Each of the summations represent a signal bandlimited to ω_m radians per second. Since $x(kT)$ represents the sample of $x(t)$, it is evident that the first summation is equal to $k_1 x(t)$. The second summation represents a noise signal $\epsilon(t)$.

Thus, the output $\hat{y}(t)$ is given by

$$\hat{y}(t) = k_1 x(t) + \epsilon(t)$$

Hence, $$s_o(t) = k_1 x(t)$$

$$S_o = k_1^2 \overline{x^2(t)}$$

$$n_0(t) = \epsilon(t)$$

and $$N_o = \overline{n_0^2(t)} = \overline{\epsilon^2(t)}$$

Now, the mean square value of a signal is equal to the mean square value of its samples. Thus $\overline{\epsilon^2(t)}$, the mean square value of $\epsilon(t)$, is equal to $\overline{\epsilon_k^2}$. But the mean square value of ϵ_k is given by equation (7.142).

Hence $$\overline{\epsilon^2(t)} = \overline{\epsilon_k^2} = \left(\frac{t_r}{A}\right)^2 \overline{n^2(t)}$$

Thus $$N_o = \left(\frac{t_r}{A}\right)^2 \overline{n^2(t)}$$

and $$\frac{S_o}{N_o} = \frac{k_1^2}{\overline{n^2(t)}} \left(\frac{A}{t_r}\right)^2 \overline{n^2(t)} = k_1^2 \frac{A^2}{\overline{n^2(t)}} \overline{x^2(t)} \, B^2$$

where B is the channel bandwidth (in Hz).

If the ratio of the pulse width τ to the period T is α, then the input signal power, S_i is given by (assuming $t_r < T$)

$$S_i = \alpha A^2$$

and the input noise power N_i is given by

$$N_i = \overline{n^2(t)}$$

Hence,
$$\frac{S_o/N_o}{S_i/N_i} = \frac{k_1^2}{\alpha}\overline{x^2(t)}\,B^2 \qquad ...(7.144)$$

If we consider signal-to-noise root mean square voltage ratios rather than power ratios, we obtain

$$\frac{(S_o/N_o)_{VT}}{(S_i/N_i)_{VT}} = k_1\left(\frac{\overline{x^2(t)}}{x}\right)^{1/2} B \qquad ...(7.145)$$

Thus, the S/N power ratio improvement is proportional to the square of the bandwidth. The behaviour is identical to that of wideband FM. In both these systems, it is possible to trade bandwidth for the signal-to-noise ratio. The S/N voltage ratio is linearly proportional to the bandwidth B. This is characteristic of encoded systems. We shall now show that in the coded systems such as PCM, the S/N voltage ratio can be increased exponentially with bandwidth. Thus, the coded systems are inherently capable of better transmission efficiency than the uncoded systems.

7.14. Coded Systems

The modulation systems can be boadly classified into two classes as under:

(*i*) uncoded systems, and

(*ii*) coded systems

In uncoded systems one symbol in message space is transformed into one symbol in the modulated signal space. Thus, for AM each possible amplitude of the original message is transformed into a particular amplitude of the modulated signal. Similarly, for PPM each amplitude of the message results in a particular displacement of the pulse position. In coded systems (such as PCM), however, each message symbol or amplitude is transformed into a number of signal symbols. It has been shown that only coded systems can achieve the most efficient exchange bandwidth for signal to noise ratio that is theoretically possible. The of uncoded systems (such as FM or PPM) are inherently incapable of efficient exchange of the bandwidth and signal-to-noise ratio. We shall now consider PCM from the point of view of exchange of bandwidth and signal-to-noise ratio.

7.15. Noise in Binary PCM

In PCM, the modulating signal to sampled. Each of the samples is transmitted by a code formed by a group of pulses. In general, the samples can assume any value in a certain continuous range of the amplitude. However, we allow the samples to take any one of the finite number of levels in this range by approximating each sample to the nearest allowable level. This approximation, known as quantization, obviously is a source of error. The signal that is transmitted is an approximation and there is an uncertainty in the signal amplitude of the amount equal to the difference between the allows levels. The error or the uncertainty can be made as small as possible by increasing the number of levels in a given amplitude range.

Now, let us consider the case of a binary pulse code where the pulses can assume amplitudes of 0 and A volts only. Also, a group of 4 binary pulses are required to represent a sample that can assume any one of the 16 possible values. This is because 4 binary pulses can form $2 \times 2 \times 2 \times 2 = 16$ distinct patterns. Similarly, we need a group of 5 binary pulses to represent a sample which can assume one of the 32 possible values. It is easy to see that to represent a sample that can assume any one of the M possible values, we need a group of $\log_2 M$ number of binary pulses per sample. Note that it will take an infinite number of pulses per sample if the samples are transmitted exactly without quantization. This is because the samples can assume any value in a continuous range making $M = \infty$.

At the receiver, the binary pulses are converted back into the samples (the quantized samples) and are now passes through a low-pass filter to recover $x(t)$.

Let us assume that the samples of the signal $x(t)$ under consideration are quantized to M levels, each separated by Δv volts. Each sample is now represented by a group of $\log_2 M$ binary pulses. In a particular example, let, $M = 16$, $\Delta v = 0.1$ volt, and $\log_2 M = 4$. Let the mean square value of the noise be $\sigma_n^{\,2}$. We shall now calculate the signal power at the input of the demodulator. This input consists of binary pulses. In order to calculate the power of the coded signal, we must know the amplitude distribution of the modulating signal $x(t)$. For convenience, let us assume that the amplitudes of $x(t)$ are uniformly distributed in* the given range (0 to 1.5 volts).

This implies that the likelihood of observing the signal amplitude at any one value in the given range is the same as that at any other value in the same range. This leads to the conclusion that the chances of occurrence of any one quantized level is the same as any other (equal to 1/M). It is therefore obvious that the likelihood of observing any one code pattern is the same as that of any other pattern. Thus, each of the code patterns occurs with the same frequency. Hence, the likelihood of observing 0 volt pulse is equal to that of observing A volt pulse (equal to $\frac{1}{2}$). On the average, half the time the pulse of A volts will be present, and the remaining half of the time there will be no pulse. It is obvious that the average signal power S_i is given by

$$S_i = \frac{A^2}{2}$$

For PCM, the signal detection is merely a question of recognizing the presence or absence of a pulse. If we let $A = K\sigma_n$, then making K sufficiently large, the probability of error can be reduced as much as desired. It can be shown that for $K = 10$, the error probability is reduced to the order of 10^{-6}. In this chapter however, we shall leave K as an unknown constant. If the pulse is equally likely to be present and absent, then

$$S_i = \frac{A^2}{2} = \frac{K^2\sigma_n^2}{2}$$

and
$$N_i = \sigma_n^{\,2}$$

Therefore,
$$\frac{S_i}{N_i} = \frac{K^2}{2} \qquad ...(7.146)$$

By making K large enough, we can make the error probability in this detection as low as possible. Hence, in effect the channel noise is eliminated. However, we do have an error in the received signal due to quantization. There is an uncertainty introduced at the transmitter itself, and the error between the actual and approximated signals acts as a noise. We call this the quantization noise.

To calculate the quantization noise, recall that the mean square value of bandlimited signal is equal to the mean square value of its samples (taken at a rate greater than or equal to the Nyquist rate):

$$\overline{x^2(t)} = \overline{x_k^2(t)}$$

We shall now calculate the mean square value of the error (quantization noise) introduced by the process of quantization. Each sample of $x(t)$ is approximated to the nearest allowed level. Let the kth sample of $x(t)$ be x_k and the approximated kth sample be $\hat{x}_k$. Obviously $\hat{x}_k$ is one of the M allowed levels closest to x_k. We can express the signal $x(t)$ in terms of its samples f_k *i.e.*,

* This assumption is not necessary. The results derived here can be shown to be valid for any amplitude distribution provided the quantizing interval is small enough.

$$x(t) = \sum_k x_k Sa(\omega_m t - k\pi) \qquad ...(7.147)$$

At the receiver, when the pulse train is decoded, it yields samples $\hat{x}_k$. These samples, when passed through a low-pass filter, yield a continuous signal $\hat{x}_k$ *i.e.*,

$$\hat{x}_k = \sum_k \hat{x}_k Sa(\omega_m t - k\pi) \qquad ...(7.148)$$

The error signal $e(t)$ (which acts as a noise) is given by

$$e(t) = x(t) - \hat{x}(t) = \sum_k (x_k - \hat{x}_k)\, Sa(\omega_m t - k\pi) \qquad ...(7.149)$$

If we denote $\quad e_k = x_k - \hat{x}_k$

then $\quad e(t) = \sum_k e_k\, Sa(\omega_m t - k\pi)$

It is obvious that the mean square value of $e(t)$ is equal to the mean square value of its samples e_k *i.e.*,

$$\overline{e^2(t)} = \overline{e_k^2}$$

It may be noted that e_k is the difference between the actual sample value $x(t)$ and the quantized sample value $\hat{x}_k$. Since the quantized levels are spaced Δv volts apart, e_k must lie in the range $(-\Delta v/2), (\Delta v/2)$. Since the amplitude distribution of $x(t)$ is assumed to be uniform in the range $[0, (M-1)\,\Delta v]$, the distribution of e_k will also be uniform in the range in the range $(-\Delta v/2), (\Delta v/2)$, that is, the likelihood of observing e_k, a typical quantization error at one value in the range $(-\Delta v/2), (\Delta v/2)$, is the same as that of observing at any other value in the same range.

The mean square value of e_k is obviously given by

$$\overline{e_k^2} = \lim_{N\to\infty} \frac{1}{N} \sum_{n=-N/2}^{N/2} \left(\frac{n\Delta v}{N}\right)^2 = \lim_{N\to\infty} \frac{2}{N} \sum_{N=0}^{N/2} \left(\frac{n\Delta v}{N}\right)^2$$

$$= \lim_{N\to\infty} \frac{2(\Delta v)^2}{N^3} \sum_{N=0}^{N/2} n^2 = \lim_{N\to\infty} \frac{2(\Delta v)^2}{N^3} \left[\frac{\frac{N}{2}\left(\frac{N}{2}+1\right)(N+1)}{6}\right]$$

or $\quad \overline{e_k^2} = \dfrac{(\Delta v)^2}{12}$

Thus, $\quad e^2(t) = \dfrac{(\Delta v)^2}{12}$

But, $\overline{e^2(t)}$ is the mean square of the quantization error or quantization noise at the output. We shall denote the quantization noise power by N_o. Hence,

$$N = N_q = \frac{(\Delta v)^2}{12} \qquad ...(7.150)$$

Next, we calculate the signal power at the output. The output signal $x(t)$ is given by

$$\hat{x}(t) = \sum_k \hat{x}_k Sa(\omega_m t - k\pi)$$

We have $\quad \overline{\hat{x}^2(t)} = \overline{(\hat{x}_k)^2}$

Thus, the mean square value of the output signal is equal to the mean square value of the quantized samples. The quantized samples can assume M discrete values, and each is equally probable. This is because we have assumed the amplitude of signal $x(t)$ equally likely to occur in the range 0 to $(M-1)\Delta v$. The sample $\hat{x}_k$ can assume any of the M values $0, \Delta v, 2\Delta v, ..., (M-1)\Delta v$, with the same probability. Clearly the mean square value of f_k is

$$\overline{(\hat{x}_k)^2} = \frac{1}{M}\{0^2 + (\Delta v)^2 + (2\Delta v)^2 + ... + [(M-1)\,\Delta v]^2\}$$

$$= \frac{(\Delta v)^2}{M} \sum_{k=0}^{M-1} k^2 = \frac{\Delta v^2}{M} \frac{M(M-1)(2M-1)}{6} = \frac{\Delta v^2 (M-1)(2M-1)}{6}$$

or

In practice, $M >> 1$ and

$$\overline{(\hat{x}_k)^2} \cong \frac{M}{3}(\Delta v)^2$$

Hence $\quad S_o = \overline{(\hat{x}_k^2)} = \dfrac{M^2}{3}(\Delta v)^2 \qquad ...(7.151)$

From equations (7.148), (7.150) and (7.151), we obtain

$$\frac{(S_o/N_o)}{(S_i/N_i)} = \frac{4M^2}{K^2/2} = \frac{8M^2}{K^2} \qquad ...(7.152)$$

We must now relate the signal-to-noise improvement ratio to the bandwidth of the system. We have seen that for M quantization levels, we need $\log_2 M$ pulses per sample. Since, for a signal bandlimited to f_mHz, we need $2f_m$ samples per second, the system must be capable of transmitting $2f_m \log_2 M$ pulse per second. We know that a system of bandwidth B is capable of transmitting $2B$ independent pulses per second.

Hence, to transmit $2f_m \log_2 M$ pulses per second, the bandwidth B required is $f_m \log_2 M$.

$$B = f_m \log_2 M \quad \text{and} \quad M^2 = (2)^{2B/f_m}$$

Thus, $\quad \dfrac{(S_o/N_o)}{(S_i/N_i)} = \dfrac{8}{K^2}(2)^{2B/f_m} \qquad ...(7.153)$

If we consider the S/N root mean square voltage ratio rather than the power ratio, we get

$$\frac{(S_o/N_o)_{VT}}{(S_i/N_i)_{VT}} = \frac{\sqrt{8}}{K}(2)^{B/f_m} \qquad ...(7.154)$$

7.16. Comparison of Coded and Uncoded Systems

Up to now we have discussed a number of wideband systems which allow exchange of S/N ratio and bandwidth. Examples of uncoded wideband systems are FM (also PM) and PPM. The pulse code modulation is an example of a coded wideband system.

In uncoded system, (FM and PPM), we have observed that the improvement in S/N voltage ratio is linear with bandwidth. On the other hand for the coded system (PCM), the improvement of S/N voltage ratio was observed to be exponential with bandwidth. Thus trading bandwidth and S/N ratio is much more efficient in coded systems than in uncoded systems. Coded systems are inherently capable of better transmission efficiency than uncoded systems.

It can shown in that the maximum theoretical attainable improvement in S/N ratio is exponential with bandwidth.* Thus, coded systems come close to realization of the theoretical maximum

* This point will be discussed in detail in next chapter.

efficiency that can be attained. The qualitative reason for the exponential relationship in PCM is not difficult to understand.

For uncoded systems, doubling the bandwidth doubles the S/N voltage ratio. But in PCM, doubling the bandwidth allows twice the number of pulses to be transmitted. If we double the number of pulses, the possible patterns that can be formed increases as a square. For example, a group of 2 binary pulses can form 4 patterns. But a group of 4 binary pulses can from 16 patterns. Hence, doubling the bandwidth squares the quantized levels. Similarly, tripling the bandwidth will allow three times the number of pulses. But, tripling the number of pulses increases the number of patterns to the power 3.

In general, increasing the bandwidth n times increases the number of patterns exponentially (to the nth power). This allows quantized levels to increase exponentially in a given range. Obviously, Δv, the quantum gap, decreases exponentially with the bandwidth. The quantizing noise power (the output noise power) is $(\Delta v/12)^2$ and reduces exponentially. This causes the S/N ratio to improve exponentially.

The most important advantage of a coded system is it convenience for long distance communication by using several repeater stations. At each repeater station, the signal mixed with noise is cleaned up by regenerating the pulses. Thus, at each repeater station a noise-free signal is transmitted. In effect, this implies that for PCM, one needs to be concerned only about the noise on the link between the repeater stations. The noise does not accumulate throughout the transmission path. Thus, the transmission requirements for PCM are almost independent of the total length of the system.

In fact, PCM can be transmitted over any distance by placing enough repeater stations. This is the basic reason for using coded systems. In contrast, for uncoded systems or analog modulation systems the signal is continuously affected by the noise and cannot be cleaned up or regenerated periodically along the path. The noise accumulates throughout the transmission path. The larger the transmission path, the larger will be the noise. This restricts the distance over which such signals can be transmitted.

Now, let us recognize the similarity between coded and uncoded systems with the digital and analog systems, respectively. In analog systems, say, analog computers, increasing the voltage representing the physical variable increases the precision linearly. In the digital systems, however, increasing the number of digits increases the precision exponentially. The errors in analog systems accumulate, whereas in digital systems (the digital computer, for example) accumulation does not occur and the noise below a certain level has no influence.

MISCELLEANOUS SOLVED EXAMPLES

Example 7.1. Given an analog baseband communication system with additive while noise having power spectral density (psd) $\eta/2$ and a distorting channel having the frequency response

$$H_c(\omega) = \frac{1}{1 + j\omega / W}$$

The distortion is equalized by a receiver filter having the frequency response

$$H_{eq}(\omega) = \begin{cases} \dfrac{1}{H_c(\omega)} & \text{for } \le |\omega| \le W \\ 0 & \text{otherwise} \end{cases}$$

obtain an expression for the output signal-to-noise ratio (SNR).

Solution: We have

$$S_o = \frac{1}{2\pi}\int_{-\infty}^{\infty} |H_c(\omega)|^2 |H_{eq}(\omega)|^2 S_{XX}(\omega)\, d\omega$$

where $S_{XX}(w)$ is the psd, of message signal $X(t)$.

$$S_o = \frac{1}{2\pi}\int_{-W}^{W} S_{XX}(\omega)\, d\omega = S_X$$

where S_X is the average signal power

The average output noise power is given by

$$N_o = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\eta}{2} |H_{eq}(\omega)|^2\, d\omega$$

Also,

$$N_o = \frac{\eta}{2\pi}\int_{0}^{W}\left[1 + \left(\frac{\omega}{W}\right)^2\right] d\omega = \frac{\eta}{2\pi}\left(\frac{4}{3}W\right) = \frac{4}{3}\eta f_m$$

or

Therefore, we have

$$\left(\frac{S}{N}\right)_o = \frac{S_o}{N_o} = \frac{S_X}{\frac{4}{3}\eta f_m} = \frac{3}{4}\frac{S_X}{\eta f_m} \quad \textbf{Ans.}$$

Example 7.2. Consider an analog baseband communication system with additive white noise. The transmission channel is assumed to be distortionless, and the power spectral density (psd) of white noise $\eta/2$ is 10^{-9} watt per hertz (W/Hz). The signal to be transmitted is an audio signal with 4 kHz bandwidth. At the receiver end, an RC low-pass filter with a 3 dB bandwidth of 8 kHz is used to limit the noise power at the output. Compute the output noise power.

Solution: We know that the frequency response $H(\omega)$ of an RC low pass filter with a 3 dB bandwidth of 8 kHz is given by

$$H(\omega) = \frac{1}{1 + j\omega / \omega_0} \qquad \text{...(i)}$$

Here, ω_0 is 3 dB bandwidth of an RC low-pass filter.

and $\omega_0 = 2\pi(8000)$

Now the output noise power N_o is given by

$$N_o = E[n^2{}_0(t)] = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\eta}{2} |H(\omega)|^2\, d\omega$$

Substituting value of $H(w)$ from equation (i), we get

$$N_o = \frac{\eta}{2}\frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{1}{1 + (\omega/\omega_0)^2}\, d\omega$$

Simplifying, we get

$$N_o = \frac{1}{4}\eta\omega_0 = \frac{1}{4}2(10^{-9})(2\pi)(8)(10^3)W = 25.2\ \mu\text{W} \quad \textbf{Ans.}$$

Example 7.3. Assuming sinusoidal modulation, prove that in an AM system with envelope detection, the output signal to noise ratio (SNR) is given by

$$\left(\frac{S}{N}\right)_o = \frac{m^2}{2 + m^2}\gamma$$

where m is the modulation index for AM and $\gamma = \dfrac{S_i}{nf_m}$. *(U.P. Tech., Sem. Exam. 2004-05)*

* $W = 2\pi f_m$.

@sTc

Solution: For sinusoidal modulation, we have

$$X(t) = \cos \omega_m t^*$$

Therefore, average signal power S_X is given by

$$S_X = E[X^2(t)] = \frac{1}{2} \quad ...(i)$$

We know that for AM system (with envelope detection), the output SNR is given by

$$\left(\frac{S}{N}\right)_o = \frac{m^2 S_X}{1 + m^2 S_X}\gamma$$

Using equation (*i*), we get

$$= \frac{m^2\left(\frac{1}{2}\right)}{1 + m^2\left(\frac{1}{2}\right)}\gamma = \frac{m^2}{2 + m^2}\gamma$$

Note that with 100 per cent modulation ($m = 1$), we have

$$\left(\frac{S}{N}\right)_o = \frac{1}{3}\gamma \quad \textbf{Hence Proved}$$

Example 7.4. Prove that the performance of an SSB system using synchronous detection is equivalent to the performance of both DSB and baseband systems.

Solution: The SSB signal $X_c(t)$ can be expressed as

$$X_c(t) = A[X(t)\cos\omega_c t + \hat{X}(t)\sin\omega_c t] \quad ...(i)$$

where $\hat{X}(t)$ denotes the Hilbert transform of $X(t)$. Note that equation (*i*) represents the lower-sideband SSB signal and that the minimum bandwidth of predetection filter is W for a single-band signal.

The input of the receiver is given by

$$X_i(t) = X_c(t) + n_i(t) \quad ...(ii)$$

$$= [AX(t) + n_c(t)]\cos\omega_c t + [A\,\hat{X}(t) - n_s(t)]\sin w_c t$$

Synchronous detection rejects the quadrature components of both the signal and noise, and thus yielding

$$Y_o(t) = A\,X(t) + n_c(t) = X_o(t) + n_o(t) \quad ...(iii)$$

where $X_o(t) = A\,X(t)$ and $n_o(t) = n_c(t)$

The output signal power is given by

$$S_o = E[X_o^2(t)] = A^2E[X^2(t)] = A^2S_X$$

The power spectral density of $n_c(t)$ is illustrated in figure 7.12 for the case of lower-sideband SSB.

The output noise power is given by

$$N_o = E[n_c^2(t)] = \frac{\eta W}{2\pi} = \eta f_m$$

Thus, the output SNR will be

$$\left(\frac{S}{N}\right)_o = \frac{A^2 S_X}{\eta f_m} \quad ...(iv)$$

The input signal power will be

$$S_i = E[X_c^2(t)]$$

* Here, we have considered $X(t)$ as a random process.

$$S_i = A^2\left\{\frac{1}{2}E[X^2(t)] + \frac{1}{2}E[\hat{X}^2(t)]\right\} = A^2E[X^2(t)] = A^2S_X$$

Since, we have

$$E[X(t)\hat{X}(t)] = 0 \quad \text{and} \quad E[X^2(t)] = E[\hat{X}^2(t)]$$

Therefore,

$$\left(\frac{S}{N}\right)_o = \frac{S_i}{\eta f_m} = \gamma \quad ...(v)$$

which implies that insofar as noise is concerned, SSB with ideal synchronous detection has the same performance as both DSB and baseband systems.

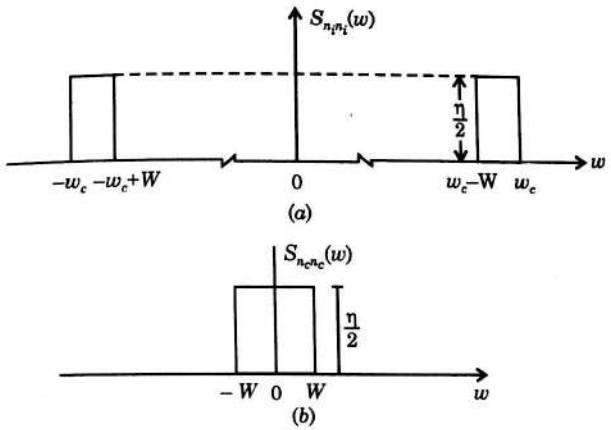


Fig. 7.12.

Example 7.5. Given an AM system with additive thermal noise having a power spectral density (psd) $\eta/2 = 10^{-12}$ W/Hz. Assume that the baseband message signal $X(t)$ has a bandwidth of 4 kHz and the amplitude distribution as shown by figure 7.13. The signal is demodulated by envelope detection and appropriate post-detection filtering. Assume $m = 1$.

(*i*) Determine the minimum value of carrier amplitude A that will yield $(S/N)_o \geq 40$ dB.

(*ii*) Find the threshold value of A.

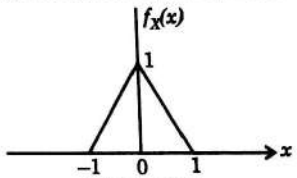


Fig. 7.13.

Solution: (*i*) We have

$$X_c(t) = A[1 + X(t)]\cos\omega_c t$$

Average signal power is given by

$$S_X = E[X^2(t)] = \int_{-\infty}^{\infty} x^2 f_X(x)\,dx = 2\int_0^1 x^2(-x+1)dx = \frac{1}{6}$$

For ($m = 1$), we have

$$\left(\frac{S}{N}\right)_o = \frac{S_X}{1 + S_X}\gamma = \frac{\frac{1}{6}}{1 + \frac{1}{6}}\gamma = \frac{1}{7}\gamma \geq 10^4$$

Hence,

$$\gamma \geq 7(10^4)$$

Also, we have

$$\frac{A^2\left(1+\frac{1}{6}\right)}{4(10^{-12})(4)(10^3)} \geq 7(10^4)$$

from which, we get $A \geq 31(10^{-3})V = 31$ mV

Hence, the minimum value of A required is 31V. **Ans.**

(*ii*) Again, $(S/N)_i$ at threshold is given by

$$\left(\frac{S}{N}\right)_{i_{th}} = 10$$

Therefore, $\gamma|_{\text{at threshold}} = \gamma_{\text{th}} = 2\left(\frac{S}{N}\right)_{i_{th}} \cong 20$

and from part (*i*), we have

$$\frac{A^2\left(1+\frac{1}{6}\right)}{4(10^{-12})(4)(10^3)} = 20$$

from which, we obtain

$$A = 0.52(10^{-3})V = 0.52 \text{ mV}$$

which is the threshold value of A. **Ans.**

Example 7.6. Compute the transmission bandwidth B_T and the required transmitted power S_T of DSB, SSB and AM systems for transmitting an audio signal which has a bandwidth of 10 kHz with an output SNR of 40 dB. Assume that the channel introduces a 40 dB power loss and channel noise is AWGN with power spectral density (psd) $\eta/2 = 10^{-9}$ W/Hz. Assume $\mu^2 S_X = 0.5$ for AM.

Solution: As a matter of fact, the bandwidth requirements are easily found as under:

$$B_T = \begin{cases} 20\text{kHz} & \text{for DSB and AM} \\ 10\text{kHz} & \text{for SSB} \end{cases}$$

For the transmitter power of DSB and SSB systems, we have

$$\left(\frac{S}{N}\right)_o = \frac{S_i}{\eta f_m} = 10^4 \ (= 40 \text{ dB})$$

and $S_i = \eta f_m(10^4) = 2(10^{-9})(10^4)(10^4) = 0.2$ W

Because the channel power loss is 40 dB, the required transmitted power S_T is

$$S_T = 0.2(10^4) = 2000 \text{ W} = 2 \text{ kW}$$

For an AM system using envelope detection with $m^2 S_X = 0.5$, we have

$$\left(\frac{S}{N}\right)_o = \frac{1}{3}\left(\frac{S_i}{\eta f_m}\right)$$

Hence, the required transmitted power is 3 times than that for the DSB or SSB system, *i.e.*,

$S_T = 6$ kW **Ans.**

Example 7.7. Prove that in an FM system the output SNR, assuming sinusoidal modulation, is given by

$$\left(\frac{S}{N}\right)_o = \frac{3}{2} m^2 \gamma \qquad \text{...(i)}$$

where m is the modulation index for FM.

Solution: For sinusoidal modulation, we have

$$X(t) = \cos \omega_m t$$

Therefore, average signal power is given by

$$S_X = E[X^2(t)] = \frac{1}{2}$$

Again, we know that $\left(\frac{S}{N}\right)_0$ in FM system is given by

$$\left(\frac{S}{N}\right)_o = 3\left(\frac{k_f^2 S_X}{W^2}\right)\gamma = \frac{3}{2}\left(\frac{\Delta\omega}{W}\right)^2 \gamma = \frac{3}{2} m_f{}^2 \gamma$$

Since $\Delta\omega = [k_f x(t)|_{\max} = k_f$ and $\frac{\Delta\omega}{W} = m_f$

Note: *It is important to note that the modulation index m_f is determined by the bandwidth W of the post-detection low-pass filter and is not related to the sinusoidal message frequency ω_m, except insofar as this filter is chosen so as to pass the spectrum of the desired filter. For a specified bandwidth W, the sinusoidal message frequency may lie anywhere 0 and W and would yield the same output SNR*

Example 7.8. Prove that narrowband FM offers no improvement in SNR over AM.

Solution: For AM with envelope detection and assuming 100 per cent sinusoidal modulation, the output SNR is given by

$$\left(\frac{S}{N}\right)_o = \frac{1}{3}\gamma$$

For FM with sinusoidal modulation, the output SNR is given by

$$\left(\frac{S}{N}\right)_o = \frac{3}{2} m_f^2 \gamma$$

Hence, we observe that the use of FM offer the possibility SNR over AM when

$$\frac{3}{2} m_f^2 > \frac{1}{3} \text{ or } m_f > 0.47$$

We know that a value of $m_f < 0.2$ is considered to define an FM signal to narrowband. We conclude that the narrowband FM offers no improvement in SNR over AM. **Hence Proved.**

Example 7.9. An audio $x(t)$ is to be transmitted over a radio frequency (RF) channel with additive white noise. It is required that the output SNR be greater than 40 dB. Assume the following characteristics for $x(t)$ and the channel:

$$E[x(t)] = 0 \ |x(t) \leq 1 \ S_x = E[x^2(t)] = \frac{1}{2} \ B = 15 \text{ kHz}$$

Power spectral density of white noise $\eta/2 = 10^{-10}$ W/Hz

Compute the transmission bandwidth B_T and required average power S_T for

(*i*) DSB modulation

(*ii*) AM with 100 per cent modulation and envelope detection

(*iii*) PM with $k_p = 3$

(*iv*) FM with a deviation ratio $D = 5$.

Solution: (*i*) For DSB modulation

$$B_T = 2B = 30 \text{ kHz}$$

We have

$$\left(\frac{S}{N}\right)_o = \gamma = \frac{S_i}{\eta B} = \frac{S_i}{2(10^{-10})(15)(10^3)} \geq 10^4 \ (= 40 \text{ dB})$$

or $\quad S_i \geq S_i(10^{-2})W$

Hence, $\quad S_T = S_i(10^5) \geq 3000 \text{ W} = 3 \text{ kW}$

(*ii*) For AM with $\mu = 1$ and envelope detection

Again $\quad B_T = 2B = 30 \text{ kHz}$

From we have

$$\left(\frac{S}{N}\right)_o = \frac{1}{3}\gamma$$

Hence, $\quad S_T \geq 9 \text{ kW}$

(*iii*) For PM with $k_p = 3$, we have

$$B_T + 2(k_p + 1)B = 120 \text{ kHz}$$

Also,

$$\left(\frac{S}{N}\right)_o = k_p{}^2 S X_g = (3^2)\left(\frac{1}{2}\right)\frac{S_i}{2(10^{-10})(15)(10^3)} \geq 104$$

or

$$S_i \geq \frac{2}{3}(10^{-2})W$$

Therefore,

$$S_T = S_i(10^5) \geq 667 \ W$$

(*iv*) For FM with $D = 5$, from

$$B_T = 2(D + 1)B = 180 \text{ kHz}$$

We have

$$\left(\frac{S}{N}\right)_o = 3D^2 S_X \gamma = 3(5^2)\left(\frac{1}{2}\right)\frac{S_i}{2(10^{-10})(15)(10^3)} \geq 10^4$$

or

$$S_i \geq \frac{2}{25}(10^{-2})W$$

Therefore, $\quad S_T = S_i(10^5) \geq 80 \text{ W}$

Example 7.10. Consider that the power spectral density of the signal $X(t)$ is

$$S_{XX}(\omega) = \begin{cases} \dfrac{1}{1 + (\omega/\omega_1)^2} & \text{for } |\omega| < 2\pi(15)(10^3) \\ 0 & \text{elsewhere} \end{cases}$$

where $\omega_1 = 2\pi(2.1)(103)$ rad/s.

(*i*) Determine the required change in modulation level with and without pre-emphasis for a fixed modulation power.

(*ii*) What is the net SNR improvement under the conditions?

Solution: (*i*) Without pre-emphasis, the modulation signal power S_X is

$$S_X = \frac{1}{2\pi}\int_{-W}^{W} \frac{1}{1 + (\omega/\omega_1)^2}\, d\omega = \frac{\omega_1}{\pi}\tan^{-1}\frac{W}{\omega_1}$$

With pre-emphasis, we have

$$S_X = \frac{1}{2\pi}\int_{-W}^{W} S_{XX}|H_{PE}(\omega)|^2\, d\omega = \frac{1}{2\pi}\int_{-W}^{W} K d\omega = K\frac{W}{\pi}$$

Hence, if the modulation signal power is to reamain fixed, we must have

$$K = \frac{\tan^{-1}(W/\omega_1)}{W/\omega_1}$$

With $B = 15$ kHz and $f_1 = 2.1$ kHz, we get

$$K = \frac{\tan^{-1} 7.14}{7.14} = 0.200 (= -6.98 \text{ dB})$$

(*ii*) Note that the noise improvement factor with preemphasis/deemphasis is 13.27 dB. However, now that modulation level must be decreased by 6.98 dB, so the net SNR improvement is 13.27 – 6.98 dB.

Example 7.11. The following sample function

$$x(t) = A_c \cos(2\pi f_c t) + \omega(t)$$

is applied to the low-pass RC filter as shown in the figure 7.14. The amplitude A_c and frequency f_c of the sinusoidal component are constants, and $\omega(t)$ is a white Gaussian noise of zero mean and power spectral density $\eta/2$. Derive an expression for the output signal-to-noise ratio with the sinusoidal component of $x(t)$ regarded as the signal of interest. *(U.P. Tech., Sem. Exam., 2004-05)*

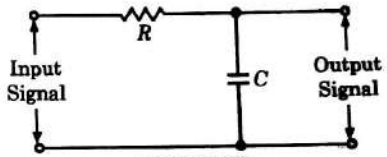


Fig. 7.14.

Solution: Frequency response $H(\omega)$ of the given RC low-pass filter is given by

$$H(\omega) = \frac{1}{j\omega RC + 1}$$

Then, output noise power, will be

$$N_0 = E[\eta_0{}^2(t)] = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_{xx}(\omega).|H(\omega)|^2\, d\omega$$

or

$$N_0 = \frac{1}{2\pi}\int_{-\infty}^{\infty} \frac{\eta}{2}\frac{d\omega}{(\omega RC)^2 + 1} \qquad \left[\text{Given, } S_{xx}(\omega) = \frac{\eta}{2}\right]$$

or

$$N_0 = \frac{1}{2\pi}\left(\frac{\eta}{2}\right)\frac{1}{RC}\left[\tan^{-1} RC\Big|_{-\infty}^{\infty}\right] = \frac{\eta}{4RC}.$$

Now, considering only sinusoidal component of $x(t)$ as input *i.e.*, $x(t) = A_c \cos(2\pi f\text{–}t)$. Then, corresponding value of the signal amplitude at the filter output will be

$$= \frac{A_c}{\left[1 + (2\pi f_c RC)^2\right]^{\frac{1}{2}}}$$

$$|SNR|_0 = \frac{2A_c^2 RC}{N_0[1 + (2\pi f_c RC)^2]} \qquad \textbf{Hence Proved.}$$

Example 7.12. The average noise power per unit bandwidth measured at the front end of an AM receiver is 10^{-3} watt per Hertz. The modulating wave is sinusoidal, with a carrier power of 80 kilowatts, and a sideband power of 10 kilowatts per sideband. The message bandwidth is 4 kHz. Assuming the use of an envelope detector in the receiver, compute the output signal-to-noise ratio of the system. By how many decibels is this system inferior to a DSB-SC modulation system?

Solution: Figure 7.15 shows an amplitude modulation system using an envelope detector.

The filtered signal $x(t)$ at the input of envelope detector is, given by

$$x(t) = s(t) + n(t)$$
$$= [A_c + A_c K_a m(t)] \,.\, \cos 2\pi f_c t + n_I(t) \,.\, \cos(2\pi f_c t) - n_Q(t) \,.\, \sin(2\pi f_c t)$$

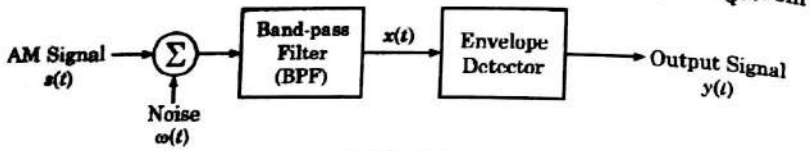


Fig. 7.15.

and receiver output

$$y(t) = \text{envelope of } x(t) = \{(A_c + A_c K_a m(t) + n_I(t)]^2 + [n_Q(t)]^2\}^{1/2}$$

For the receiver to operate satisfactorily, the signal term $A_C[1 + K_a m(t)]$ would be large compared to noise terms $n_I(t)$ and $n_Q(t)$, then

$$y(t) \simeq A_c + A_c K_a m(t) + n_I(t).$$

Because, dc term A_c bears no relation to the message signal $m(t)$, it can ignored, then average power of the information-bearing component,

$$A_c K_a m(t) = \left[A_c^2 K_a^2 \frac{P}{2}\right] 4$$

where P is the average power of the message signal $m(t)$.

The average power of the noise in message bandwidth is WN_0.

Then we have $$(SNR)_0 = \frac{A_c^2 K_a^2 P}{2WN_0} \quad \text{...(i)}$$

Now, carrier power $= \dfrac{A_c^2}{2} = 80 \times 10^3$ watts

Power in each sideband of the AM wave $= \dfrac{m^2 A_c^2}{8} = 10 \times 10^3$ watts

Modulating factor is given by

$$m = \sqrt{0.5} = 0.707$$

From equation (i), the output signal-to-noise ratio of an AM receiver using envelope detection is given by

$$(SNR)_0 = \frac{A_c^2 k_a^2 P}{2WN_0}$$

where $$k_a{}^2 P = \frac{1}{2} k_a^2 A_m^2 = \frac{1}{2} m^2 = 0.25$$

Substituting values, we get

$$(SNR)_0 = \frac{80 \times 10^3 \times 0.25}{4 \times 10^3 \times 10^{-3}} = 5 \times 10^3$$

For a DSB-SC system, figure of merit will be

$$= \frac{(SNR)_0}{(SNR)_c} = \frac{\left(\dfrac{A_c^2 P}{2WN_0}\right)}{\left(\dfrac{A_c^2 P}{2WN_0}\right)} = 1$$

and for an AM system using envelope detection, we have

$$\frac{(SNR)_0}{(SNR)_c} = \frac{m^2}{m^2 + 2} = 0.2$$

This means that the AM system is inferior to the DSB-SC system by an amount = equal to $10 \log_{10} 5 = 7$ dB

$$(SNR)_2 = \frac{1}{2} \times \frac{0.99 \times 10^{-6}}{4} \times 8 \times 10^3$$
$$= 0.99 \times 10^{-3} \text{ watts.}$$

Average output signal power $= \dfrac{A_c^2 P}{16}$

where P is the message signal power

The SSB modulated wave is

$$s(t) = \frac{1}{2} A_c m(t) \cos(2\pi f_c t) - \frac{1}{2} A_c \hat{m} \sin(2\pi f_c t)$$

Average power of $s(t) = \dfrac{A_c^2 P}{4} = 10$ watts

$\therefore$ The output signal power = 2.5 watts.

$$(SNR)_0 = \frac{2.5}{0.99 \times 10^{-3}} = \mathbf{2.53 \times 10^3}$$

Expressing this results in decibels, we thus have an output signal-to-noise of 44 dB. **Ans.**

Example 7.13. An unmodulated carrier of amplitude A_c and frequency f_c and bandlimited white noise are summed and then passed through an ideal detector. Assume the noise spectral density to be of height $N_0/2$ and bandwidth 2 W, centred about the carrier frequency f_c. Find the output signal-to-noise ratio for the case when the carrier-to-noise ratio is high.

Solution: The received signal is,

$$x(t) = A_c \cos(2\pi f_c t) + n(t)$$
$$= A_c \cos(2\pi f_c t) + n_I(t).\cos(2\pi f_c t) - n_Q(t).\sin(2\pi f_c t)$$
$$= [A_c + n_I(t)] .\cos(2\pi f_c t) - n_Q(t).\sin(2\pi f_c t)$$

Then, the envelope detector output is,

$$a(t) = \text{envelope of } x(t)$$
$$= \left\{[A_c + n_1(t)]^2 + n_Q^2(t)\right\}^{\frac{1}{2}}$$

Now, when the carrier-to-noise ratio is high, then signal term $[A_C + n_1(t)]$ will be compared with noisy term $n_Q(t)$.

Then, $$a(t) \cong A_c + n_c(t).$$

The term A_c represent the useful signal component, thus output signal power is $A_c{}^2$.

The power spectral densities of $n(t)$ and $n_I(t)$ have been shown in figures 7.16 and 7.17.

From figure 7.17, it may be observed that output noise power is $2N_0W$.

Therefore, output signal-to-noise power ratio is $(\text{SNR})_0 = \dfrac{A_c^2}{2N_0W}$.

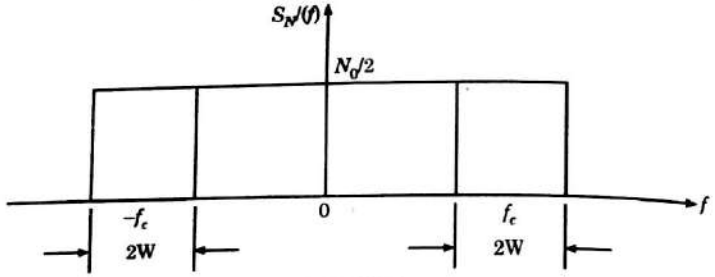


Fig. 7.16.

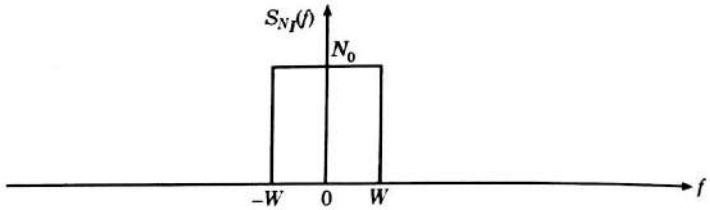


Fig. 7.17.

Example 7.14. Let a message signal $m(t)$ be transmitted using single-sideband modulation. The power spectral density of $m(t)$ is

$$S_M(f) = \begin{cases} a\dfrac{|f|}{W} & \text{for} \quad |f| \le W \\ 0, & \text{elsewhere} \end{cases}$$

where a and W constants. White Gaussian noise of zero mean and power spectral density $\frac{N_0}{2}$ is added to the SSB modulated wave at the receiver input. Derive an expression for the output signal-to-noise ratio of the receiver.

Solution: The power spectral density of the message signal $m(t)$ has been shown in figure 7.18.

Average signal power is given by

$$P = \int_{-\infty}^{\infty} S_M(f)\, df$$

$$= 2\int_0^W \frac{af}{W}\, dW = aW \qquad ...(i)$$

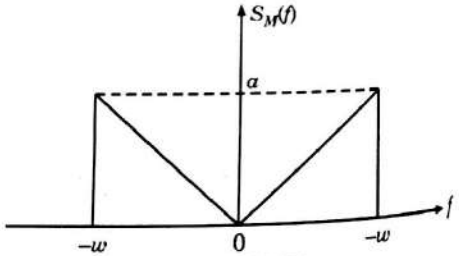


Fig. 7.18.

Now, output signal-to-noise a system, using SSB modulation in the transmitter and coherent detection in the receiver, is given by

$$(SNR)_0 = \frac{A_c^2 P}{4WN_0}, \text{ here } A_c \text{ is carrier amplitude}$$

Using equation (*i*), we get

$$(SNR)_0 = \frac{A_c^2 . aW}{4WN_0} = \frac{aA_c^2}{4N_0} \qquad \textbf{Hence Proved.}$$

Example 7.15. Given a phase modulation (PM) system, with the modulated wave defined by

$$s(t) = A_c \cos[2\pi f_c t + K_p m(t)]$$

where k_p is a constant and $m(t)$ is the message signal. The additive noise $n(t)$ at the phase detector input is

$$n(t) = n_I(t)\cos(2\pi f_c t) - n_Q(t)\sin(2\pi f_c t) \qquad ...(i)$$

Assuming that the carrier-to-noise ratio at the detector input is high compared with unity, determine (*i*) the output signal-to-noise ratio and (*ii*) the figure of merit of the system. Compare your results with the FM system for the case of the sinusoidal modulation.

(GATE Examination, 1999)

Solution: Figure 7.19 below for a angle modulation system,

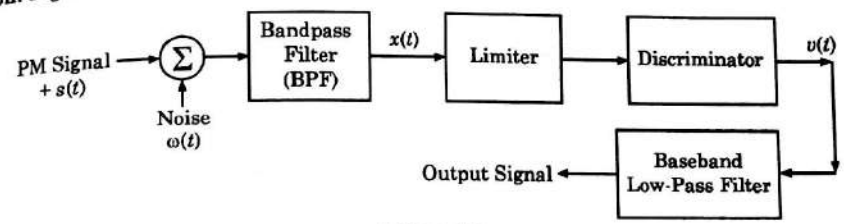


Fig. 7.19.

The transmitted signal has the form

$$s(t) = A_c \cos[\omega_c t + \phi(t)]$$

where $\phi(t) = K_p m(t)$

The noisy signal at the bandpass filter output is

$$x(t) = s(t) + n(t)$$

where $n(t)$ is narrowband noise.

$$= A_c \cos[\omega_c t + \phi(t)] + n_I(t).\cos(\omega_c t) - n_Q(t).\sin(\omega_c t)$$

$$= A_c \cos[\omega_c t + \phi(t)] + r(t).\cos[2\omega_c t + \psi(t)] \qquad ...(i)$$

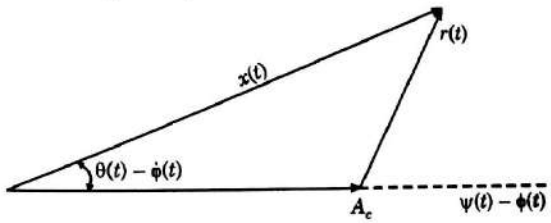


Fig. 7.20.

From the phasor, we have

$$\theta(t) = \phi(t) + \tan^{-1}\left\{\frac{r(t).\sin[\psi(t) - \phi(t)]}{A_c + r(t).\cos[\psi(t) - \phi(t)]}\right\} \qquad ...(ii)$$

Now, proceeding from equation (*i*), we get

$$x(t) = V(t).\cos[\omega_c t + \theta(t)]$$

where $V(t) = \{[A_c \cos\phi + r(t)\cos\psi(t)]^2 + [A_c \sin\phi + r(t).\sin\psi(t)]^2\}^{1/2}$

and $\theta(t) = \tan^{-1}\dfrac{A_c \sin\phi + r(t).\sin\psi(t)}{A_c \cos\phi + r(t)\cos\psi(t)}$

The limiter suppresses any amplitude variation in $V(t)$. Hence, inphase modulation, SNRs are derived from considering only $\theta(t)$.

Then, assuming detector to be ideal, its output will be

$$Y_0(t) = \theta(t)$$

Now, because carrier to noise is high compared to that of unity at the detector input, then, for equation (*ii*), we have

$$\theta(t) \equiv \phi(t) + \frac{r(t)}{A_c}.\sin[\psi(t) - \phi(t)]$$

Then, for PM; $\phi(t) = K_p m(t)$,

Then, we have $\theta(t) \equiv K_p m(t) + \frac{r(t)}{A_c}.\sin[\psi(t) - \phi(t)]$

Replacing $\psi(t) - \phi(t)$ with $\psi(t)$ will not affect the result as carrier-to-noise is high.

$$\theta(t) = K_p m(t) + \frac{r(t)}{A_c}.\sin[\psi(t)]$$

Therefore, output at phase discriminator will be

$$Y(t) = k_p m(t) + \frac{n_Q(t)}{A_C}$$

The message signal component of $y(t) = K_p m(t)$

Hence, the average output signal power $= K_p{}^2 p$

where p is the message signal power.

Now, because the phase detector is restricted to the message bandwidth W, the average output noise power $= \frac{2WN_0}{A_c^2}$.

Hence, the output signal-to-noise ratio of the PM system is given by

$$(SNR)_0 = \frac{K_p^2 P A_c^2}{2WN_0}$$

(*ii*) The channel signal-to-noise ratio of the PM system is the same as that of the corresponding FM system, *i.e.*,

$$(SNR)_0 = \frac{A_c^2}{2WN_0}$$

Therefore, the figure of merit of the PM system $= K_p{}^2 P$.

For the case of sinusoidal modulation, we get

$$m(t) = A_m \cos(2\pi f_m t)$$

Thus, we have $P = \frac{A_m^2}{2}$

Therefore, corresponding value of the figure of merit for a PM system $= \frac{1}{2} m \frac{2}{p}$

where $m_p = K_p A_m$.

On the other hand, the figure of merit for an FM system with sinusoidal wave is

$$= \frac{3}{2} m^2.$$

Note: *We observe, therefore, that for a specified phase deviation, the FM system is 3 times as good as the PM system.*

Example 7.16. An AM receiver, operating with a sinusoidal modulating signal and 80 per cent modulation, has an output signal-to-noise ratio of 30 dB.

(*i*) What is the corresponding carrier-to-noise ratio?

(*ii*) By how many decibels can we decrease the carrier-to-noise ratio so that the system is operating just above threshold?

Solution: The output signal-to-noise ratio of an AM receiver using envelope detection is,

$$(SNR)_0 \cong \frac{A_c^2 K_a^2 P}{2WN_0} = \frac{A_c^2 K_a^2 A m^2}{4WN_0} = \frac{A_c^2 m^2}{4WN_0}$$

where m is the modulating factor.

The carrier-to-noise ratio is defined as under:

$$\rho = \frac{A_c^2/2}{2WN_0}$$

Thus, we have

$$(SNR)_0 = m^2 \rho$$

With an output signal-to-noise ratio of 30 dB, we have

$$(SNR)_0 = 10^3$$

Hence, with $m = 0.8$, we get

$$\rho = 1563$$

$\therefore$ Carrier-to-noise ratio = **31.9. dB.**

(*ii*) With a carrier-to-noise ratio of 6.6 dB, the envelope detector is expects to operate satisfactorily. Therefore, by reducing the carrier-to-noise ratio by an amount equal to 3.19 – 6.6 = 25.3 dB, we expect the receiver to operate above threshold.

SUMMARY

1. Noise analysis in continuous wave (CW) modulation is carried out in the form of a parameter known as figure of merit denoted by γ. This parameter figure of mert γ is the ratio of output singal-to-noise ratio to the input signal-to-noise ratio of the receiver. This means that we compare output signal-to-noise ratio (S_0/N_0) with the input signal-to-noise ratio (S_i/N_i) of the receiver to evaluate the noise performance.
2. In any practical communication system, the front end of the receiver consists of a tuned circuit also known as bandpass filter. This bandpass filter processes the incoming desired signal and the noise.
3. Inphase noise component
$$n_c(t) = n(t) \cos w_c t + \hat{n}(t) \sin w_c t$$
4. Quadrature noise component
$$n_s(t) = \hat{n}(t) \cos w_c t - n(t) \sin w_c t$$
5. The figure of merit 'γ' may be defined as the ratio of output signal to noise ratio to input signal to noise ratio of a receiver.
Mathematically,
$$\gamma = \frac{\text{Output signal to noise ratio}}{\text{Input signal to noise ratio}} = \frac{S_0/N_0}{S_i/N_i}$$
6. Total noise power N = white noise power spectrum density × Bandwidth
or
$$N = \frac{n}{2} \times \text{Bandwidth}$$
Thus, the noise has a Gaussian distribution.
7. Channel noise is always additive. We assume that the disturbing effect of channel noise is always additive. This means that the effect of channel noise may be obtained by simple addition of signal $x(t)$ and noise $n(t)$.

8. The noise performance depends on the relative magnitudes of the signal and noise. The analysis will be carried out for two cases, as under
   (*i*) Small noise case,
   (*ii*) Large noise case

SHORT QUESTIONS WITH ANSWERS

Q.1. What do you mean by figure of merit?

Ans. The figure of merit 'γ' may be defined as the ratio of output signal to noise ratio to input signal to noise ratio of a receiver.

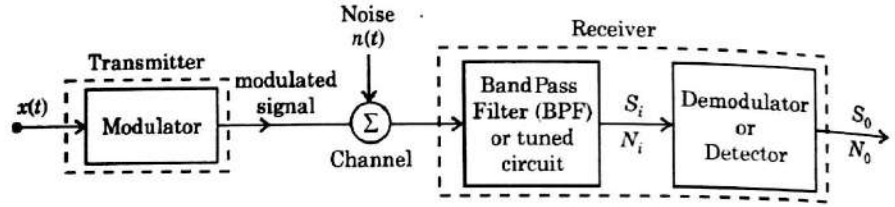


Fig. 7.21. *Generalised equivalent model of a communication system for noise calculation.*

Mathematically, $\gamma = \dfrac{\text{Output signal to noise ratio}}{\text{Input signal to noise ratio}}$

$$\gamma = \frac{S_0 / N_0}{S_i / N_i} \quad \text{...(7.8)}$$

From this expression, it is evident that a communication system with higher γ will have a better noise performance, *i.e.*, the adverse effect of noise is small.

Q.2. What is threshold effect in an envelope detector? Explain.

Ans. When a noise is large compared to the signal at the input of the envelope detector, the detected output has a message signal completely mingled with noise. It means that if the input signal-to-noise ratio (S_i/N_i) is below a certain level, called *threshold level*, the noise dominates over the message signal, *Threshold* is defined as value of the input *signal to noise ratio* (S_o/N_o) *below which the output signal to noise ratio* (S_i/N_i) *deteriorates much more rapidly than the input* signal to noise ratio. The threshold effect starts in an envelope detector whenever the carrier power-to-noise power ratio approaches unity or less. It is important to state that the threshold effect is the property of an envelope detector, on the other hand, *such an effect is not observed in a coherent detector.*

Q.3. Define coded and uncoded system.

Ans. The modulation systems can be boadly classified into two classes as under:

(*i*) uncoded systems, and

(*ii*) coded systems

In uncoded systems one symbol in message space is transformed into one symbol in the modulated signal space. Thus, for AM each possible amplitude of the original message is transformed into a particular amplitude of the modulated signal. Similarly, for PPM each amplitude of the message results in a particular displacement of the pulse position. In coded systems (such as PCM), however, each message symbol or amplitude is transformed into a number of signal symbols. It has been shown that only coded systems can achieve the most efficient exchange bandwidth for signal to noise ratio that is theoretically possible. The of uncoded systems (such as FM or PPM) are inherently incapable of efficient exchange of the bandwidth and signal-to-noise ratio. We shall now consider PCM from the point of view of exchar ge of bandwidth and signal-to-noise ratio.

Q.4. Briefly compare Coded and Uncoded Systems.

Ans. (*i*) In uncoded system, (FM and PPM), we have observed that the improvement in S/N voltage ratio is linear with bandwidth. On the other hand for the coded system (PCM), the improvement of S/N voltage ratio was observed to be exponential with bandwidth. Thus trading bandwidth and S/N ratio is much more efficient in coded systems than in uncoded systems. Coded systems are inherently capable of better transmission efficiency than uncoded systems.

(*ii*) For uncoded systems, doubling the bandwidth doubles the S/N voltage ratio. But in PCM, doubling the bandwidth allows twice the number of pulses to be transmitted. If we double the number of pulses, the possible patterns that can be formed increases as a square. For example, a group of 2 binary pulses can form 4 patterns. But a group of 4 binary pulses can from 16 patterns. Hence, doubling the bandwidth squares the quantized levels. Similarly, tripling the bandwidth will allow three times the number of pulses. But, tripling the number of pulses increases the number of patterns to the power 3.

In general, increasing the bandwidth n times increases the number of patterns exponentially (to the nth power). This allows quantized levels to increase exponentially in a given range. Obviously, Δv, the quantum gap, decreases exponentially with the bandwidth. The quantizing noise power (the output noise power) is $(\Delta v/12)^2$ and reduces exponentially. This causes the S/N ratio to improve exponentially.

The most important advantage of a coded system is it convenience for long distance communication by using several repeater stations. At each repeater station, the signal mixed with noise is cleaned up by regenerating the pulses. Thus, at each repeater station a noise-free signal is transmitted. In effect, this implies that for PCM, one needs to be concerned only about the noise on the link between the repeater stations. The noise does not accumulate throughout the transmission path. Thus, the transmission requirements for PCM are almost independent of the total length of the system.

REVIEW QUESTIONS

1. Explain the bandpass noise model.
2. Explain, how noise can be calculated in a communication system.
3. Calculate the figure of merit γ for a DSB-SC system.
4. Calculte the figure of merit γ for a SSB-SC system.
5. Evaluate the noise performance of FM system.

NUMERICAL PROBLEMS

1. Rewrite in terms of $\gamma_p = S_p/(\eta B)$, where S_p is the peak envelope power of the DSB of AM signal.

   [Ans. DSB: $\left(\frac{S}{N}\right)_o = \frac{1}{2} S_X \gamma_p$ AM: $\left(\frac{S}{N}\right)_o = \frac{1}{8} S_X \gamma_p$]

2. An AM receiver operates with a tone modulation, and the modulation index μ = 0.3. The message signal is 20 cos 1000π*t*.
   (*i*) Calculate the output SNR relative to the baseband performance.
   (*ii*) Calculate the improvement (in decibels) in the output SNR that results if m is increased from 0.3 to 0.7.

   [Ans. (*a*) $\left(\frac{S}{N}\right)_o = 0.043\gamma$, (*b*) 6.6 dB]

3. An AM system with envelope detection is operating at threshold. Determine the power gain in decibels needed at the transmitter to produce (S/N)° = 30 dB for tone modulation with m = 1.

   [Ans. ≡ 22 dB]

4. An AM system with envelope has $(S/N)_o = 30$ dB and tone modulation with $m = 1$ with $B = 8$ kHz. If all bandwidths are increased accordingly while other parameters remain fixed, what is the largest usable value of B? [Ans. 1.2 MHz]
5. Determine the detector gain α_d for an SSB system. [Ans. $\alpha_d = 1$]
6. Determine the output SNR in a PM system for the modulation. [Ans. $\left(\frac{S}{N}\right)_o = \frac{1}{2}\beta^2\gamma$]
7. Determine the detection gain α_d for an FM system with $\beta = 2$. [Ans. $\alpha_d = 36$]
8. Prove that for tone modulation, FM is superior to PM by a factor of 3 from the SNR point of view.
9. For a modulating signal $x(t) = \cos^3\omega_m t$, show that PM is superior to FM by a factor of 2.25 from the output SNR point of view.
10. Consider a communication system with the following characteristics:

$$S_X = E[x^2(t)] = \frac{1}{2} \quad B = 10 \text{ kHz} \quad \frac{\eta}{2} = 10^{-12} \text{ W/Hz}$$

Transmission loss = 70 dB.

Calculate the required transmission power S_T needed to achieve $(S/N)_o = 40$ dB when the modulation is (*i*) SSB, (*ii*) AM with $m = 1$ and $m = 0.5$ (*iii*) PM with $kp = \pi$, (*iv*) FM with $D = 1$ and $D = 5$.

[Ans. (*i*) SSB, $S_T = 1$ kW; (*ii*) AM; $m = 1$, $S_T = 3$ kW; $m = 0.5$, $S_T = 9$ kW; (*iii*) PM, $S_T = 202.6$ W; (*iv*) FM ; $D = 1$, $S_T = 667$ W; $D = 5$, $S_T = 26.7$ W]

□□□

CHAPTER 8

AM and FM Receivers

Inside this Chapter

8.1. Introduction 449
8.2. Classification of Radio Receivers 450
8.3. Tuned Radio Frequency (TRF) Receiver 450
8.4. Superheterodyne Receiver 452
8.5. AM Receivers 453
8.6. Sensitivity 453
8.7. Selectivity 453
8.8. Double Spotting 454
8.9. Tracking or Tuning of a Superheterodyne Receiver 454
8.10. Image Frequency and its Rejection 455
8.11. AM Superheterodyne Receiver (Description of Various Blocks) 457
8.12. RF Amplifier 457
8.13. A Series Resonance Circuit 459
8.14. Characteristics of a Series Resonant Circuit 460
8.15. A Parallel Resonance Circuit or Tuned Circuit 461
8.16. Characteristics of Parallel Resonant or Tuned Circuit 462
8.17. The Single Tuned Voltage Amplifier 464
8.18. The Frequency Response of a Single Tuned Voltage Amplifier 465
8.19. Limitation of a Single Tuned Voltage Amplifier 466
8.20. Double Tuned Voltage Amplifier 466
8.21. The Frequency Response of Double Tuned Voltage Amplifier 467
8.22. Stagger Tuned Voltage Amplifier 467
8.23. Local Oscillator 468
8.24. Linear Diode Detector with π-Filter and Simple AGC 473
8.25. Delayed AGC 473
8.26. Amplified and Delayed AGC 474
8.27. AGC Characteristics 475
8.28. LIC Superheterodyne AM Receivers 475
8.29. AM Transmitter 476

8.1. Introduction

We know that in a communication system, a radio transmitter radiates or transmits a modulated carrier signal. This modulated carrier signal travels down the channel *i.e.* transmission medium

and reaches at the input of radio receiver. This means that the modulated carrier signal is picked up by the antenna of the radio receiver. This modulated signal so received is generally very weak. Therefore, inside the receiver, this weak signal is first amplified in an R.F. (radio frequency) amplifier stage of the radio receiver. Also, since the received modulated signal contains a lot of noise or unwanted signals at adjacent frequencies, it must be selected and the noise must be rejected. Finally, in receiver, the R.F. carrier or modulated signal must be demodulated to get back the original modulating or baseband signal. Further, since the demodulated or detected signal (*i.e.* audio signal in case of broadcast receiver) is generally weak, it has to be amplified in one or more stages of audio amplifier.

From the above discussion, we can summarize the main functions of a radio receiver as under:

(*i*) intercept the incoming modulated signal (*i.e.* electromagnetic waves) by the receiving antenna.

(*ii*) select the desired signal and reject the unwanted signals.

(*iii*) amplify this selected R.F. signal.

(*iv*) detect the modulated signal to get back the original modulating or baseband signal.

(*v*) amplify the modulating frequency signal.

This means that a radio receiver is an electronic equipment which picks up the desired signal, rejects the unwanted signals, amplifies the desired signal, demodulates the modulated signal to get back the original modulating frequency signal.

8.2. Classification of Radio Receivers

We can classify the radio receiver in two ways as under:

(A) Depending upon the applications, the radio receivers may be classified as follows:

(*i*) *Amplitude Modulation (AM) Broadcast Receivers:* These receivers are used to receive the broadcast of speech or music transmitted from amplitude modulation broadcast transmitters which operate on long wave, medium wave or short wavebands.

(*ii*) *Frequency Modulation (FM) Broadcast Receivers:* These receivers are used to receive the broadcast programmes from FM broadcast transmitters which operate in VHF or UHF bands.

(*iii*) *Communication Receivers:* Communication receivers are used for reception of telegraph and short wave telephone signals. This means that communication receivers are used for various purposes other than broadcast services.

(*iv*) *Television Receivers:* Television receivers are used to receive television broadcast in VHF or in UHF bands.

(*v*) *Radar Receivers:* Radar receivers are used to receive radar (*i.e.*, radio detection and ranging) signals.

(B) Depending upon the fundamental aspects, the radio receivers may also be classified as under:

(*i*) Tuned Radio Frequency (TRF) Receivers.

(*ii*) Superheterodyne Receiver.

In fact, various forms of receivers have been proposed from time to time. However, only two of them became popular for commercial applications. These are Tuned Radio Frequency (TRF) receiver and superheterodyne receiver. Presently, the superheterodyne receiver is the most popular and most widely used. The TRF receiver was used earlier in the 1940s. The TRF receiver had some inherent drawbacks which were removed in superheterodyne receiver. Therefore, we shall start our discussion with TRF receiver and then come to the superheterodyne receiver.

8.3. Tuned Radio Frequency (TRF) Receiver

Tuned radio frequency (TRF) receiver is the simplest radio receiver. Figure 8.1 shows the block diagram of a tuned radio frequency receiver. The very first block of this receiver is an RF stage. This stage generally contains two or three RF amplifiers. Actually, these RF (radio frequency) amplifiers

are tuned RF amplifiers *i.e.* they have variable tuned circuit at the input and output sides. At the input of the receiver, there is a receiving antenna as shown in block diagram in figure 8.1. At this antenna signals from different sources (*i.e.* stations) are present. However, with the help of input variable tuned circuit of RF amplifiers the desired signal (*i.e.* station) is selected. But this selected signal is usually very weak of the order of μV. This selected weak signal is amplified by the RF amplifier (*i.e.* RF stage).

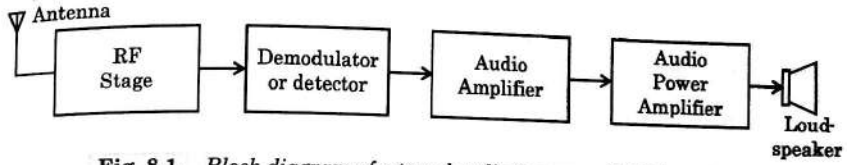


Fig. 8.1. *Block diagram of a tuned radio frequency (TRF) receiver.*

Thus, the function of RF stage is to select the desired signal and amplify it.

After this, the amplified incoming modulated signal is applied to the demodulator. The demodulator or detector demodulates the modulated signal and thus at the output of the demodulator, we get modulating or baseband signal (*i.e.* audio signal). This audio signal is amplified by audio amplifier. After that, this audio signal is further amplified by a power amplifier upto desired power level to drive the loudspeaker. The last stage of this receiver is the loudspeaker. A loudspeaker is a transducer which changes electrical signal into sound signal.

8.3.1. Drawbacks of Tuned Radio Frequency (TRF) Receiver

As discussed above, although TRF receiver is cheaper and the simplest one, it has certain drawbacks as under:

(*i*) The TRF receiver suffers from a tendency to oscillate at higher frequencies from the multistage RF amplifiers with high gain and operating at same frequency. If such an amplifier has a gain of 20,000 then if a small portion of the output leaked back to the input of the RF stage, then positive feedback and oscillation will result. This type of leakage could result from power supply coupling, stray capacitance coupling, radiation coupling or coupling through any other element common to the input and output stages. Definitely, this type of condition is undesirable for a good receiver.

This problem is also termed as instability of the receiver.

(*ii*) The selectivity of a receiver is its ability to distinguish between a desired signal and an undesired signal. The selectivity of TRF receiver is poor. In fact, it is difficult to achieve sufficient selectivity at high frequencies due to the enforced use of single-tuned circuits.

(*iii*) Another problem associated with the TRF receiver is the bandwidth variation over the tuning range. For example, in AM broadcast system, let us consider that a tuned circuit is required to have a bandwidth of 10 kHz at a frequency of 540 kHz.

According to the definition, the Quality factor Q of this tuned circuit must be

$$Q = \frac{\text{Resonance frequency}}{\text{Bandwidth}} = \frac{540}{10} = 54$$

Now, at the other end of this AM broadcast band (*i.e.* 1640 kHz), the Quality factor Q of the coil, according to above equation, must increase by a factor of 1640/535 (*i.e.* 3) to a value of 164. However, in practice due to several losses dependent upon frequency would prevent such a large increase. Thus, practically, the Quality factor Q of this tuned circuit is unlikely to exceed 120 and hence providing a bandwidth of the tuned circuit equal to

$$\Delta f = \frac{f_r}{Q} = \frac{1640}{120} = 13.8 \text{ kHz}$$

Therefore, due to this increased bandwidth of 13.8 kHz in place of a fixed bandwidth of 10 kHz, the receiver would pick up or select adjacent frequencies (i.e. stations) with the desired frequency or station. This means that the bandwidth of the TRF receiver varies with the incoming frequency.

8.4. Superheterodyne Receiver

Figure 8.2 shows the block diagram of a superheterodyne receiver. All the drawbacks in TRF receiver have been removed in a superheterodyne receiver. The basic superheterodyne receiver is most widely used. This means that the superheterodyne principle is used in all types of receiver like television receiver, radar receiver etc.

In a superheterodyne receiver, the incoming RF signal frequency is combined with the local oscillator signal frequency through a mixer and is converted into a signal of lower fixed frequency. This lower fixed frequency is known as **intermediate frequency**.

However, the intermediate frequency signal contains the same modulation as the original signal. This intermediate frequency signal is now amplified and demodulated to reproduce the original signal.

The word heterodyne stands for mixing. Here we have mixed the incoming signal frequency with the local oscillator frequency. Therefore this receiver is called superheterodyne receiver.

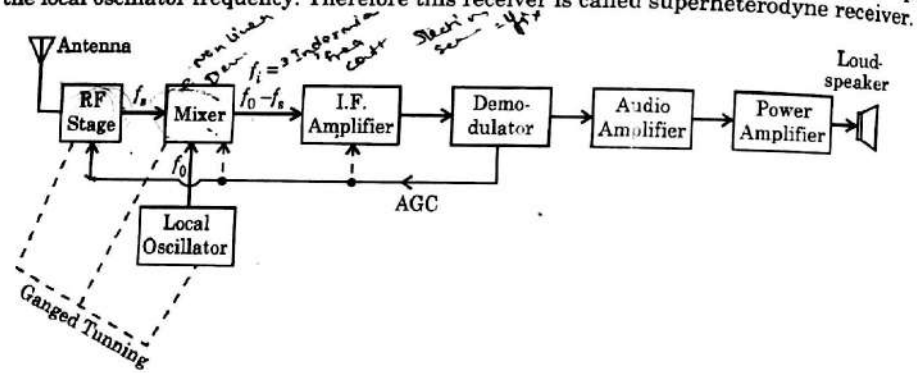


Fig. 8.2. *Block diagram of a superheterodyne receiver.*

Thus, in a superheterodyne receiver, a constant frequency difference is maintained between the local oscillator signal frequency and incoming RF signals frequency through capacitance tuning in which the capacitances are ganged together and operated by a common control knob. The intermediate frequency (IF) amplifier generally contains a number of transformers each consisting of a pair of mutually coupled tuned circuits. Thus, with this large number of double-tuned circuits, operating at a specially chosen frequency, the IF amplifier provides most of the gain (*i.e.* sensitivity) and band-width requirements (*i.e.* selectivity) of the receiver. This means that the IF amplifier determines the sensitivity and selectivity of the superheterodyne receiver.

Also, since the characteristics of the IF amplifier are independent of the incoming frequency to which the receiver is tuned, the selectivity and sensitivity of the superheterodyne receiver are quite uniform throughout its tuning range and not subject to the variations like a TRF receiver. Further since the IF amplifier works at a fixed IF frequency, the design of this system is quite easy to provide high gain and constant bandwidth.

Because of its narrow bandwidth, the IF amplifier rejects all other frequencies except intermediate frequency (IF). Actually, this rejection process reduces the risk of interference from other stations or sources. Infact, this selection process is the key to the superheterodyne receiver's exceptional performance.

After the IF amplifier, the signal is applied at the input of demodulator which extracts the original modulating signal (*i.e.* audio signal). This audio signal is amplified by an audio amplifier

to get a particular voltage level. This amplified audio signal is further amplified by a power amplifier to get a specified power level so that it may activate the loudspeaker. The loudspeaker is a transducer which converts this audio electrical signal into audio sound signal and thus the original signal is reproduced *i.e.* the original transmission is received.

The advantages of the superheterodyne receiver make it the most suitable for the majority of radio receiver applications like AM, FM, communications, single-sideband, television and even radar receiver; all use superheterodyne principle. This means that it can be considered as today's standard form of radio receiver.

8.5. AM Receivers

Because the type of receiver is almost the same for various forms of modulation or system, therefore, it is generally most convenient to explain the various principles of a superheterodyne receiver while dealing with AM receivers. Thus, with the discussion of AM receiver, a basis is formed for the more complex versions of superheterodyne receiver. In this section, let us discuss few superheterodyne receiver characteristics.

8.6. Sensitivity

The sensitivity of a radio receiver may be defined as its ability to amplify weak signals. It is generally, defined in terms of the voltage which must be applied at receiver input terminals to provide a standard output power measured at the output terminals. For AM broadcast receivers, several relevant quantities have been standardized. A signal modulated by a 400 Hz sine wave and modulation index of 30% is applied through standard coupling network known as a **dummy antenna**.

In addition to this, the loudspeaker is replaced by an equivalent load resistance. After this the output is measured across this resistance and it must be equal to the standard value of 50 mW.

Sensitivity is also expressed in microvolts or in decibels below 1 volt and is measured at three points along the tuning range when a production receiver is lined up.

Figure 8.3 shows the sensitivity curve over the tuning band. At 1 MHz, this particular receiver has a sensitivity of 12.7 μ volt or −98 dB volt. Sometimes, the sensitivity defintion is extended, and the manufacture of this receiver may quote it to be, nor merely 12.7 μV, but "12.7 μV for an SNR of 20 dB in the output of the receiver".

However, for professional receivers, the sensitivity is generally quoted in terms of signal power required to produce a minimum acceptable output signal with a minimum acceptable output noise level. Few factors determining the sensitivity of a superheterodyne receiver are as under:

(*a*) The gain of the IF amplifiers.

(*b*) The gain of the RF amplifiers.

(*c*) The noise figure of the receiver.

It may be noted that the typical values of sensitivity are 150 μ volt for small broadcast band receivers, and 1 μ volt or below for high quality communication receiver in the HF band.

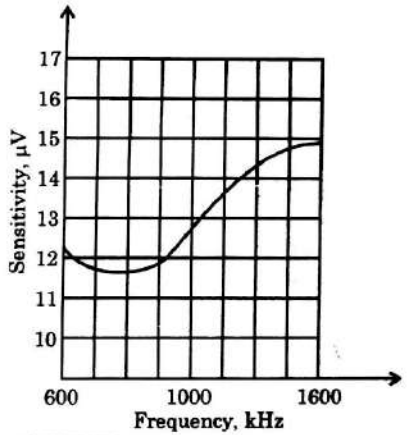


Fig. 8.3. *A typical sensitivity curve for a good domestic receiver.*

8.7. Selectivity

The selectivity of a receiver may be defined as the ability to reject unwanted signals. It also expresses the attenuation that the receiver offers to signal at frequencies adjacent to the one to which it is tuned. It is generally, expressed as a curve as shown in figure 8.4. In selectivity measurement, the

frequency of the generators is varied to either side of the frequency to which the receiver is tuned. Naturally, the output of the receiver falls since the input frequency is not incorrect. Thus, the input voltage must be increased until the output is the same as it was originally. The ratio of the voltage required of resonance to the voltage required. When the generator is tuned to the receiver's frequency it is calculated at a number of points and then plotted in decibels to give a curve as shown in figure 8.4. It may be noted that selectivity depends upon the following factors:

(i) Selectivity varies with receiving frequency and becomes somewhat worse when the receiving frequency is raised.

(ii) In general, it is mainly determined by the response of the IF section, with the mixer and RF amplifier input circuits playing a small but significant part.

(iii) Selectivity is the main factor which determines the adjacent channel rejection of a receiver.

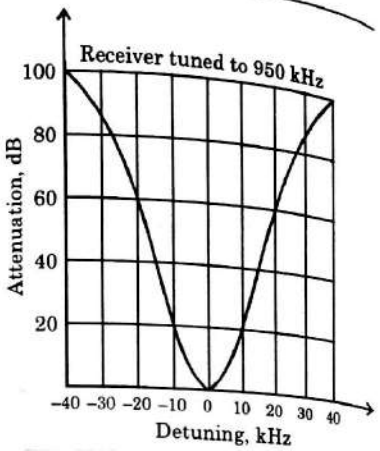


Fig. 8.4. A Typical selectivity curve

8.8. Double Spotting

When a receiver picks up the same short wave station at two nearby points on the receiver dial, the double spotting phenomenon takes place. The main cause for double spotting is poor front-end selectivity, *i.e.*, inadequate image-frequency rejection. The front-end of the receiver does not select different adjacent signal very well. However, fortunately, the IF stage.

The adverse effect of double spotting is that a weak station may be marked by the reception of a nearby strong station at the spurious point on the dial. On the other hand, double spotting may be used to calculate the IF of an unknown receiver. The spurious point on the dial is precisely $2f_i$ below the correct frequency. If image-frequency rejection is improved, then certainly there will be a corresponding decrease in the double spotting occurrence.

8.9. Tracking or Tuning of a Superheterodyne Receiver

In a superheterodyne receiver, the local oscillator frequency is made to track with the tuned circuits which are tuned to the incoming signal frequency in order to make a constant frequency difference at the output of mixer. For general AM broadcast system, the intermediate frequency (IF) is 455 kHz. This indicates that the local oscillator should always be set at a frequency which is 455 kHz above the incoming signal frequency. For purpose the front end of the receiver tuned circuits are made to track together simply by mechanically linked or ganged capacitors. A ganged capacitor has three capacitor sections, one each for the RF amplifier, mixer and the local oscillator. In addition to this, small variable capacitances known as trimmers are connected in parallel with each section. These capacitances can be adjusted for proper operation at highest frequency. However for lowest frequency adjustment, small variable capacitors known as padders are connected in series with the inductor of the tank circuit.

The various tuned circuits are mechanically coupled so that only one tuning control and dial are required. This means that no matter what is the incoming signal frequency, the RF and mixer input tuned circuits must be tuned to it. The local oscillator must simultaneously be tuned to a frequency which is precisely higher than the signal frequency by the intermediate frequency. However, any error that may exist in the frequency difference would result in an incorrect frequency being fed to the intermediate frequency (IF) amplifier. This error must naturally be avoided. Such type of errors are known as tracking errors. These tracking errors result in stations appearing away from their correct position on the dial.

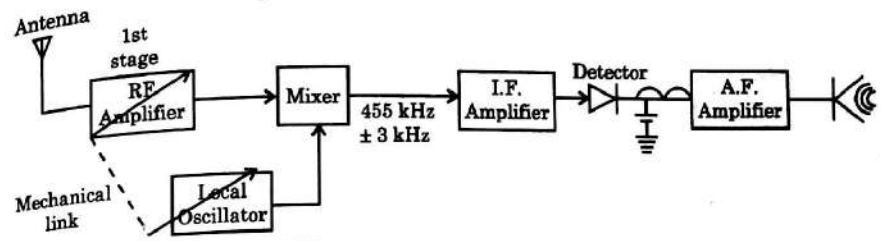


Fig. 8.5. *Tuning of a superheterodyne receiver.*

The various tuned circuits are mechanically coupled so that only one tuning control and dial are required. This means that no matter what is the incoming signal frequency, the RF and mixer input tuned circuits must be tuned to it. The local oscillator must simultaneously be tuned to a frequency which is precisely higher than the signal frequency by the intermediate frequency. However, any error that may exist in the frequency difference would result in an incorrect frequency being fed to the intermediate frequency (IF) amplifier. This error must naturally be avoided. Such type of errors are known as tracking errors. These tracking errors result in stations appearing away from their correct position on the dial.

It is quite possible to keep the maximum tracking error below 3 kHz. A value as low as this is quite acceptable.

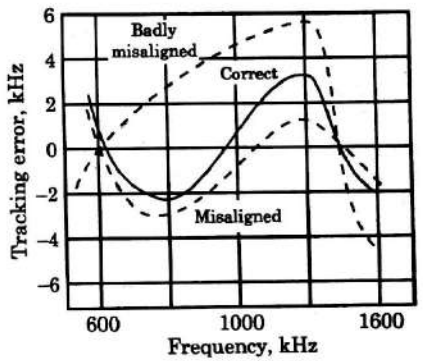


Fig. 8.6. *Tracking curves.*

8.10. Image Frequency and its Rejection

We know that a superheterodyne receiver is a better receiver than a Tuned Radio Frequency (TRF) receiver. However, a superheterodyne receiver suffers from a major drawback known as Image Frequency problem. This problem of image frequency is inherent to a superheterodyne receiver and arises because of the use of heterodyne principle. In fact, the frequency conversion process carried out by the local oscillator and the mixer often allow an undesired frequency in addition to the desired incoming frequency.

In a standard broadcast receiver, the local oscillator frequency is always made higher than the incoming signal frequency. It is kept equal to the signal frequency plus the intermediate frequency (IF).

Mathematically,

$$f_o = f_s + f_i \qquad ...(8.1)$$

where f_o = local oscillator frequency

f_s = desired incoming frequency

f_i = intermediate frequency

From equation (8.1), we have

$$f_i = f_o - f_s$$

Hence, the intermediate frequency is the difference between the local oscillator frequency and the signal frequency.

Now, if a frequency f_{si} manages to reach the mixer, such that

$$f_{si} = f_o + f_i \qquad ...(8.2)$$

then this frequency f_{si} would also produce f_i when it is mixed with f_o. This undesired or spurious intermediate frequency signal will also be amplified by the IF stage and thus would cause

interference. This has the effect of two sources or stations being received simultaneously. This situation is obviously undesirable.

The term f_{si} is known as the image frequency and is defined as the signal frequency plus twice the intermediate frequency.

Putting the value of f_o in equation (8.2) from equation (8.1), we get

$$f_{si} = f_o + f_i$$

$$f_{si} = f_s + f_i + f_i$$

or

$$f_{si} = f_s + 2f_i \quad \text{...(8.3)}$$

Thus this spurious frequency signal cannot be distinguished by the IF stage and hence would be treated in the same manner as the desired frequency signal.

The rejection of an image frequency signal by a single tuned circuit may be defined as the ratio of the gain at the signal frequency to the gain at the image frequency. This is given as

$$\alpha = \sqrt{1+Q^2\rho^2} \quad \text{...(8.4)}$$

Here

$$\rho = \frac{f_{si}}{f_s} - \frac{f_s}{f_{si}} \quad \text{...(8.5)}$$

and Q = Quality factor of the tuned circuit in a loaded condition.

If the receiver has an RF stage, then there is only a single tuned circuit and the rejection will be calculated using equation (8.4). However if the receiver has an RF stage then there are two tuned circuits both tuned to f_s. The image frequency rejection of each stage will be calculated by using equation (8.4). The total or overall rejection will be the product of the two.

The image-frequency rejection of the receiver depends upon the front-end selectivity of the receiver. The rejection of image frequency must be achieved before the IF stage. Once an undesired or spurious frequency enters the first IF amplifier, it would become impossible to remove it from the desired signal.

It may be observed that if $\frac{f_{si}}{f_s}$ is large as is the case for AM broadcast band the use of an RF stage is not necessary for good image frequency rejection. However, it would become essential above about 3 MHz.

Example 8.1. For a broadcast superheterodyne AM receiver having no RF amplifier, the loaded Quality factor Q of the antenna coupling circuit is 100. Now if the intermediate frequency is 455 kHz, then determine the following:

(i) the image frequency and its rejection ratio at an incoming frequency of 1000 kHz.

(ii) the image frequency and its rejection ratio at an incoming frequency of 25 MHz.

Solution: Given that

$$Q = 100$$

and

$$f_i = 455 \text{ kHz}$$

$$f_s = 1000 \text{ kHz}$$

The image-frequency is given as

$$f_{si} = f_s + 2f_i = 1000 + 2 \times 455$$

or

$$f_{si} = 1000 + 910 = 1910 \text{ kHz}$$

Further,

$$\rho = \frac{f_{si}}{f_s} - \frac{f_s}{f_{si}} = \frac{1910}{1000} - \frac{1000}{1910}$$

or

$$\rho = 1.910 - 0.524 = 1.386$$

Since the given receiver has no RF amplifier, therefore there is only single tuned circuit.

The rejection ratio is given as

$$\alpha = \sqrt{1+Q^2\rho^2} = \sqrt{1+(100)^2 \times (1.386)^2}$$

or

$$\alpha = \sqrt{1+(138.6)^2} = 138.6 \qquad \textbf{Ans.}$$

(ii) For second case, it is given that

$$Q = 100$$

$$f_i = 455 \text{ kHz} = 0.455 \text{ MHz}$$

and

$$f_s = 1000 \text{ kHz}$$

The image frequency is given as

$$f_{si} = f_s + 2f_i = 25 + 2 \times 0.455$$

or

$$f_{si} = 25.91 \text{ MHz}$$

$$\rho = \frac{f_{si}}{f_s} - \frac{f_s}{f_{si}} = \frac{25.91}{25} - \frac{25}{25.91}$$

or

$$\rho = 1.0364 - 0.9649 = 0.0715$$

The rejection ratio α is given as

$$\alpha = \sqrt{1+Q^2\rho^2} = \sqrt{1+(100)^2 \times (0.0715)^2}$$

$$\alpha = \sqrt{1+(7.15)^2} = 5.22 \qquad \textbf{Ans.}$$

8.11. AM Superheterodyne Receiver (Description of Various Blocks)

In this section, let us discuss the principle and working of different blocks used in a superheterodyne AM receiver.

8.12. RF Amplifier

R.F. amplifier is a small signal tuned amplifier with tuned circuits both in the input side and the output side. Both these input and output tuned circuits are tuned to the desired incoming carrier frequency. Accordingly the tuned circuits select the desired carrier frequency and reject all undesired frequencies including the image frequency. Hence the RF amplifier provides image frequency rejection. Also the gain provided by the RF amplifier will result in improved signal/noise ratio in the output of the receiver. This is due to the fact that the incoming weak signal is raised to a higher level with the help of RF amplifier before it is fed at the input of the mixer stage which contributes most of the noise generated in the receiver. However, if the incoming weak signal is fed directly to the frequency mixer, the signal/noise ratio at the output of the mixer stage is quite poor and hence any amount of subsequent amplification cannot improve S/N ratio. Thus the one important function of the RF amplifier is to improve S/N ratio.

There are some cases also where an RF amplifier is not used in the receiver rather its use is uneconomical there. The best example of this kind of receiver is a domestic receiver used in a high-signal-strength area like a metropolitan city. Since in a metropolitan city like Delhi, several stations are situated and in such places strength is obviously very high and thus, there is no need for the use of RF amplifier. In such cases the tuned circuit connected to the antenna is the actual input circuit of the mixer.

However, a receiver having an RF amplifier is obviously superior in performance to a receiver without RF amplifier.

We may summarize the advantages of RF amplifier as under:

(i) Greater gain *i.e.* better sensitivity

(ii) Improved rejection of adjacent undesired signals *i.e.* better selectivity

(iii) Improved signal/noise ratio
(iv) Improved image frequency rejection
(v) Improved coupling of the receiver to the antenna
(vi) Prevention of reradiation of the local oscillator voltage through the antenna
(vii) Prevention of spurious frequencies from entering the mixer and heterodyne to produce interfering frequency equal to IF

Circuit of RF Amplifier

Figure 8.7 shows the circuit diagram of one stage RF amplifier using an NPN transistor. It is a small signal amplifier using parallel tuned circuit as the load impedance. This parallel output tuned circuit is tuned to the incoming desired signal frequency. The output from the receiving antenna is transformer coupled to the base of the transistor. The secondary coil of the input tuned circuit is tuned to the incoming desired signal frequency with the help of ganged tuning capacitor. In fact the tuning capacitors *i.e.* variable air capacitors in the input side and the output side of the RF amplifier are ganged together. In addition to this, small trimmer capacitors are connected in shunt with these tuning capacitors for the purpose of RF alignment.

A self-bias is provided with the help of resistors R_1 and R_2 and $R_E - C_E$ assembly. A decoupling network consisting of resistor R_b and capacitor C_b is placed in the collector supply lead.

The amplified R.F. signal developed across the collector tuned circuit is coupled through a step down transformer to the input of the frequency mixer. This step down transformer provides the impedance matching between the high impedance of the RF amplifier collector circuit and the low impedance of the base to emitter circuit of the following stage. Also the collector is connected to a suitable point on the primary of the output transformer so that load impedance to the collector is optimum.

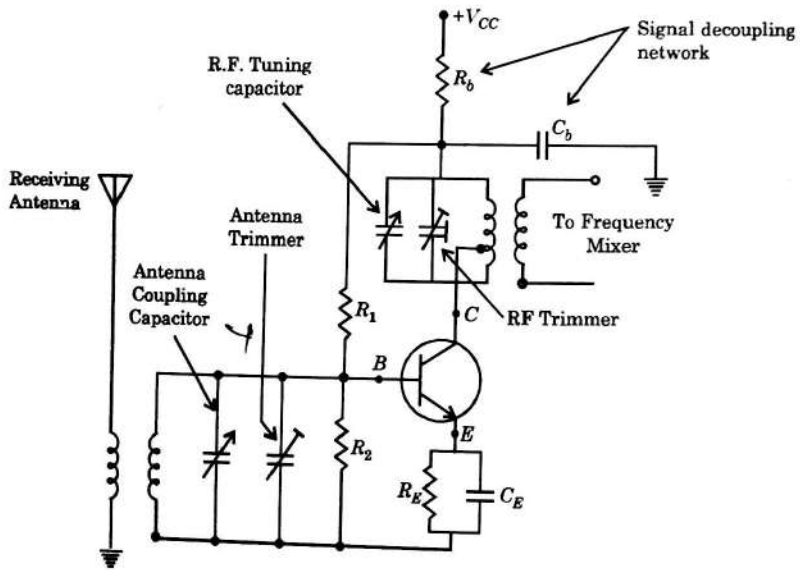


Fig. 8.7. *Circuit diagram of RF amplifier.*

8.13. A Series Resonance Circuit

As discussed in last article, in a series resonance circuit an inductor (L) and a capacitor (C) are connected in series across an ac source as shown in figure 8.8. The frequency of the supply source can be varied. At any given frequency f, the inductive reactance is given by

$$X_L = 2\pi fL \qquad ...(8.6)$$

and the capacitive reactance is given by

$$X_C = \frac{1}{2\pi fC} \qquad ...(8.7)$$

The total impedance of the circuit is given as

$$Z_S = \sqrt{R^2 + (X_L - X_C)^2} \qquad ...(8.8)$$

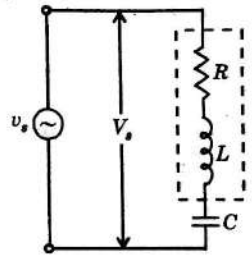


Fig. 8.8. *A series resonance circuit.*

It is evident from equations (8.6) and (8.7) that the value of X_L and X_C depends upon the source supply frequency *i.e.*, when frequency increases, the value of X_L increases since X_L is proportional to f whereas the value X_C decreases with the increase in frequency as X_C is inversely proportional to the frequency. The variation in X_L and X_C with frequency has been shown in figure 8.9. At a particular frequency f_r, we find that $X_L = X_C$ (*i.e.*, Point where the two curves intersect each other. This frequency is known as **resonant frequency.** Figure 8.9 shows variation of X_L and X_C with frequency.

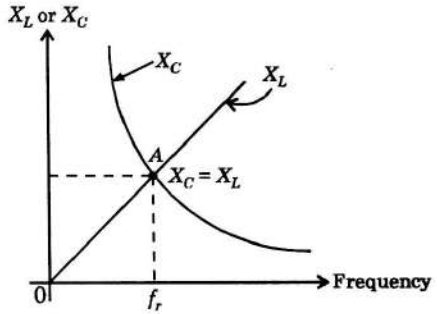


Fig. 8.9. *Variation of X_L and X_C with frequency.*

Resonant Frequency

The frequency at which an LC circuit shows resonance *i.e.* the resultant current drawn by the circuit comes in phases with source voltage is known as **resonant frequency f_r.**

If f_r is the resonant frequency, then in LC series circuit the condition of resonance is obtained when

$$X_L = X_C$$

Substituting the values of X_L and X_C, we get

or $$2\pi f_r L = \frac{1}{2\pi f_r C} \quad \text{or} \quad f_r = \frac{1}{2\pi\sqrt{LC}} \qquad ...(8.9)$$

At resonant frequency, the value of circuit impedance will be

$$Z_r = R \qquad (\text{since } X_L - X_C = 0) \qquad ...(8.10)$$

This is minimum value of impedance. However, the current drawn by the circuit from the source will be maximum at resonance *i.e.*

$$I_r = \frac{V_S}{R} \qquad ...(8.11)$$

From the above discussions, the following points may be observed about the series resonance circuit:

(i) $X_L = X_C$ (ii) $f_r = \frac{1}{2\pi\sqrt{LC}}$

(iii) $Z_r = R$ (iv) $I_r = \frac{V}{R}$

8.14. Characteristics of a Series Resonant Circuit

In this section let us consider some of the important characteristics of a series resonant circuit as under:

(i) Effect of Frequency on the Impedance of the Series Resonant Circuit

As discussed in last article that at resonant frequency f_r, the value of inductive reactance X_L is equal to the value of capacitive reactance (*i.e.*, $X_L = X_C$). Therefore, the value of impedance at this frequency is minimum and equal to resistance R of the circuit, *i.e.*

$$Z_S = \sqrt{R^2 + (X_L - X_C)^2}$$

where $X_L = 2\pi f L$

and $X_C = \dfrac{1}{2\pi f C}$

At $f_r, X_L = X_C$

Thus, $Z_r = \sqrt{R^2} = R$

If frequency decreases than resonant frequency, the value of X_L decreases (as $X_L = 2\pi fL$) but the value of X_C increases (as $X_C = 1/2\pi fC$). Therefore, the impedance of the circuit increases. On the other hand, if frequency increases more than the resonant frequency, the value X_L increases but the value of X_C decreases. Hence, it is observed that the value of circuit impedance is minimum at resonant frequency and its value increases when frequency decreases or increases than the resonant frequency. This effect has been illustrated by the curve plotted between impedance of series circuit (Z_S) and the supply frequency (f), as shown in figure 8.10.

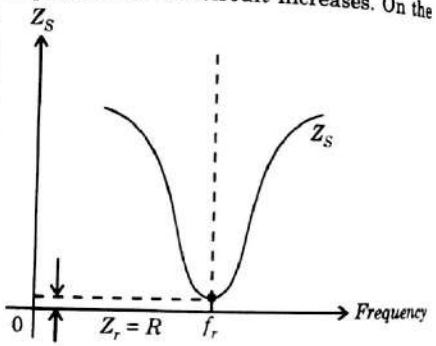


Fig. 8.10. *Plot of impedance of series circuit versus supply frequency.*

(ii) Resonance Curve

The curve plotted between current drawn by the series resonance circuit and the supply frequency is known as **resonance curve** of series resonance circuit.

As discussed earlier, at resonant frequency the circuit impedance is minimum and hence the current drawn by the circuit will be maximum. The magnitude of current will decrease as the frequency increases or decreases than the resonant frequency value. If graph is drawn between the current and supply frequency, we will get curve as shown in figure 8.11. Two curves have been drawn for different values of R (where $R_1 < R_2$). The curve reveals few points as under:

(a) The rate at which the current decreases as we move away from the resonant, frequency depends upon the resistance of the circuit. For smaller values of resistance the current falls more sharply, whereas for higher values of resistance the curve is flat.

(b) The magnitude of current is inversely proportional to resistance of the series resonant circuit. It plays an important role in the selecting property (selectivity) of the resonant circuit.

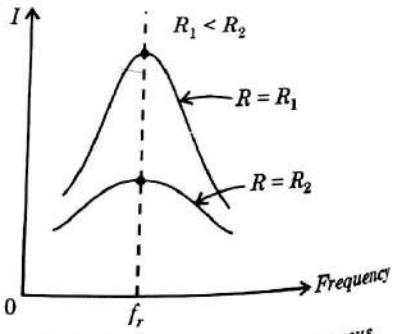


Fig. 8.11. *Plot of current versus supply frequency.*

(iii) The Quality Factor Q and its Significance

The ratio of voltage drop across L or C (*i.e.*, I_rX_L or I_rX_C) to the voltage drop across R (I_rR) at series resonance is known as **quality factor Q** of the circuit. It is also known as **figure of merit.**

It has been observed that at series resonance, the circuit draws largest current from the mains. This produces a heavy voltage drop across L or C. The factor by which the p.d. across L or C rises to that of the applied voltage is known as Q-factor of the series resonant circuit.

Mathematically, the Quality factor is given as

$$\text{Quality factor, } Q = \frac{I_rX_L}{I_rR} = \frac{X_L}{R} = \frac{\omega_r L}{R} = \frac{2\pi f_r L}{R} \quad \text{...(8.12)}$$

We know that at resonance, the current is given by

$$I_r = \frac{V}{R}$$

and

$$V_L = I_r X_L = \frac{V}{R} \times \omega_r L = V \times \frac{\omega_r L}{R} = QV \quad \text{...(8.13)}$$

Equation (8.13) shows that Q of the circuit varies inversely with the circuit resistance. In figure 8.12, variation of current with frequency has been shown for different values of the resistance. The curve shown for $R = R_1$ will represent a circuit with high Q. It shows that the current falls rapidly as we move off the resonance (for low value of R or high value of Q). This means that such a circuit provides better frequency discrimination.

Thus, it can be said that Q is the measure of ability of a resonant circuit to select a band of frequency. The higher the Q of a series resonant circuit, the greater is its ability as a frequency selective network.

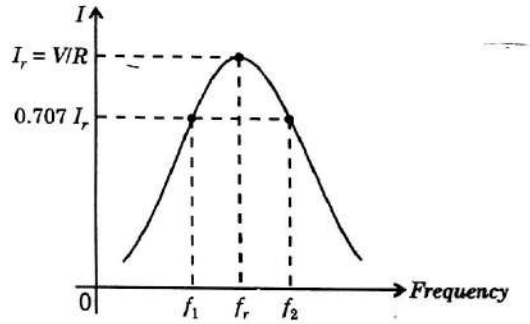


Fig. 8.12.

(iv) Bandwidth

The range of frequency over which the circuit current is equal to or more than 70.7% of maximum circuit current at resonance is called the bandwidth of a series resonant circuit.

Figure 8.12 shows a typical series resonance circuit curve. Here, the circuit current is equal to or greater than 70.7% of maximum current ($I_r = V/R$) between frequency range from f_1 to f_2.

Therefore, according to the definition

Bandwidth, $BW = f_2 - f_1$

Bandwidth represents the frequency range at which the circuit offers low impedance path. Because, the Q of the circuit determines the overall steepness of the curve, the bandwidth may also be determined in terms of the resonant frequency and Q of the circuit by the expression

$$BW = (f_1 - f_2) = \frac{f_r}{Q} \quad \text{...(8.14)}$$

This means that higher the value of Q, the smaller is the bandwidth.

8.15. A Parallel Resonance Circuit or Tuned Circuit

In a parallel resonance circuit or tuned circuit, an inductor and a capacitor are connected in parallel across an ac source as shown in figure 8.13. The frequency of the supply source can be varied. If the frequency of applied source voltage is equal to the natural or resonance frequency of L_C circuit, then electrical resonance will take place. Under such conditions, the impedance of the tuned circuit becomes maximum and the circuit draws minimum current from the signal source.

Considering figure 8.13, let the current drawn by inductor be I_L and current drawn by capacitor be I_C. Since inductor also has some resistance, it draws current I_L which lags behind the voltage by some angle (say ϕ_L). At resonance, the resultant current (I) drawn by the circuit must be in phase with the applied voltage. It is clear from the phasor diagram shown in figure 8.14 that at resonance, we have

$$I_C = I_L \sin \phi_L \qquad ...(8.15)$$

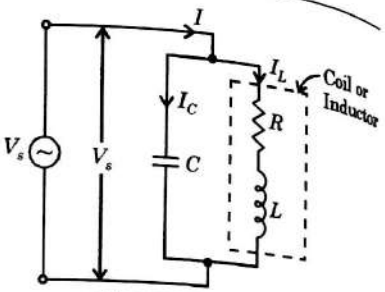


Fig. 8.13. *A parallel resonance circuit.*

Resonant Frequency

The frequency at which an LC circuit shows resonance, *i.e.* the resultant current drawn by the circuit comes in phase with the source voltage is known as **resonant frequency**.

If f_r is the resonant frequency, then in LC parallel circuit, the condition of resonance is obtained when

$$I_C = I_L \sin \phi_L$$

where, $$I_L = \frac{V}{Z_L}$$

$$\sin \phi_L = \frac{X_L}{Z_L}$$

and $$I_C = \frac{V}{X_C}$$

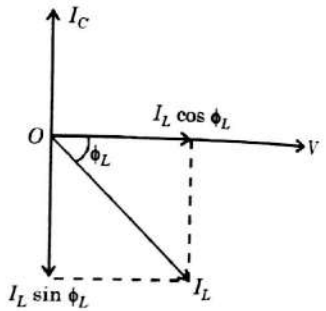


Fig. 8.14. *Phasor diagram.*

Thus, $$\frac{V}{X_C} = \frac{V}{Z_L} \times \frac{X_L}{Z_L}$$

or $$Z_L^2 = X_L X_C$$

or $$(R^2 + X^2_L) = X_L X_C$$

or $$R^2 + (2\pi f_r L)^2 = \frac{2\pi f_r L}{2\pi f_r C}$$

or $$(2\pi f_r L)^2 = \frac{L}{C} - R^2$$

or $$f_r = \frac{1}{2\pi L}\sqrt{\frac{L}{C} - R^2}$$

or $$f_r = \frac{1}{2\pi}\sqrt{\frac{1}{LC} - \frac{R^2}{L^2}} \qquad ...(8.16)$$

In practice, the resistance of inductor of coil is very small therefore, in equation (8.16) the value of R^2/L^2 can be neglected.

Hence, resonant frequency will be

$$f_r = \frac{1}{2\pi\sqrt{LC}} \qquad ...(8.17)$$

8.16. Characteristics of Parallel Resonant or Tuned Circuit

Now let us consider few important characteristics of parallel resonant circuit or tuned circuit as under:

(i) Effect of Frequency on the Impedance of the Circuit

As discussed earlier that at resonant frequency f_r the line current (*i.e.*, current at resonant) I_r is in phase with the supply voltage and its amplitude minimum. It clearly shows that at resonant frequency, the impedance of the parallel resonant circuit is maximum. This has been shown by the impedance-frequency curve in figure 8.15. However, the impedance of the circuit decreases rapidly when the frequency is changed above or below the resonant frequency f_r. This is an important characteristic of a parallel tuned circuit which provodes it the selective properties, *i.e.* to select the resonant frequency and reject all other frequencies.

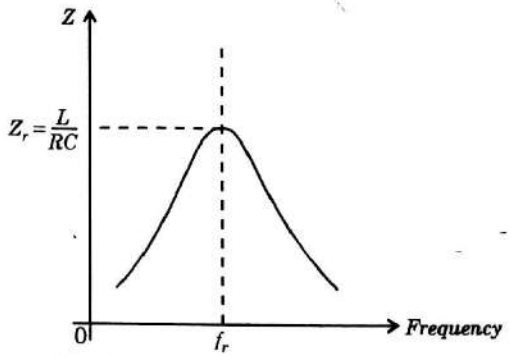


Fig. 8.15. *Impedance-frequency curve.*

Mathematically, we can write

At resonance, line current is given by

$$I_r = I_L \cos \phi_L$$

or $$\frac{V}{Z_r} = \frac{V}{Z_L} \times \frac{R}{Z_L}$$

$$Z_r = \frac{Z_L^2}{R}$$

But since $$Z_L^2 = X_L X_C = \frac{L}{C}$$

Therefore, $$Z_r = \frac{L/C}{R}$$

Hence, circuit impedance at resonance $Z_r = \frac{L}{CR}$.

Z_r will be in ohms when L, C and R are in henry, farads and ohms respectively.

(ii) Resonance Curve

The curve plotted between the current drawn by the parallel by the parallel resonance circuit and the supply frequency is known as resonance curve of parallel resonance circuit.

The line current drawn by the parallel resonance circuit at resonant frequency is given by expression

$$I_r = \frac{V}{Z_r}$$

Because Z_r is very high, the line current is quite small (*i.e.*, minimum) at the resonant frequency. Its value increases rapidly when the frequency is varied above or below the resonant frequency f_r. This has been shown in the curve drawn figure 8.16.

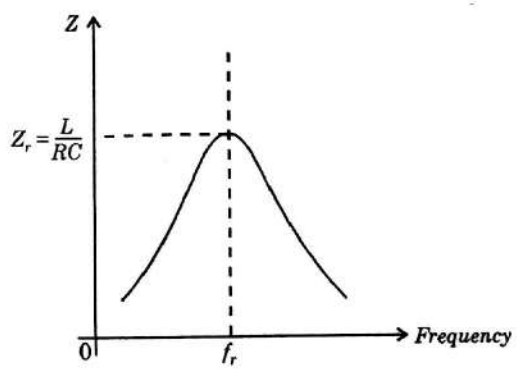


Fig. 8.16. *Impedance-frequency curve (i.e., resonance curve)*

(iii) Quality Factor Q

The most important characteristic of parallel tuned circuit is its selectivity. The sharpness of the resonance (impedance v/s frequency) curve denotes its property of selectivity and the sharpness depends upon the resistance of the coil. Smaller the resistance of the coil, sharper will be the resonance curve. This is due to the reason that a small resistance consumes less power and draws smaller line current.

The ratio of inductive reactance of the coil at resonance to its resistance by is called **quality factor Q or figure of merit.**

Mathematically, $$Q = \frac{X_L}{R}$$

or $$Q = \frac{2\pi f_r L}{R}$$

Thus quality factor is just a measure of quality of the tuned circuit.

(iv) Bandwidth

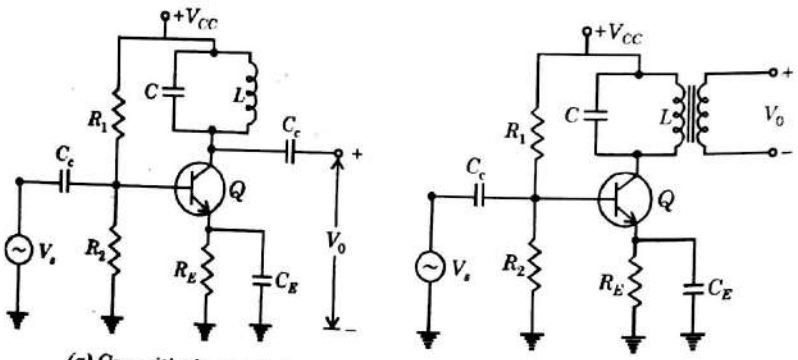
The concept of bandwidth can be explained in a similar manner as for series resonance circuit discussed earlier. The only difference is that in series resonant circuit, the response is taken in terms of the line current whereas in parallel resonant circuit, it is taken in terms of impedance.

Also, the relation between resonant frequency and bandwidth will be

$$BW = \frac{f_r}{Q}$$

8.17. The Single Tuned Voltage Amplifier

Figure 8.17 shows the circuit of a single-tuned voltage amplifier which makes uses of a bipolar junction transistor (BJT). The circuit shown in figure 8.17(*a*) is known as capacitively coupled tuned amplifier since the output is taken through a coupling capacitor (C_C). Similarly, the circuit shown in figure 8.17 known as inductively coupled tuned amplifier since the output is taken across an inductor. Both these circuits consist of a transistor amplifier and tuned circuit as the load. The values of capacitance (*C*) and inductance (*L*) of the tuned circuit are selected in such a manner that the resonant frequency of the tuned circuit is equal to the frequency to be selected and amplified. The resistors R_1, R_2 and R_E form biasing resistors. These resistors provides the d.c. operating currents and voltages for the transistor.



(a) Capacitively coupled (b) Inductively coupled

Fig. 8.17. *A single-tuned voltage amplifier.*

It may be noted that the other active devices like field-effect transistor (FETs) and metal-oxide semiconductor field-effect transistors (MOSFET's) can also be used in tuned voltage amplifier circuit instead of bipolar junction transistor (BJT). Figure 8.18 shows the circuit of single-tuned voltage amplifier which makes use of a field affect transistor (FET).

The working of tuned voltage amplifier can be understood by considering a radio frequency signal, to be amplified, applied at the input of the amplifier. The resonant frequency of the tuned circuit, is made equal to the frequency of the input signal by changing the value of capacitance (*C*) or inductance (*L*). When the frequency of the tuned circuit becomes equal to that of the input signal, a large signal appears across the output terminals. It can be noted that if the input signal is a complex wave (*i.e.* it contains many frequency components), in that case the signal with frequency equal to the resonant frequency would be **amplified** whereas all the other frequencies will be rejected by the tuned circuit.

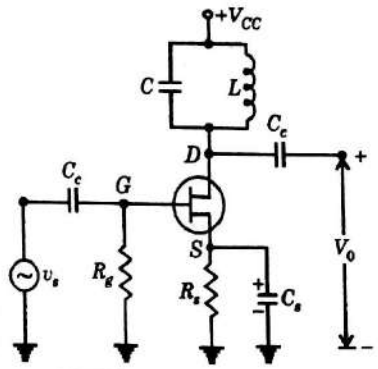


Fig. 8.18. *A single-tuned voltage amplifier using field-effect transistor (FET)*

8.18. The Frequency Response of a Single Tuned Voltage Amplifier

As already discussed in the chapter on single stage transistor amplifier that voltage gain of an amplifier depends upon current gain (β), input resistance (R_i) and effective or a.c. load resistance. Thus, the voltage gain is given by expression

$$A_V = \beta \times \frac{r_L}{R_i}$$

We know that the ac load resistance of a parallel resonant resonate circuit (*i.e.* tuned circuit) is expressed as

$$r_L = Z_P = \frac{L}{CR}$$

here L = Value of inductance,

C = Value of capacitance, and

R = Value of effective resistance of the inductor.

Therefore, voltage gain of a tuned voltage amplifier is given by relation

$$A_V = \beta \times \frac{\frac{L}{CR}}{R_i}$$

We also know that the value of the quantity L/CR (known as impedance of the tuned circuit) is very high at the resonant frequency and it decreases as the frequency changes above or below the resonant frequency f_r. Therefore, voltage gain of a tuned amplifier is very high at the resonant frequency and decreases as the frequency changes above or below the resonant frequency.

The above facts have been shown in the form of a voltage gain versus frequency plot shown in figure 8.19. Such a plot is known as frequency response curve of a tuned voltage amplifier.

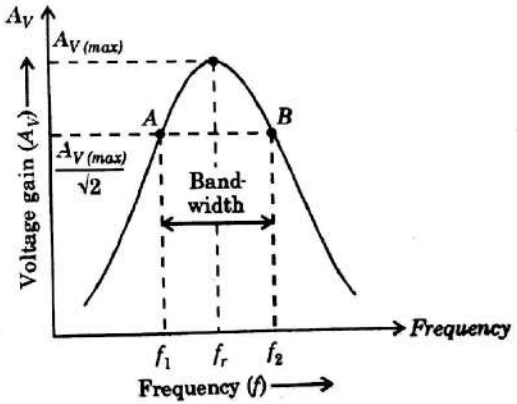


Fig. 8.19. *The frequency response curve*

The bandwidth (BW) of an amplifier is equal to the frequency difference between the points A and B on either side of the resonant frequency, where the value of voltage gain drops to $1\sqrt{2}$ of its maximum value at resonance. Thus, bandwidth will be

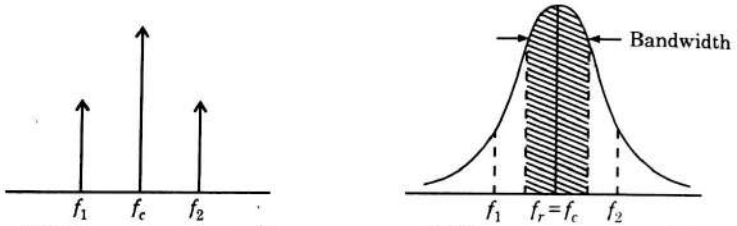
$$BW = \Delta f = f_2 - f_1 = \frac{f_0}{Q_0}$$

where Q_r is the quality factor (or Q-factor) of the tuned circuit.

8.19. Limitation of a Single Tuned Voltage Amplifier

The tuned voltage amplifiers are used in the wireless communication systems. Usually, the communication transmitters do not transmit the audio (*i.e.* voice or music) signal directly. Instead, the audio signal is first raised to some high frequency level. This high frequency is in the radio frequency (RF) range and it serves as carrier of the audio signal. The process of raising the audio signal to the radio frequency level is called **modulation**. The communication transmitters transmit the modulated wave.

The modulated wave has a relatively narrowband of frequencies centred around the carrier frequency (f_c) as shown in figure 8.20(*a*). The frequencies (f_1 and f_2) in the narrowband actually represent the information and hence is an important part of transmitted signal.



(*a*) *Frequency spectrum of the transmitted signal*

(*b*) *Frequency response of the tuned voltage amplifier*

Fig. 8.20.

The tuned voltage amplifier, in communication receivers is used to select the desired carrier frequency and amplifying the complete band of frequencies around the selected carrier frequency. In other words, tuned voltage amplifiers are required to be highly selective. But the high selectivity requires a tuned circuit with a high Q-factor. We also know that a high Q circuit will give a high voltage gain. But, at same time, it will give reduced bandwidth (because bandwidth is reciprocal to the Q-factor). It means that a tuned voltage amplifier with reduced bandwidth may not be able to amplify equally the complete band of the transmitted signal as shown in figure 8.20(*b*). In other words, narrow bandwidth or smaller passband of the amplifier will result in a poor reproduction of the audio signal. It is the major limitation of a single tuned voltage amplifier and is overcome by using a double tuned circuit. In double-tuned voltage amplifiers, we use two tuned circuits which are inductively coupled to each other.

8.20. Double Tuned Voltage Amplifier

Figure 8.21 shows the circuit of a double tuned voltage amplifier. It consists of a transistor amplifier with two tuned circuits. One of the tuned circuits (*i.e.* L_1C_1) has been shown as the collector load and other (*i.e.* L_2C_2) as the output. The resistor R_1, R_2 and R_E are used to provide d.c. currents and voltages for the transistor operation.

The signal to be amplified is applied at the input terminal through the coupling capacitor (C_C). The resonant frequency of the tuned circuit L_1C_1 is made equal to that of the signal (by adjusting L_1 or C_1). Under these conditions, the tuned circuit offers a very high impedance to the input signal. As a result of this, a large output appears across the tuned circuit L_1C_1. The output from this tuned circuit is inductively coupled to the L_2 C_2 tuned circuit.

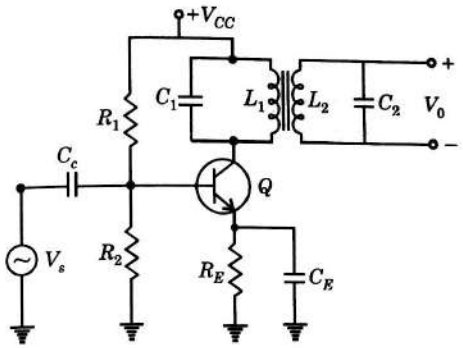
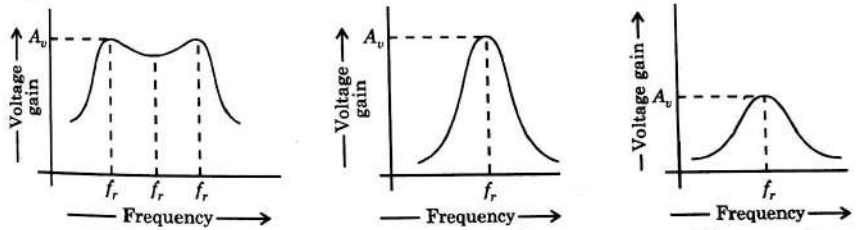


Fig. 8.21. *Double tuned voltage amplifier.*

The double tuned from voltage amplifiers are extensively used in intermediate frequency (IF) amplifiers in radio and television receivers.

8.21. The Frequency Response of Double Tuned Voltage Amplifier

The frequency response of a double tuned voltage amplifier depends upon its degree of coupling (also known coefficient of coupling) between the two tuned circuits. Infact, the degree of coupling gives us an idea of the amount of energy transferred between the two tuned circuits. Figure 8.22 shows the frequency response of tuned circuit with different degrees of coupling. Figure 8.22(*a*) shows the response curve for tight coupling. Here the resonance is shown to occur at two new frequencies (f'_r and f''_r) which are different from the resonance frequency (f_r). As the degree of coupling is decreased, the two frequencies (f'_r and f''_r) come closer and merge into one frequencies (f_r) at critical coupling. The frequency response curve for such type of coupling is shown in figure 8.22(*b*). When the degree of coupling is coupled is decreased below the critical coupling, a single a single peak of reduced height is obtained as shown in figure 8.22(*c*).



(*a*) Tight coupling (*b*) Critical coupling (*c*) Loose coupling

Fig. 8.22.

The tuned circuits of a double-tuned voltage amplifier are tightly coupled to each other. Therefore, the frequency response of a double-tuned voltage amplifier will be as shown in figure 8.23. This type of response curve provides a high selectivity, high gain and relatively large bandwidth to the tuned voltage amplifier.

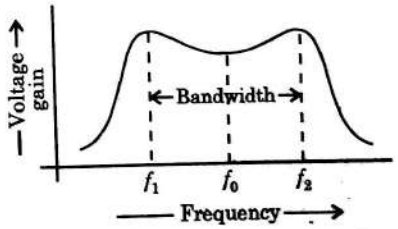


Fig. 8.23. *Frequency response of a double-tuned voltage amplifier.*

8.22. Stagger Tuned Voltage Amplifier

It has been observed that if two of more tuned circuits, which are synchronously tuned are cascaded, the overall bandwidth

decreases. However, if the different tuned circuits which are cascaded, are tuned to slightly different frequencies, it is possible to obtain an increased T bandwidth with a flat passband with steep sides. This technique is known as **stagger-tuning**.

Figure 8.24 shows a two-stage tuned voltage amplifier. The stagger tuning in this circuit, can be achieved by resonating the tuned circuits L_1C_1 and L_2C_2 to slightly different frequencies.

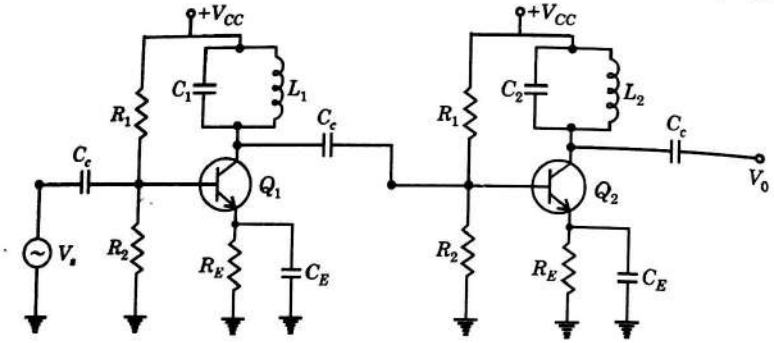


Fig. 8.24. *A two-stage stagger tuned voltage amplifier.*

Figure 8.25 shows the frequency response of a stagger-tuned amplifier. In this figure, curve 'a' shows the gain versus frequency response of the L_1C_1 tuned circuit. Similarly, the curve 'b' shows the response of the L_2C_2 tuned circuit. The curve 'c' indicates the combined (or the overall) response of the circuit. It is obvious from this curve that the amplifier hs a greater bandwidth and flatter passband. It has been found that more the number of stages used, the flatter will be the passband steeper will be the gain fall-off outside the passband. It can be noted that because of stagger-tuning, there is a loss of voltage gain (the gain reduces from A_V to A'_V). If an optimum stagger-tuning is employed, the response curve of the amplifier is very close to a rectangular response curve. Such a response curve is called **Butterworth response.**

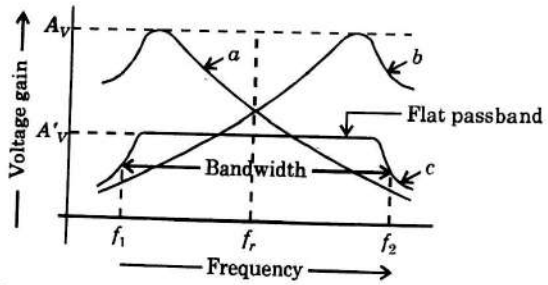


Fig. 8.25. *Frequency response curve of a stagger-tuned voltage amplifier.*

8.23. Local Oscillator

The most common types of local oscillator used in medium wave and short wave AM broadcasting receivers are the Armstrong and the Hartley, the Colpitts and the clapp oscillators. The Colpitts and Clapp oscillators are utilized for higher frequency receiver systems. All local oscillators are L-C oscillators and use a single tuned circuit to determine the frequency of oscillation. Figure 8.26 shows a typical local oscillator circuit used in a superheterodyne receiver along with the mixer circuit.

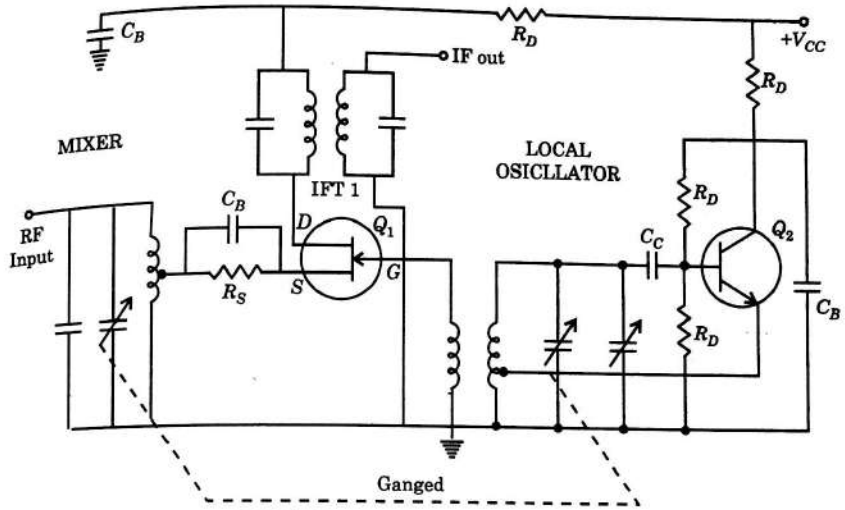


Fig. 8.26. *A local oscillator circuit.*

Local oscillator frequency of the standard broadcast receiver is usually made higher than the incoming signal frequency by an amount equal to intermediate frequency (IF) of 455 kHz. To clarify this let us consider the standard AM broadcast band which extends from 540 kHz to 1650 kHz. Let the local oscillator frequency is made smaller than the incoming signal frequency. In such case, the local oscillator should be capable of varying frequency in the range of 85 kHz to 1195 kHz. This gives a ratio of maximum to minimum frequency equal to 14:1. However, this ratio cannot be achieved by normal tunable capacitance. The capacitance of normal tunable capacitors may be varied in the ratio of maximum capacitance to minimum capacitance equal to 10:1 which gives a maximum frequency ratio of 3.2:1. This means that if the local oscillator frequency is kept lower than the signal frequency by an amount of 455 kHz, the normal tunable capacitors cannot be used.

On the other hand, if the local oscillator frequency is made larger than the incoming signal frequency by an amount of 455 kHz then the variation of the capacitance over the complete AM band would be from 995 kHz to 2105 kHz which gives a ratio of 2.2:1. This ratio can be easily achieved by using normal tunable capacitor. Thus, due to this reason, the superheterodyne receivers used for commercial broadcasting use local oscillator frequency which is higher than the incoming signal frequency which is higher than the incoming signal frequency by an amount equal to the intermediate frequency.

8.23.1. Frequency Mixers

A frequency mixer is a nonlinear device which produces a number of frequencies when two different frequencies are applied at the input of a mixer. In the frequency mixer shown schematically in figure 8.27, an incoming RF signal voltage of frequency f_s and a local oscillator voltage of frequency f_o are applied at its input. The frequency mixer uses a device which has non-linear dynamic characteristics. Due to this, the two input voltages beat together or heterodyne within the mixer to produce an output current which has components of frequency f_s, f_o, $mf_o \pm nf_s$, where m and n are integers. Out of all these terms, one of interest is the component of deference frequency $(f_o - f_s)$. This component is selected in the outside of mixer by a tuned circuit which is tuned to this difference frequency term. This difference frequency term $(f_o - f_s)$ is called **intermediate frequency** (IF).

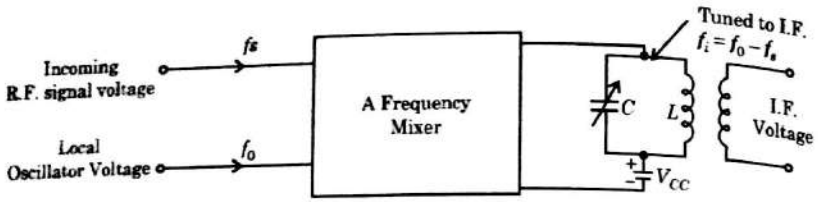


Fig. 8.27. *Schematic diagram of a frequency mixer.*

Thus every incoming RF signal voltage irrespective of its frequency is reduced to this standard intermediate frequency (455 kHz for AM). This means that it is essential that the local oscillator frequency f_o must be made to vary in such a manner so as to maintain the difference frequency $(f_o - f_s)$ which is always equal to the intermediate frequency f_i (455 kHz for AM).

8.23.2. Intermediate Frequencies and IF Amplifiers

In this section, we shall discuss intermediate frequencies used in different communication systems and intermediate frequency amplifier.

8.23.2.1. Choice of Intermediate Frequency

In any communication system, the intermediate frequency (IF) of a receiver is generally a compromise because there are certain reasons why it must be neither low nor high. Following are major factors which affect the choice of intermediate frequency in any communication system:

(*i*) If the intermediate frequency (IF) is quite high then a poor selectivity and poor adjacent-channel rejection will result unless sharp cut-off filters are used in the IF stage of the receiver.

(*ii*) A very high value of intermediate frequency (IF) causes tracking problems in receiver.

(*iii*) If the intermediate frequency is lowered, the image-frequency rejection will become poorer. The rejection of an image frequency is given by

$$\alpha = \sqrt{1+Q^2\rho^2} \qquad ...(8.18)$$

where, Q = Quality factor

$$p = \frac{f_{si}}{f_s} - \frac{f_s}{f_{si}} \qquad ...(8.19)$$

and

$$f_{si} = f_s + 2f_i \qquad ...(8.20)$$

From equations (8.18) and (8.19) it is clear that image frequency rejection is improved if the ratio of image frequency to signal frequency is increased and this requires a high intermediate frequency (IF). It may also be observed that image frequency rejection will become worse if the signal frequency is increased.

(*iv*) If the intermediate frequency is very low, it will make the selectivity too sharp and thus cutting off the sidebands.

(*v*) If the intermediate frequency (IF) is very low, then the frequency stability of the local oscillator will have to be made correspondingly higher since now any frequency drift is a larger proportion of the low IF than that of a high IF.

(*vi*) The intermediate frequency must not fall within the tuning range of the receiver otherwise instability would occur and a heterodyne whistles would be heard and thus would make it impossible to tune to the frequency band just adjacent to the intermediate frequency (IF).

8.23.2.2. Intermediate Frequencies Used

Following are the different intermediate frequencies which are used in different communication systems:

(*i*) In case of standard broadcast AM receivers, which are tuned to 540 kHz to 1650 kHz, the intermediate frequency is 455 kHz.

(*ii*) AM, SSB and other receivers used for shortwave or VHF reception use a first IF in the range of 1.6 to 2.3 MHz, or else above 30 MHz. Actually, such type of receivers use two or more different intermediate frequencies.

(*iii*) FM broadcast receivers tuned to a standard band of 88 MHz to 108 MHz use an IF which is almost 10.7 MHz.

(*iv*) Television receivers in the VHF band (54 MHz to 223 MHz) and in the UHF band (470 MHz to 940 MHz) use an IF between 26 MHz and 46 MHz with 36 MHz and 46 MHz are the two most popular values.

(*v*) Microwave and radar receivers using a frequency range of 1 GHz to 10 GHz use IFs which depend upon the application with 30 MHz, 60 MHz and 70 MHz are the most popular values.

8.23.2.3. Intermediate Frequency (IF) Amplifier

This stage of a superheterodyne receiver generally contains two or more stages of small signal tuned amplifier using CE configuration. The output of the frequency mixer produced across the tuned circuit is coupled to the input of this amplifier. Each such tuned circuit has a passband characteristic and the overall response of the n-stages of IF amplifier provides the required 10 kHz bandwidth. Infact, the number of tuned circuits inductively coupled to the next stage is $(n + 1)$, here n is the number of stages of IF amplifier. This type of tuned circuit tuned to the intermediate frequency and inductively coupled to the next stage is generally known as IF transformer (*i.e.* IFT). In majority of the commercial broadcast receivers two stages of IF amplifier are used. In these stages, the quality factor Q of the tuned circuit is kept high because the band of frequencies accommodated is quite small being only 10 kHz in amplitude modulation broadcast receiver.

Figure 8.28 shows the circuit diagram of a two-stage IF amplifier. The two stages are quite similar. The IF voltage from the first IF transformer is fed to the base of the transistor T_1. The AVC bias is applied at the lower end of the first IF transformer secondary. For the first IF amplifier stage, the collector terminal is grounded through resistor R_4 whereas the emitter terminal is provided a positive bias from the 6 volt d.c. supply through a potential divider constituted by the resistors R_1 and R_2. Also, since the base terminal is grounded through the AVC bias, it results in the required negative bias at the base with respect to the emitter terminal. $R_3 - C_3$ combination makes the AVC decoupling network. For the given values of resistors R_1 and R_2, the emitter to base bias is at +3 volts. The base is connected through the IF transformer secondary to a part of the positive d.c. supply through the potential divider containing resistors R_5 and R_6. Hence, there results the required negative collector-to-base bias V_{CCB} (*i.e.* reverse bias) and positive emitter-to-base bias V_{EEB} (*i.e.* forward bias). The output of the second IF amplifier stage is fed to the demodulator. Also the R_7 - C_7 combination provides the emitter self-bias. The R_4 - C_4 combination provides the collector self-bias in the first stage. In addition to this, impedance matching of the high collector circuit impedance of one stage to the low input impedance of the following stage is obtained by using a step-down transformer.

Neutralization

Every stage of intermediate-frequency (IF) amplifier provides a high gain. Due to this, there results a tendency of oscillation in common emitter amplifier because of positive feedback from the output to the input through collector junction capacitance. Thus, the oscillations so generated cause hawling in the audio output. So hence, the operation of the IF amplifier becomes unstable. To balance this, we use negative feedback in IF amplifier stages. This negative feedback is provided by a capacitor C_2 or C_6 connected between the base and the one end of the primary tuned circuit, the collector being connected to the other.

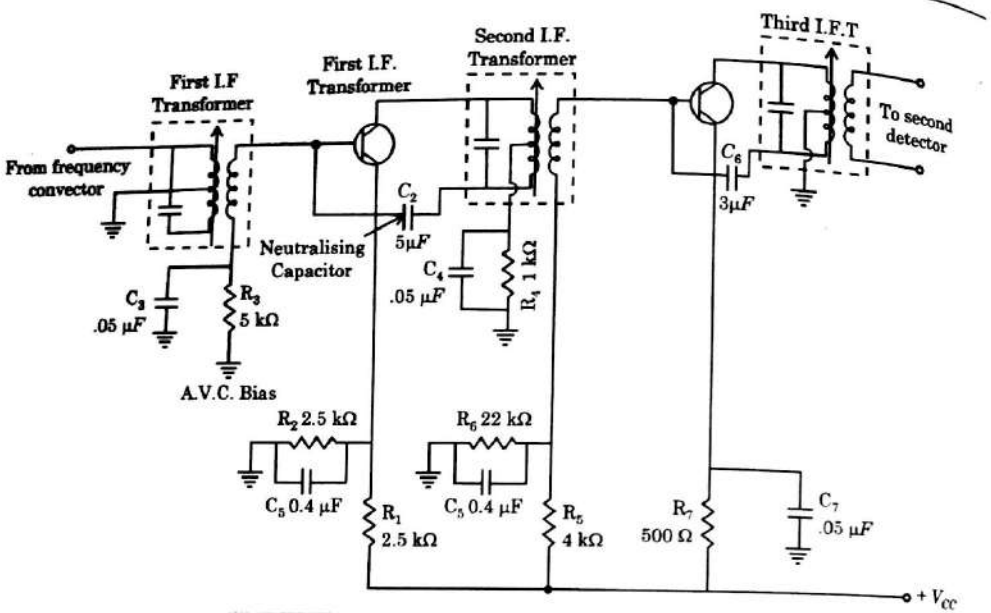


Fig. 8.28. *Two-stage transistorized IF amplifier.*

8.23.3. Automatic Gain Control (AGC) or Automatic Volume Control (AVC)

It is generally observed that as a result of fading, the amplitude of the IF carrier signal at the detector input may vary as much as 30 or 40 dB. This results in corresponding variations in general level of reproduced programme at the receiver output. At a carrier minimum, the loudspeaker output becomes inaudible and gets mixed in noise. On the other hand, at the carrier maximum, the output of the loudspeaker becomes intolerably large. The simplest and the most effective means adopted to balance the fading is the automatic gain control (AGC) or automatic volume control (AVC). Infact, this is universally used in all broadcast and radio receivers. It has been observed that a properly designed AGC system reduces the amplitude variation due to fading from a high value of 30 to 40 dB to a small value of 3 to 4 dB.

Principle of Automatic Gain Control (AGC)

The principle of operation of AGC contains the following steps:

(*i*) to derive by rectification of carrier voltage in a linear diode detector, a d.c. voltage proportional to the carrier amplitude.

(*ii*) to apply this d.c. voltage as a reverse-biased voltage at the input of the RF amplifier, frequency mixer and the IF amplifier.

Thus, now if the carrier signal amplitude increases, the AGC bias increases and the gains of all the tuned stages preceding the detector decrease resulting in decrease in carrier amplitude at the input of the detector bringing it back to its original or normal value. Now, if the carrier amplitude decreases due to some reason, then the reverse action takes place. Hence, the AGC smoothens out the variations in the carrier amplitude to a very large extent.

8.23.4. Linear Diode Detector with Capacitor Filter and Simple AGC

Figure 8.29 shows the circuit of a linear diode detector with simple AGC. In this circuit, the half-wave rectified voltage is developed across load resistor R. Capacitor C filters the RF components due to which only the d.c. and the modulating frequency voltage are obtained across the load resistor R.

The d.c. component is removed from the output by the use of coupling capacitor C_C. AGC is picked up from the diode end of the load resistor R. But, since this voltage consists of modulating frequency component as well, therefore an AGC filter consisting of a series resistor R_A and shunt capacitor C_A is used to remove the modulating or baseband frequency component and thus leaving only a positive d.c. voltage as the required AGC bias.

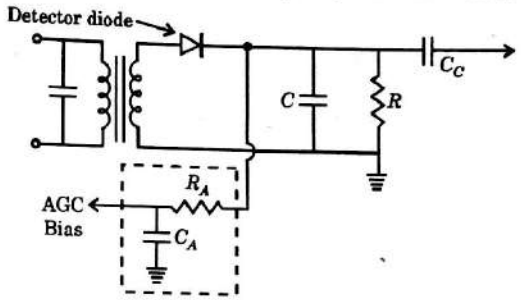


Fig. 8.29. *Linear diode detector with capacitor filter and simple AGC.*

The time-constant of this AGC filter is suitable selected to remove all modulating frequency components. This time-constant R_AC_A must be large enough to remove even the lowest modulating frequency component from the AGC bias. However, AGC must be on the other hand, small enough to enable the AGC bias to follow the change in carrier amplitude. A typical time-constant of AGC filter is in the range of 0.1 to 0.2 second. Now, this positive AGC bias is applied at the base of PNP transistors of preceding tuned stages. This positive AGC bias then reduces the net forward bias at emitter junction thereby reducing the gain of the amplifier. However, in case of NPN transistors, a negative AGC bias is applied at the bases of the transistors of preceding tuned stages. In this case, the detector circuit is similar to that shown in figure 8.29 except that the polarity of the diode is reversed.

8.24. Linear Diode Detector with π-Filter and Simple AGC

To provide the better removal of RF components from the modulating frequency output, a π-filter is used in place of a simple capacitor filter. Figure 8.30 shows the circuit diagram of linear diode detector with π-filter and simple AGC. Also, a manual volume control is generally provided by using a variable resistor at the input of the first audio amplifier as shown in the figure.

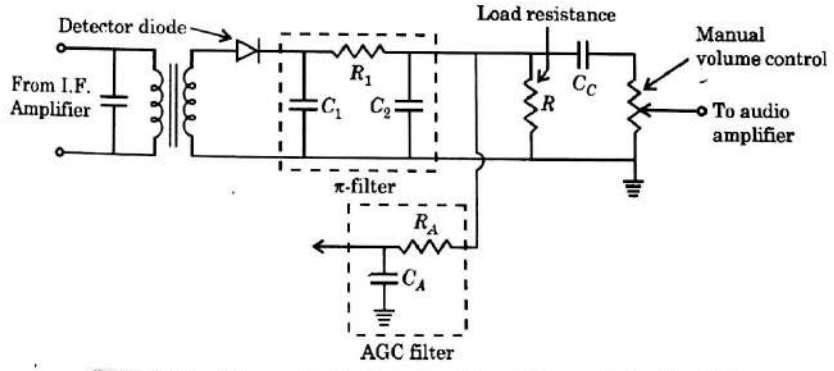


Fig. 8.30. *Linear diode detector with π-filter and simple AGC.*

8.25. Delayed AGC

Simple AGC systems discussed till now (figures 8.29 and 8.30) suffer from a major drawback that the AGC becomes operative even for very week signals. The result of this is that the receiver gain starts falling as soon as detector diode starts producing the output. On the other hand, an ideal AGC system must remain in operative until the input carrier voltage reaches a reasonable large predetermined voltage. Subsequently, the AGC must come into operation to maintain output level

constant instead of variation in input level of carrier voltage. If a delay is produced in AGC operation, then it serves some purpose. Figure 8.31 shows an arrangement to produce delay in AGC operation. With zero and small signal voltages, diode D_2 conducts due to which the AGC bias just equals the potential of cathode of this diode. Hence AGC remains fixed at a low positive value. As the input carrier voltage increases, the AGC bias produced due to rectification of carrier voltage in detector diode D_1 increases. Also, when this rectified bias magnitude exceeds the magnitude of the positive cathode voltage of diode D_2, then diode D_2 stops to conduct and the AGC system works normally.

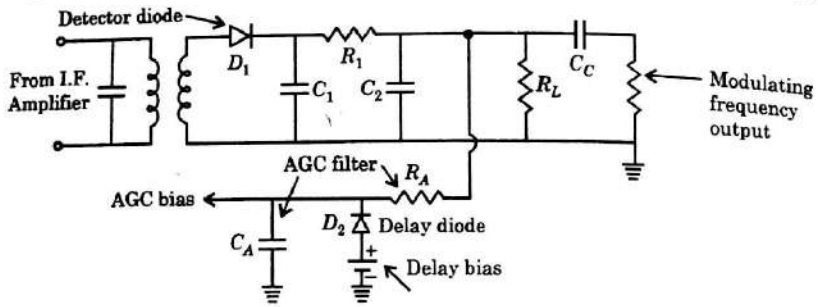


Fig. 8.31. *Linear diode detector with delayed AGC.*

8.26. Amplified and Delayed AGC

Going one step further, if the delayed AGC bias is amplified before application as reverse bias to the tuned amplifiers, the AGC behaviour or characteristics closely approaches the ideal delayed AGC.

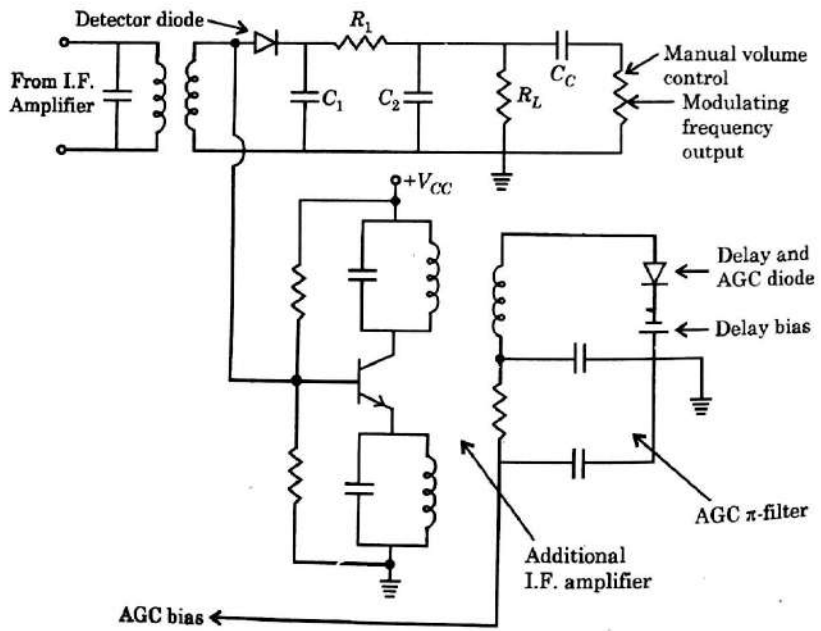


Fig. 8.32. *Linear diode detector with amplified and delayed AGC.*

8.27. AGC Characteristics

The AGC characteristic *i.e.* the curve giving audio output of receiver plotted against input carrier voltage is shown in figure 8.33 for different cases.

Simple AGC is given by the curve *S* whereas an ideal AGC is represented by the curve ID. The curve PD gives the AGC characteristics of delayed AGC and the curve AD gives the amplified and delayed AGC characteristics.

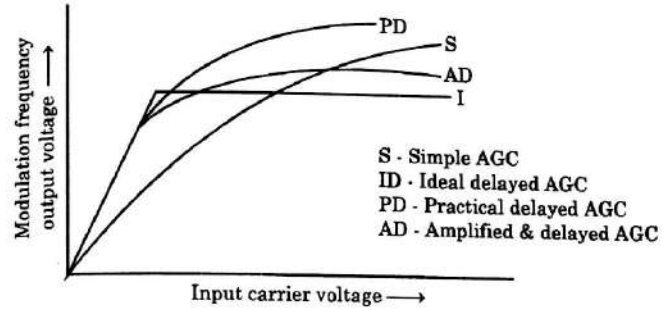


Fig. 8.33. *AGC Characteristics.*

8.28. LIC Superheterodyne AM Receivers

A complete superheterodyne AM receiver is now available in the form of linear IC. Because the inductor cannot be fabricated in the IC form, therefore, we have to make tuning circuit arrangement externally for coupling the in-built RF amplifier, mixer and IF amplifier stages of the Linear ICs. A number of linear IC superheterodyne AM receivers are available in the market. The National Semiconductor Corporation offers LM 1820, a 14 pin IC which contains the RF amplifier, mixer

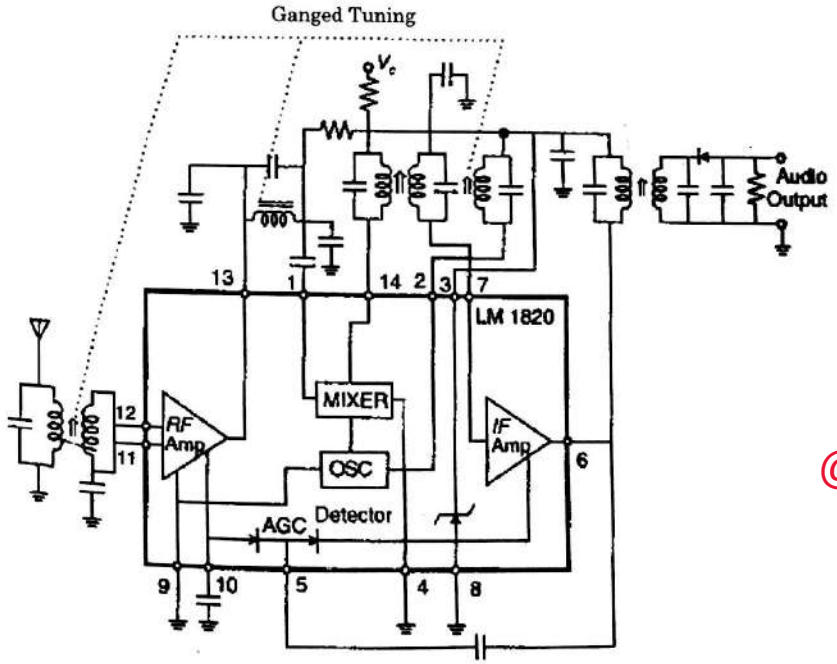


Fig. 8.34. *Linear IC (LM 1820) superheterodyne AM receiver.*

@sTc

and local oscillator and IF amplifier. With proper external connections, this IC can be used as a super heterodyne AM receiver. Figure 8.34 shows a typical arrangement of IC LM 1820 superheterodyne AM receiver. These linear IC receivers are not so popular they are neither low cost nor small sized. The size and cost are found out mainly by the frequency tuning circuit. The linear IC AM radios find applications in the high quality AM receivers. However, ceramic filters may be used in place of LC tuned circuit.

These types of filters are IC compatible. Use of these ceramic filters is theoretically possible to fabricate a functional AM broadcast band receiver using just the chip and two external potentiometers (for volume control and station selection) and antenna.

8.29. AM Transmitter

In this article, let us describe a typical AM transmitter briefly. This transmitter can be either high level, or low level. The *RF* oscillator invaribly makes use of a crystal controlled oscillator to ensure high accuracy of the carrier frequency as recommended by the FCC. It is followed by a Class-A RF buffer amplifier which provides a high impedance load for the oscillator to minimize drift. Then the carrier signal is subsequently amplified using a class-C RF amplifier. For high level modulation, the final stage is another class-C RF output amplifier at which the modulating signal $x(t)$ is amplified. The modulating signal is processed before it is applied at the final stage. The modulating signal is filtered so as to occupy the correct bandwidth of 10 kHz after modulation. The modulating signal is then amplified by audio amplifier and power audio frequency (AF) amplifier. The modulating signal culminates in the modulator amplifier, which is the highest power audio amplifier. The difference between high-level and low-level modulation depends upon the point at which modulation takes place. For low-level modulator, the modulation is done at some stage before the final stage. In the block diagram, an class-B RF linear amplifier before transmitting antenna has been shown in the figure 8.35.

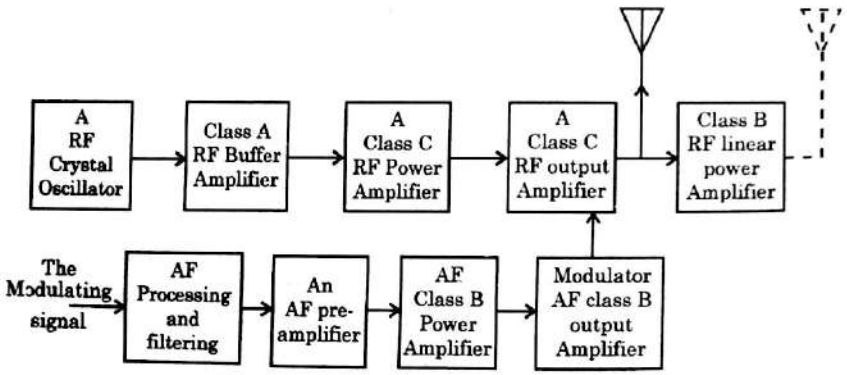


Fig. 8.35. *Block diagram of a typical AM transmitter.*

With the development of linear integrated circuit (LIC) technology, it has become possible to achieve LIC transmitters also. Instead of choosing discrete design of transmitter systems, one can directly choose LIC transmitters which requires minimum circuit connection. LP 2000 is a nomolithic transmitter offered by Lithic systems Inc. The block diagram of the internal arrangement of the internal arrangement of this IC has been shown in figure 8.36. Schematic circuit incorporating external discrete components to use the chip as on AM transmitter has been shown in figure 8.37.

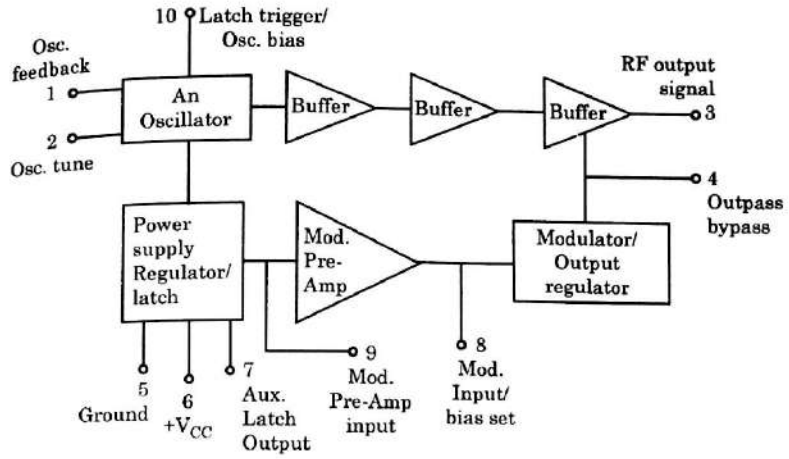


Fig. 8.36. *Block diagram of the internal arrangement of a LP 2000 micro transmitter.*

The outputs of such transmitters are generally low of the order of a 50 mW and can be used for short range applications of the order of 1 km or 50.

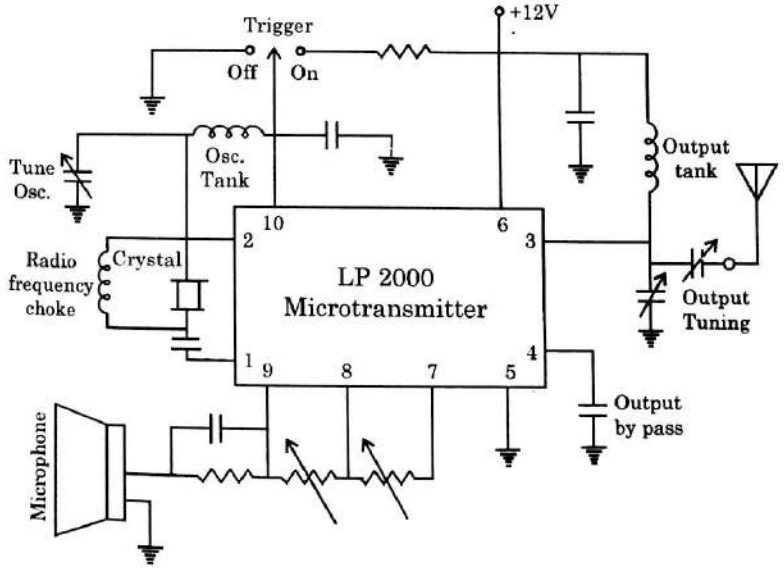


Fig. 8.37. *Schematic circuit arrangement for using LP 2000 as AM transmitter.*

Example 8.2. A tuned radio frequency receiver (TRF) has to be designed with a single tuned circuit using a 9×10^{-6} Henry inductor. The ideal 10 kHz bandwidth occurs at 1100 kHz.

(i) Determine the capacitance range of the variable capacitor in the LC tank circuit required to tune the receiver in 550 to 1550 kHz range.

(ii) Determine the bandwidth (BW) of the receiver at 550 and 1550 kHz.

(Karnataka University-1998)

Solution: We know that the Q of the L-C tank circuit is given by

$$Q = \frac{f_c}{BW}$$

Substituting, the given values, we get

$$Q = \frac{1100}{10} = 110$$

(*i*) The resonant frequency of the L-C tank circuit is expressed as

$$f_c = \frac{1}{2\pi\sqrt{LC}} \quad ...(i)$$

At 550 kHz resonant frequency, the required capacitance value can be ontained as under:

$$550 \times 10^3 = \frac{1}{2\pi\sqrt{9 \times 10^{-6} \times C}}$$

Solving, we get

$$C = 9.31 \times 10^{-9} \text{ Farad} = 9.31 \text{ nano Farad}$$

At 1550 kHz resonant frequency, the required capacitance value can be obtained as under:

$$1550 \times 10^3 = \frac{1}{2\pi\sqrt{9 \times 10^{-6} \times C}}$$

Solving, we get

$$C = 1.17 \times 10^{-9} \text{ Farad} = 1.17 \text{ nano Farad}$$

(*ii*) The bandwidth of the TRF receiver when tuned to 550 kHz will be given by

$$BW = \frac{f_c}{Q} = \frac{550}{110} \text{ kHz} = 5 \text{ kHz}$$

The bandwidth of the TRF receiver when tuned to 1550 kHz will be given by

$$BW = \frac{f_c}{Q} = \frac{1550}{110} \text{ kHz} \cong 14 \text{ kHz} \quad \textbf{Ans.}$$

Example 8.3. A superheterodyne receiver is tuned to receive a 1 MHz carrier amplitude modulated by 1 kHz sine wave. It is also assumed that the IF frequency is 455 kHz. Find the frequency components at the input and output of the IF amplifier. Assume the IF bandwidth to be 10 kHz.

Solution: Because the superheterodyne receiver makes use of 455 kHz IF frequency, therefore the local oscillator (LO) frequency will be

$$f_{LO} = f_c + 455 \quad ...(i)$$

When the receiver is tuned to 1000 kHz, the local oscillator frequency becomes

$$f_{LO} = f_c + 455$$

or

$$f_{LO} = 1000 + 455 = 1455 \text{ kHz}$$

Now, the mixer is a non-linear device, it would generate the following frequencies at the output which will be presented to the input of the IF amplifier:

(*a*) Frequencies at the original inputs 1000 kHz, upper side frequency is 100 kHz, lower side frequency is 999 kHz and local oscillator frequency is 1455 kHz.

(*b*) Sum and difference components of the above input frequencies, *i.e.* 2454 kHz, 2455 kHz, 2456 kHz, 454 kHz, 455 kHz and 456 kHz.

(*c*) Harmonics of all frequencies given above in (*a*) and (*b*) plus d.c. component, it would accept only the components 454 kHz and 456 kHz.

Example 8.4. Find the image frequency for a standard broadcast band AM receiver using a 455 kHz IF and tuned to a station at 640 kHz.

Solution: We know that the local oscillator frequency is given by

$$f_{LO} = f_s + f_i$$

Substituting the given values, we get

$$f_{LO} = 640 + 455 = 1095 \text{ kHz}$$

The image frequency f_{si} can be expressed as

$$f_{si} = f_{LO} + f_i$$

or

$$f_{si} = 1095 + 455 = 1550 \text{ kHz}$$

The image frequency f_{si} can also be found directly using the following relatinship

$$f_{si} = f_s + 2f_i$$

The image frequency f_{si} is thus far away from the desired station frequency (f_s). Therefore, image frequency is not a major problem for standard AM broadcast band.

Example 8.5. In a broadcast superheterodyne receiver having no RF amplifier, the loaded Q of the antenna coupling circuit is 100. If the IF frequency is 455 kHz, determine

(*i*) The image frequency and its rejection ratio for tuning at 1.1 kHz station.

(*ii*) The image frequency and its rejection ratio for tuning at 25 MHz.

Solution: (*i*) We know that

$$f_{si} = f_s + 2f_i$$

Putting the given values, we get

$$f_{si} = 1100 + 2 \times 455 = 2010 \text{ kHz}$$

Also, since

$$\rho = \frac{f_{si}}{f_s} - \frac{f_s}{f_{si}}$$

therefore,

$$\rho = \frac{2010}{1100} - \frac{1100}{2010} = 1.28$$

The image rejection is given by the expression

$$\alpha = \sqrt{1 + Q^2\rho^2} = \sqrt{1 + (100)^2(1.28)^2} = 128$$

(*ii*) The image frequency can be obtained from equation $f_{si} = f_s + 2f_i$ as under :

$$f_{si} = 25 + 2 \times 0.455 = 25.91 \text{ MHz}$$

Using the expression,

$$\rho = \frac{f_{si}}{f_s} - \frac{f_s}{f_{si}}$$

we get

$$\rho = \frac{25.91}{25} - \frac{25}{25.91} = 0.0715$$

Again, the image rejection can be found from the following expression

$$\alpha = \sqrt{1 + Q^2\rho^2} = \sqrt{1 + (100)^2(0.0715)^2} = 5.22$$

Thus rejection is very poor for practical receivers in the HF band.

Image rejection is not a major problem for a standard AM broadcast receiver without an RF stage. Image rejection causes problem for reception of high frequency signal using the same IF. We can get rid of this shortcoming by using extra RF stage or by increasing the intermediate frequency.

Example 8.6. For a superheterodyne receiver having no RF amplifier, the loaded Q of the antenna coupling circuit is 100. If the intermediate frequency is 455 kHz. The superheterodyne receiver is to be improved for HF reception so that its image rejection at 25 MHz is as good as it was at 1100 kHz. Obtain

(i) The loaded Q which an RF amplifier for this receiver would have to have

(ii) The new intermediate frequency that will be needed in the absence of RF amplifier.

(Nagpur University-1999)

Solution: (*i*) Let the image frequency rejection of the RF amplifier to be added be α'. The image rejection ratio α at 1100 kHz can be found as under:

Since $\quad f_{si} = f_s + 2f_i = 1100 + 2 \times 455 = 2010$ kHz

Therefore, $\quad \rho = \dfrac{f_{si}}{f_s} - \dfrac{f_s}{f_{si}} = \dfrac{2010}{1100} - \dfrac{1100}{2010} = 1.28$

Thus, the image rejection ratio is given by

$$\alpha = \sqrt{1+\rho^2 Q^2} = \sqrt{1+(1.28)^2(100)^2} = 128$$

Here, $\alpha = 128$ is the image rejection ratio at frequency 1100 kHz.

The image rejection ratio α at frequency 25 MHz is found as under:

$$f_{si} = f_s + 2f_i = 25000 + 2 \times 455 = 25910 \text{ kHz}$$

and $\quad \rho = \dfrac{f_{si}}{f_s} - \dfrac{f_s}{f_{si}} = \dfrac{25910}{25000} - \dfrac{25000}{25910} = 0.0715$

Thus, $\quad \alpha = \sqrt{1+\rho^2 Q^2} = \sqrt{1+(0.0715)^2(100)^2} = 5.22$

Hence, rejection is very poor for practical receivers in the HF band.

According to our requirement, we have

$$\alpha\alpha' = 128$$

or $\quad \alpha = 5.22$ for tuning at 25 MHz

or $\quad \alpha' = \dfrac{128}{\alpha} = \dfrac{128}{7.22} = 15.7$

Further, using expression $\alpha = \sqrt{1+Q'^2\rho^2}$, we get the loaded Q' required for the RF stage as under:

$$a' = \sqrt{1+Q'^2\rho^2}$$

$$15.5. = \sqrt{1+Q'^2(0.0715)^2}$$

or $\quad Q' \cong 247$

(*ii*) The image rejection is to be made 128 for tuning at 25 MHz by changing IF frequency only. This means that ρ has to remain same as in part (*i*)

Now, if f_i' is the new *IF* frequency then the new image frequency f'_{si} becomes

$$f'_{si} = f_s' + 2f_i'$$

Since ρ remains constant, we can write

$$\frac{f'_{si}}{f'_s} - \frac{f'_s}{f'_{si}} = 1.28$$

or $\quad \dfrac{f'_{si}}{f'_s} = 1.827$

Substituting $f'_s = 25$ MHz, we get

$$\frac{25 + 2f'_i}{25} = 1.827 \quad \text{or} \quad f'_i = 10.3 \text{ MHz}$$

Hence, the new *IF* frequency must be chosen as 10.3 MHz to achieve the desired image rejection (in the absence of RF stage). The use of RF amplifiers in the superheterodyne receiver is not compulsory. Receivers having RF amplifier stages are definitely superior in performance to the receiver without one. But all the other things are identical. In some cases the RF amplifier may be uneconomical in the sense that it increases only the price of the receiver without significantly improving the receiver's performance. **Ans.**

Example 8.7. A radio receiver used in the AM system is shown in figure 8.38. The mixer translates the frequency f_c to a fixed intermediate frequency (IF). If 455 kHz by using a local oscillator of frequency f_{LO}. The broadcase-band frequencies range from 540 to 1600 kHz.

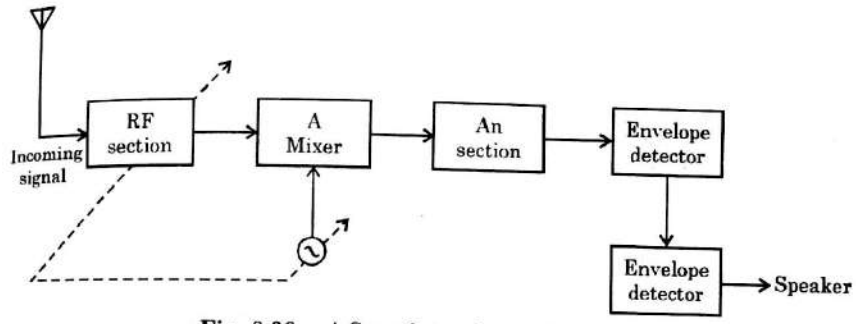


Fig. 8.38. *A Superheterodyne AM receiver.*

(a) Determine the range of tuning that must be provided in the local oscillator (i) when f_{LO} is higher than f_c (superheterodyne receiver) and (ii) when f_{LO} is lower than f_c.

(b) Explain why the usual AM radio receiver uses a superheterodyne system.

Solution: (*a*) (*i*) When $f_{LO} > f_c$,

$$540 < f_c < 1600$$

$$f_{LO} - f_c = 455$$

where both f_c and f_{LO} are expressed in kilohertz. Thus

$$f_{LO} = f_c + 455$$

When $f_c = 540$ kHz, we get $f_{LO} = 995$ kHz; and $f_c = 1600$ kHz, we get $f_{LO} = 2055$ kHz. Thus, the required tuning range of the local oscillator is 995 – 2055 kHz.

(*ii*) When $\quad f_{LO} < f_c$,

$$f_{LO} = f_c - 455$$

When $f_c = 540$ kHz, we get $f_{LO} = 85$ kHz; and when $f_c = 1600$ kHz, we get $f_{LO} = 1145$ kHz. Thus the required tuning range of the local oscillator for this case is 85 – 1145 kHz.

(*b*) The frequency ratio, that is, the ratio of the highest f_{LO} to the lowest f_{LO} is 2.07 for case (*i*) and 13.47 for case (*ii*). It is much easier to design an oscillator that is tunable over a smaller frequency ratio, that is the reason why the usual AM radio receiver uses the superheterodyne system.

SUMMARY

1. A radio transmitter radiates or transmits a modulated carrier signal.
2. The main function of a radio receiver as:
   (*i*) Intercept the incoming modulated signal (*i.e.* electromagnetic waves) by the receiving antenna.
   (*ii*) Select the desired signal and reject the unwanted signals.
   (*iii*) Amplify this selected RF signal.

(iv) Detect the modulated signal to get back the original modulating or baseband signal.
(v) Amplify the modulating frequently signal.

3. Depending upon the applications, the radio receivers may be classified as follows :
   (i) *Amplitude Modulation (AM) Broadcast Receivers:* These receivers are used to receive the broadcast of speech or music transmitted from amplitude modulation broadcast transmitters which operate on long-wave, medium-wave or short-wave bands.
   (ii) *Frequency Modulation (FM) Broadcast Receivers:* These receivers are used to receive the broadcast programmes from F.M. broadcast transmitters which operate in VHF or UHF bands.
   (iii) *Communication Receivers:* Communication receivers are used for reception of telegraph and short-wave telephone signals. This means that communication receivers are used for various purposes other than broadcast services.
   (iv) *Television Receivers:* Television receivers are used to receive television broadcast in VHF or in UHF bands.
   (v) *Radar Receivers:* Radar receivers are used to receive Radar (*i.e.* Radio detection and ranging) signals.
4. Depending upon the fundamental aspects, the radio receivers may also be classified as under :
   (i) Tuned Radio Frequency (TRF) Receivers.
   (ii) Superheterodyne Receiver.
5. Tuned radio frequency (TRF) receiver is the simplest radio receiver. Figure 8.1 shows the block diagram of a tuned radio frequency receiver. The very first block of this receiver is an RF stage. This stage generally contains two or three RF amplifiers.
6. Although TRF receiver is cheaper and the simplest one, it has certain drawbacks as under:
   (i) The TRF receiver suffers from a tendency to oscillate at higher frequencies from the multistage RF amplifiers with high gain and operating at same frequency. If such an amplifier has a gain of 20,000 then if a small portion of the output leaked back to the input of the RF stage, then positive feed back and oscillation will result. This type of leakage could result from power supply coupling, stray capacitance coupling, radiation coupling or coupling through any other element common to the input and output stages. Definitely, this type of condition is undesirable for a good receiver. This problem is also termed as instability of the receiver.
   (ii) The selectivity of a receiver is its ability to distinguish between a desired signal and an undesired signal. The selectivity of TRF receiver is poor. In fact, it is difficult to achieve sufficient selectivity at high frequencies due to the enforced use of single-tuned circuits.
   (iii) Another problem associated with the TRF receiver is the bandwidth variation over the tuning range. For example, in AM broadcast system, let us consider that a tuned circuit is required to have a bandwidth of 10 KHz at a frequency of 540 kHz.
7. In a superheterodyne receiver, the incoming RF signal frequency is combined with the local oscillator signal frequency through a mixer and is converted into a signal of lower fixed frequency. This lower fixed frequency is known as intermediate frequency.
8. The advantages of the superheterodyne receiver make it the most suitable for the majority of radio receiver applications like AM, FM, communications, single-sideband, television and even radar receiver all use superheterodyne principle. This means that it can be considered as today's standard form of radio receiver.
9. The sensitivity of a radio receiver may be defined as its ability to amplify weak signals.
10. The selectivity of a receiver may be defined as the ability to reject unwanted signals. It also expresses the attenuation that the receiver offers to signal at frequencies adjacent to the one to which it is tuned.
11. Selectivity depends upon the following factors:
    (i) Selectivity varies with receiving frequency and becomes some what worse when the receiving frequency is raised.
    (ii) In general, it is mainly determined by the response of the IF section, with the mixer and RF amplifier input circuits playing a small but significant part.
    (iii) Selectivity is the main factor which determines the adjacent channel rejection of a receiver.

12. When a receiver picks up the same short wave station at two nearby points on the receiver dial, the double spotting phenomenon takes place. The main cause for double spotting is poor front-end selectivity, *i.e.*, inadequate image-frequency rejection.
13. A superheterodyne receiver suffers from a major drawback known as Image Frequency problem. This problem of image frequency is inherent to a superheterodyne receiver and arises because of the use of heterodyne principle.
14. The advantages of RF amplifier as under:
    (i) Greater gain *i.e.* better sensitivity
    (ii) Improved rejection of adjacent undesired signals *i.e.* better selectivity
    (iii) Improved signal/noise ratio
    (iv) Improved image frequency rejection
    (v) Improved coupling of the receiver to the antenna
    (vi) Prevention of re-radiation of the local oscillator voltage through the antenna
    (viii) Prevention of spurious frequencies from entering the mixer and heterodyne to produce interfering frequency equal to IF
15. The most common types of local oscillator used in medium wave and short wave AM broadcasting receivers are the Armstrong and the Hartley, the colpitts and the clapp oscillators.
16. A frequency mixer is a nonlinear device which produces a number of frequencies when two different frequencies are applied at the input of a mixer.

SHORT QUESTIONS WITH ANSWERS

Q.1. Write the main functions of a radio receiver.

Ans. The main functions of a radio receiver as under:
(i) intercept the incoming modulated signal (*i.e.* electromagnetic waves) by the receiving antenna.
(ii) select the desired signal and reject the unwanted signals.
(iii) amplify this selected R.F. signal.
(iv) detect the modulated signal to get back the original modulating or baseband signal.
(v) amplify the modulating frequency signal.

This means that a radio receiver is an electronic equipment which picks up the desired signal, rejects the unwanted signals, amplifies the desired signal, demodulates the modulated signal to get back the original modulating frequency signal.

Q.2. How can you classify radio receiver?

Ans. Depending upon the fundamental aspects, the radio receivers may also be classified as under:
(i) Tuned Radio Frequency (TRF) Receivers.
(ii) Superheterodyne Receiver.

Q.3. Write the drawbacks of Tuned Radio Frequency (TRF) Receiver.

Ans. TRF receiver is cheaper and the simplest one, it has certain drawbacks as under:
(i) The TRF receiver suffers from a tendency to oscillate at higher frequencies from the multistage RF amplifiers with high gain and operating at same frequency. If such an amplifier has a gain of 20,000 then if a small portion of the output leaked back to the input of the RF stage, then positive feedback and oscillation will result. This type of leakage could result from power supply coupling, stray capacitance coupling, radiation coupling or coupling through any other element common to the input and output stages.
(ii) The selectivity of a receiver is its ability to distinguish between a desired signal and an undesired signal. The selectivity of TRF receiver is poor. In fact, it is difficult to achieve sufficient selectivity at high frequencies due to the enforced use of single-tuned circuits.

(*iii*) Another problem associated with the TRF receiver is the bandwidth variation over the tuning range. For example, in AM broadcast system, let us consider that a tuned circuit is required to have a bandwidth of 10 kHz at a frequency of 540 kHz.

Q.4. What do you mean by Sensitivity?

Ans. The sensitivity of a radio receiver may be defined as its ability to amplify weak signals. It is generally, defined in terms of the voltage which must be applied at receiver input terminals to provide a standard output power measured at the output terminals. For AM broadcast receivers, several relevant quantities have been standardized. A signal modulated by a 400 Hz sine wave and modulation index of 30% is applied through standard coupling network known as a **dummy antenna**.

In addition to this, the loudspeaker is replaced by an equivalent load resistance. After this the output is measured across this resistance and it must be equal to the standard value of 50 mW.

Sensitivity is also expressed in microvolts or in decibels below 1 volt and is measured at three points along the tuning range when a production receiver is lined up.

Q.5. Define Selectivity for a receiver?

Ans. The selectivity of a receiver may be defined as the ability to reject unwanted signals. It also expresses the attenuation that the receiver offers to signal at frequencies adjacent to the one to which it is tuned. In selectivity measurement, the frequency of the generators is varied to either side of the frequency to which the receiver is tuned. Naturally, the output of the receiver falls since the input frequency is not incorrect. Thus, the input voltage must be increased until the output is the same as it was originally.The ratio of the voltage required of resonance to the voltage required.

Q.6. Write the advantages of a R.F. amplifier.

Ans. The advantages of RF amplifier as under:

(*i*) Greater gain *i.e.* better sensitivity

(*ii*) Improved rejection of adjacent undesired signals *i.e.* better selectivity

(*iii*) Improved signal/noise ratio

(*iv*) Improved image frequency rejection

(*v*) Improved coupling of the receiver to the antenna

(*vi*) Prevention of reradiation of the local oscillator voltage through the antenna

(*vii*) Prevention of spurious frequencies from entering the mixer and heterodyne to produce interfering frequency equal to IF

REVIEW QUESTIONS

1. What are the salient features of broadcast radio receivers?
2. State the principle of a heterodyne receiver.
3. Draw the block diagram of a superheterodyne receiver and explain the function of each block.
4. Draw the circuit of a frequency mixer and local oscillator using transistors and explain its working.
5. What is meant by the term 'tracking error'?
6. Draw a typical tracking error curve.
7. Draw the circuit of an IF amplifier and explain its working.
8. What are the main functions served by an IF amplifier?
9. What factors govern the choice of intermediate frequency?
10. What is meant by the term adjacent channel selectivity?
11. Draw the circuit of a linear diode detector and explain its working.
12. Distinguish between simple AGC and delayed AGC.

CHAPTER 9

Sampling Theory and Pulse Modulation

Inside this Chapter

9.1. Introduction
9.2. The Sampling Theorem 485
9.3. Proof of Sampling Theorem 486
9.4. Nyquist Rate and Nyquist Interval 486
9.5. Reconstruction Filter (Low-Pass Filter) 489
9.6. Signal Reconstruction: The Interpolation Formula 492
9.7. Effect of Under Sampling: Aliasing 493
9.8. Sampling of Bandpass Signals 495
9.9. Sampling Techniques 495
9.10. Aperture Effect 500
9.11. Comparison of Various Sampling Techniques 507
9.12. Analog Pulse Modulation Methods 509
9.13. Pulse Amplitude Modulation (PAM) 511
9.14. Pulse Time Modulation 512
9.15. Comparison of Various Pulse Analog Modulation Methods 519
526

9.1. Introduction

As we know that broadly, there are two types of signals, continuous time signal and discrete-time signals. Due to some recent advance development in digital technology over the past few decades, the inexpensive, light weight, programmable and easily reproducible discrete-time systems are available. Therefore, the processing of discrete-time signals is more flexible and is also preferable to processing of continuous-time signals.

This means that in practice, although we have a large number of continuous-time signals, but we prefer processing of discrete-time signals. For this purpose we should be able to convert a continuous-time signal into discrete-time signal.

This problem is solved by a fundamental mathematical tool known as sampling theorem. The sampling theorem is extremely important and useful in signal processing. With the help of sampling theorem, a continuous-time signal may be completely represented and recovered from the knowledge of samples taken uniformly. This means that sampling theorem provides a mechanism for representing

a continuous-time signal by a discrete-time signal. Therefore, sampling theorem may be viewed as a bridge between continuous-time signals and discrete-time signals.

The concept of sampling provides a widely used method for using discrete-time system technology to implement continuous-time systems and process the continuous-time signals. We utilize sampling to convert a continuous-time signal to a discrete-time signal, process the discrete-time signal using a discrete-time system and then convert back to continuous-time signals.

9.2. The Sampling Theorem

(U.P. Tech-Semester Exam. 2002-2003)

As discussed earlier, sampling of the signals is the fundamental operation in signal-processing. A continuous time signal is first converted to discrete-time signal by sampling process. The sufficient number of samples of the signal must be taken so that the original signal is represented in its samples completely. Also, it should be possible to recover or reconstruct the original signal completely from its samples. The number of samples to be taken depends on maximum signal frequency present in the signal. Sampling theorem gives the complete idea about the sampling of signals. Different types of samples are also taken like ideal samples, natural samples and flat-top samples.

Let us discuss the sampling theorem first and then we shall discuss different types of sampling processes. The statement of sampling theorem can be given in two parts as:

(*i*) A bandlimited signal of finite energy, which has no frequency-component higher than f_m *Hz*, is completely described by its sample values at uniform intervals less than or equal to $\frac{1}{2f_m}$ second apart.

(*ii*) A bandlimited signal of finite energy, which has no frequency components higher than f_m Hz, may be completely recoverd from the knowledge of its samples taken at the rate of 2 f_m samples per second.

The first part represents the representation of the signal in its samples and minimum sampling rate required to represent a continuous-time signal into its samples.

The second part of the theorem represents reconstruction of the original signal from its samples. It gives sampling rate required for satisfactory reconstruction of signal from its samples.

Combining the two parts, the sampling theorem may be stated as under:

"A continuous-time signal may be completely represented in its samples and recovered back if the sampling frequency is $f_s \geq 2 f_m$. Here, f_s is the sampling frequency and f_m is the maximum frequency present in the signal".

9.3. Proof of Sampling Theorem

(U.P. Tech-Semester Examination 2002-2003)

To prove the sampling theorem, we shall show that a signal whose spectrum is bandlimited to f_m *Hz*, can be reconstructed exactly without any error from its samples taken uniformly at a rate $f_s > 2 f_m$ *Hz*.

Let us consider a continuous time signal $x(t)$ whose spectrum is bandlimited to f_mHz. This means that the signal $x(t)$ has no frequency components beyond f_mHz. Therefore, $X(w)$ is zero for $|w| > w_m$, *i.e.*,

$$X(w) = 0 \text{ for } |w| > w_m$$

where $$w_m = 2\pi f_m$$

Figure 9.1 (*a*) shows this continuous-time signal $x(t)$. Let $X(w)$ represents its Fourier transform or frequency spectrum as shown in figure 9.1(*b*).

Sampling of $x(t)$ at a rate of f_s*Hz* (f_s samples per second) may be achieved by multiplying $x(t)$ by an impulse train $\delta_{Ts}(t)$. The impulse train $\delta_{Ts}(t)$ consists of unit impulses repeating periodically every T_s seconds, where $T_s = \frac{1}{f_s}$.

Figure 9.1(*c*) shows this impulse train. This multiplication results in the sampled signal $g(t)$ shown in figure 9.1(*e*).

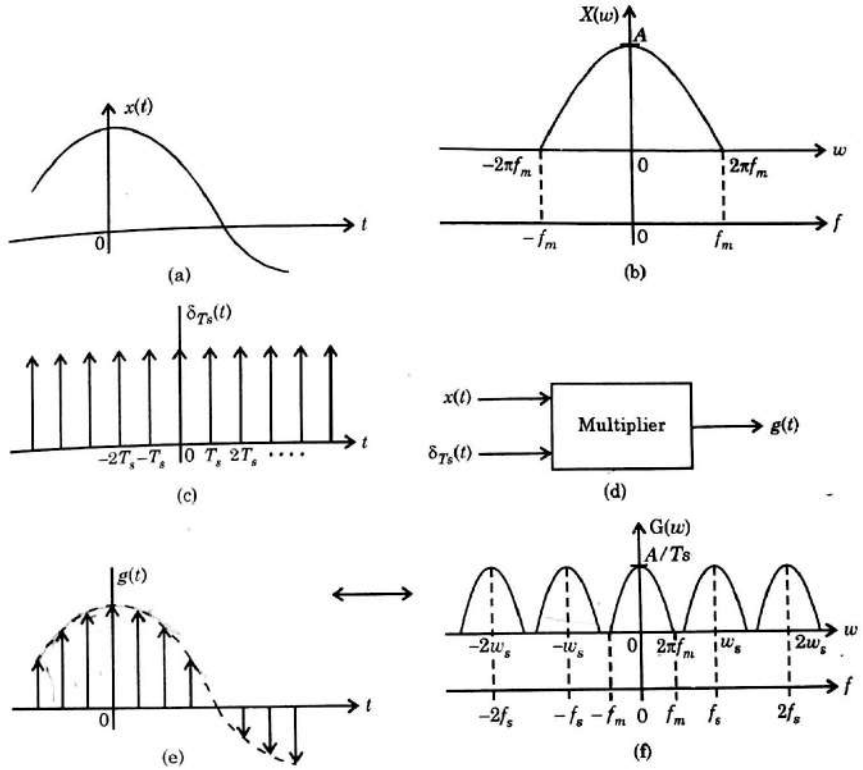


Fig. 9.1. (*a*) *A continuous-time signal.*
(*b*) *Spectrum of continuous-time signal.*
(*c*) *Impulse train as sampling function.*
(*d*) *Multiplier.*
(*e*) *Sampled signal.*
(*f*) *Spectrum of sampled signal.*

This sampled signal consists of impulses spaced every T_s seconds (the sampling interval). The resulting or sampled signal may be written as

$$g(t) = x(t)\, \delta_{Ts}(t) \qquad ...(9.1)$$

Again, since the impulse train $\delta_{Ts}(t)$ is a periodic signal of period T_s, it may be expressed as a Fourier series.

The trigonometric Fourier series expansion of impulse-train $\delta_{Ts}(t)$ is expressed as

$$\delta_{Ts}(t) = \frac{1}{T_s}\,[1 + 2\cos w_s t + 2\cos 2\, w_s t + 2\cos 3\, w_s t + ...] \qquad ...(9.2)$$

Here $$w_s = \frac{2\pi}{T_s} = 2\pi f_s$$

Putting the values of $\delta_{T_s}(t)$ from equation (9.2) in equation (9.1), the sampled signal is

$$g(t) = \frac{1}{T_s}\,[x(t) + 2\,x(t)\cos w_s t + 2x(t)\cos 2\,w_s t + 2x(t)\cos 3\,w_s t + \ldots] \qquad \ldots(9.3)$$

Now, to obtain $G(w)$, the Fourier transformation of $g(t)$, we will have to take the Fourier transform of right hand side.

Fourier transform of $x(t)$ is $X(w)$.

Fourier transform of $2x(t)\cos w_s t$ is $[X(w - w_s) + X(w + w_s)]$.

Fourier transform of $2x(t)\cos 2\,w_s t$ is $[X(w - 2w_s) + X(w + 2w_s)]$ and so on.

Therefore, on taking Fourier transformation, the equation (9.3) becomes

$$G(w) = \frac{1}{T_s}\,[X(w) + X(w - w_s) + X(w + w_s) + X(w - 2\,w_s) + X(w + 2\,w_s) + X(w - 3\,w_s) + X(w + 3\,w_s) + \ldots] \qquad \ldots(9.4)$$

or

$$G(w) = \frac{1}{T_s}\sum_{n=-\infty}^{\infty} X(w - n\,w_s) \qquad \ldots(9.5)$$

From equations (9.4) and (9.5), it is clear that the spectrum $G(w)$ consists of $X(w)$ repeating periodically with period $w_s = \frac{2\pi}{T_s}$ rad/sec. *or* $f_s = \frac{1}{T_s}$ Hz as shown in figure 9.1 (*f*).

Now if have to reconstruct $x(t)$ from $g(t)$, we must be able to recover $X(w)$ from $G(w)$. This is possible if there is no overlap between successive cycles of $G(w)$. Figure 9.1 (*f*) shows that this requires

$$f_s > 2\,f_m \qquad \ldots(9.6)$$

But the sampling interval $T_s = \frac{1}{f_s}$

Hence,

$$T_s < \frac{1}{2\,f_m} \qquad \ldots(9.7)$$

Therefore, as long as the sampling frequency f_s is greater than twice the maximum signal frequency f_m (signal, bandwidth, f_m), $G(w)$ will consist of non-overlapping repetitions of $X(w)$. If this is true, figure 9.1 (f) shows that $x(t)$ can be recovered from its samples $g(t)$ by passing the sampled signal $g(t)$ through an ideal law-pass filter of bandwidth f_m Hz. This proves the sampling theorem.*

9.3.1. Few Points about Sampling Theorem

(*i*) Figure 9.1 (*f*) shows the spectrum of sampled signal. According to the figure, as long as, the signal is sampled at rate $f_s > 2\,f_m$, the spectrum $G(w)$ will repeat periodically without overlapping.

(*ii*) The spectrum of sampled signal extends upto infinity and the ideal bandwidth of sampled signal is infinite. But here our purpose is to extract our original spectrum $X(w)$ out of the spectrum $G(w)$.

(*iii*) The original *or* desired spectrum $X(w)$ is centred at $w = 0$ and is having bandwidth *or* maximum frequency equal to w_m. The desired spectrum may be recovered by passing the sampled signal with spectrum $G(w)$ through a low pass filter with cut-off frequency w_m. This means that since a low-pass filter allows to pass only low frequencies up to cut-off frequency (w_m) and rejects all other higher frequencies, the original spectrum $X(w)$ extended upto w_m will be selected and all other successive higher frequency cycles in the sampled-spectrum will be rejected. Therefore, in this way, original spectrum $X(w)$ will be extracted out of spectrum $G(w)$. This original spectrum $X(w)$ can now be converted into time-domain signal $x(t)$.

* Maximum signal frequency f_m means that the bandwidth of this signal is simply f_m.

(*iv*) It may also be observed from figure that for the case $f_s > 2f_m$, the successive cycles of $G(w)$ are not overlapping each other. Hence in this case, there is no problem in recovering the original spectrum $X(w)$.

(*v*) For the case $f_s = 2\,f_m$, although the successive cycles of $G(w)$ are not overlapping each other, but they are touching each other. In this case also, the original spectrum $X(w)$ can be recovered from the sampled spectrum $G(w)$ using a low-pass filter with cut-off frequency w_m.

(*vi*) For the case $f_s < 2\,f_m$, the successive cycles, of the sampled spectrum will overlap each other and hence in this case, the original spectrum $X(w)$ cannot be extracted out of the spectrum $G(w)$.

Hence, For reconstruction without distortion, we must have

$$f_s \geq 2\,f_m$$

9.4. Nyquist Rate and Nyquist Interval

When the sampling rate becomes exactly equal to $2\,f_m$ samples per second, then it is called Nyquist rate. Nyquist rate is also called the *minimum* sampling rate. It is given by

$$f_s = 2\,f_m \qquad \ldots(9.8)$$

Similarly, maximum sampling interval is called *Nyquist interval*. It is given by

$$\text{Nyquist interval } T_s = \frac{1}{2\,f_m} \text{ seconds} \qquad \ldots(9.9)$$

When the continuous-time bandlimited signal is sampled at Nyquist rate ($f_s = 2\,f_m$), the sampled-spectrum $G(w)$ contains non-overlapping $G(w)$ repeating periodically. But the successive cycles of $G(w)$ touch each other as shown in figure 9.2. Therefore, the original spectrum $X(w)$ can be recovered from the sampled spectrum by using a low pass filter with a cut-off frequency w_m.

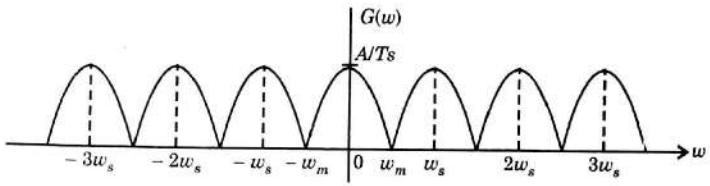


Fig. 9.2. *Sampled spectrum at Nyquist rate.*

Example 9.1. An analog signal is expressed by the equation $x(t) = 3\cos 50\pi t + 10\sin 300\pi t - \cos 100\pi t$. Calculate the Nyquist rate for this signal.

Solution: The given signal is expressed as

$$x(t) = 3\cos 50\pi t + 10\sin 300\pi t - \cos 100\pi t \qquad \ldots(i)$$

Let three frequencies present be w_1, w_2 and w_3

So that the new equation for signal,

$$x(t) = 3\cos w_1 t + 10\sin w_2 t - \cos w_3 t \qquad \ldots(ii)$$

Comparing equations (*i*) and (*ii*) we have

or $\quad w_1 t = 50\,\pi t;\ w_1 = 50\,\pi$

$\therefore \quad 2\,\pi f_1 = 50\,\pi$ or $2\,f_1 = 50$

$f_1 = 25$ Hz

Similarly, for second factor

$$w_2 t = 300\pi t \quad \text{or} \quad w_2 = 300\pi$$

or $\quad 2\pi f_2 = 300\pi \quad$ or $\quad 2\pi f_2 = 300\pi$

$\therefore \quad f_2 = 150$ Hz.

Again, for third factor

$$w_3 t = 100\pi t \quad \text{or} \quad 2\pi f_3 t = 100\pi t$$

or $\quad 2\pi f_3 = 100\pi$

$\therefore \quad f_3 = 50$ Hz

Therefore, the maximum frequency present in $x(t)$ is,

$$f_2 = 150 \text{ Hz}$$

Nyquist rate is given as

$$f_s = 2 f_m$$

where f_m = Maximum frequency present in the signal.

Here $\quad f_m = f_2 = 150$ Hz

Therefore, Nyquist rate

$$f_s = 2 f_2 = 2 \times 150 = 300 \text{ Hz} \qquad \textbf{Ans.}$$

Example 9.2. Find the Nyquist rate and the Nyquist interval for the signal

$$x(t) = \frac{1}{2\pi} \cos(4000\pi t) \cos(1000\pi t)$$

Solution: Given signal is

$$x(t) = \frac{1}{2\pi} \cos(4000\pi t) \cos(1000\pi t) = \frac{1}{4\pi} [2 \cos(4000\pi t) \cos 1000\pi t)]$$

or $$x(t) = \frac{1}{4\pi} [\cos(4000\pi t + 1000\pi t) + \cos(4000\pi t - 1000\pi t)]$$

$$[\because 2 \cos A \cos B = \cos(A + B) + \cos(A - B)]$$

or $$x(t) = \frac{1}{4\pi} [\cos 5000\pi t + \cos 3000\pi t] \qquad ...(i)$$

Let the two frequencies present in the signal be w_1 and w_2 so that the new equation for the signal will be

$$x(t) = \frac{1}{4\pi} [\cos w_1 t + \cos w_2 t] \qquad ...(ii)$$

Comparing equations (*i*) and (*ii*), we have

$$w_1 t = 5000\pi t$$

or $\quad 2\pi f_1 t = 5000\pi t \quad$ or $\quad 2 f_1 = 5000$

$\therefore \quad f_1 = 2500$ Hz

Similarly, for second factor

$$w_2 t = 3000\pi t$$

or $\quad 2\pi f_2 t = 3000\pi t \quad$ or $\quad 2\pi f_2 = 3000$

$\therefore \quad f_2 = 1500$ Hz

Therefore, the maximum frequency present in $x(t)$ is

$$f_1 = 2500 \text{ Hz}$$

Nyquist rate is given as

$$f_s = 2 f_m$$

where f_m = Maximum frequency present in the signal.

Here $\quad f_m = f_1 = 2500$ Hz

Therefore Nyquist rate

$$f_s = 2 f_m = 2 \times 2500 = 5000 \text{ Hz} = 5 \text{ kHz} \qquad \textbf{Ans.}$$

Nyquist interval is given as

$$T_s = \frac{1}{2 f_m} = \frac{1}{2 \times 2500} = \frac{1}{5000}$$

or $$T_s = 0.2 \times 10^{-3} \text{ seconds} = 0.2 \text{ m sec.} \qquad \textbf{Ans.}$$

Example 9.3. A continuous-time signal is given below:

$$x(t) = 8 \cos 200\pi t$$

Determine:

(*i*) Minimum sampling rate *i.e.*, Nyquist rate required to avoid aliasing.

(*ii*) If sampling frequency f_s = 400 Hz. What is the discrete-time signal $x[n]$ or $x[nT_s]$ obtained after sampling?

(*iii*) If sampling frequency f_s = 150 Hz. What is the discrete-time signal $x[n]$ or $x[nT_s]$ obtained after sampling?

(*iv*) What is the frequency $0 < f < f_s/2$ of sinusoidal that yields samples identical to those obtained in part (*iii*)?

Solution:

(*i*) The highest frequency component of continuous-time signal is f = 100 Hz. Hence minimum sampling rate required to avoid aliasing is called Nyquist rate and is given as

$$\text{Nyquist rate} = 2f = 2 \times 100 = 200 \text{ Hz} \qquad \textbf{Ans.}$$

(*ii*) The continuous-time signal $x(t)$ is sampled at f_s = 400 Hz. The frequency of the discrete-time signal will be

$$F = \frac{\text{Frequency of continuous-time signal } f}{\text{Sampling frequency, } f_s} = \frac{100}{400} = \frac{1}{4}$$

Then the discrete-time signal will be given as

$$x[n] = 8 \cos 2\pi F n$$

$$= 8 \cos 2\pi \times \frac{1}{4} n = 8 \cos \frac{\pi n}{2} \qquad \textbf{Ans.}$$

(*iii*) The continuous-time signal $x(t)$ is sampled at f_s = 150 Hz. The frequency of discrete-time will be

$$F = \frac{f}{f_s} = \frac{100}{150} = \frac{2}{3}$$

Then, the discrete-time signal will be given as

$$x[n] = 8 \cos 2\pi f n$$

$$= 5 \cos 2\pi \left(\frac{2}{3}\right) n = 8 \cos \frac{4\pi}{3} n$$

$$= 8 \cos \left(2\pi - \frac{2\pi}{3}\right) n$$

$$= 8 \cos \frac{2\pi n}{3} = 8 \cos \frac{2\pi n}{3} \qquad \textbf{Ans.}$$

(*iv*) For sampling rate of f_s = 150 Hz

$$F = \frac{f}{f_s} \text{ or } f = f_s \times F = \frac{1}{3} \times 150 = 50 \text{ Hz}$$

Then, the sinusoidal signal will be

$$y(t) = 8 \cos 2\pi ft = 8 \cos 2\pi \times 50 \times t = 8 \cos 100\,\pi t$$

Sampling at $f_s = 150$ Hz, yields identical samples hence $f = 100$ Hz is an alias of $f = 50$ Hz for sampling rate $f_s = 150$ Hz. **Ans.**

Example 9.4. Determine the Nyquist rate for a continuous-time signal

$$x(t) = 6 \cos 50\,\pi t + 20 \sin 300\,\pi t - 10 \cos 100\,\pi t$$

Solution: In a general form, any continuous-time signal may be expressed as

$$x(t) = A_1 \cos \omega_1 t + A_2 \cos \omega_2 t + A_3 \cos \omega_3 t \quad ...(i)$$

And the given signal is

$$x(t) = 6 \cos 50\pi t + 2 \sin 300\pi t - 10 \cos 100\,\pi t \quad ...(ii)$$

On comparing given signal (*ii*) with standard form of a signal (*i*), we obtain the frequencies for the given signal as

$$f_1 = \frac{\omega_1}{2\pi} = \frac{50\pi}{2\pi} = 25 \text{ Hz}$$

$$f_2 = \frac{\omega_2}{2\pi} = \frac{300\pi}{2\pi} = 150 \text{ Hz}$$

$$f_3 = \frac{\omega_3}{2\pi} = \frac{100\pi}{2\pi} = 50 \text{ Hz}$$

Thus, the highest frequency component of the given message signal will be

$$f_{max} = 150 \text{ Hz}$$

Therefore, Nyquist rate $= 2\,f_{max} = 2 \times 150 = 300$ Hz **Ans.**

9.5. Reconstruction Filter (Low-Pass Filter)

The low-pass filter is used to recover original signal from its samples. This is also known as **interpolation filter**.

A low-pass filter is that type of filter which passes only low-frequencies upto a specified cut-off frequency and rejects all other frequencies above cut-off frequency. Figure 9.3 shows the frequency response of low-pass filter.

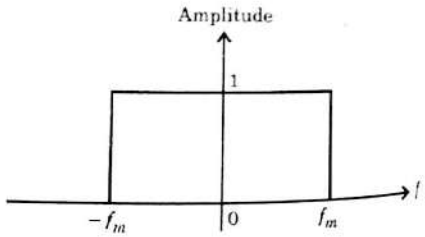


Fig. 9.3. *Ideal low-pass filter (LPF).*

From figure 9.3, it may be observed that in case of low-pass filter, there is sharp-change in response at cut-off frequency, that is amplitude response becomes suddenly zero at cut-off frequency which is not possible practically. This means that an ideal low-pass filter is not physically realizable. In place of ideal-low pass filter, we use practical filter.

Figure 9.4 shows the frequency response of practical low-pass filter. From figure 9.4, it may be observed that in case of practical filter, the amplitude response decreases slowly to become zero. This means that there is a transition band in case of practical filter. Figure 9.5 shows the use of practical low-pass filter in reconstruction of original signal from its sample.

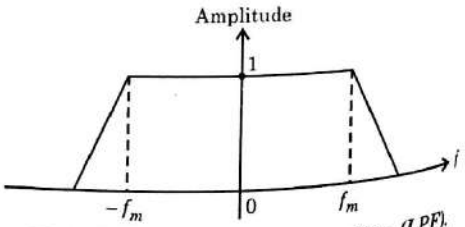


Fig. 9.4. *Practical low-pass filter (LPF).*

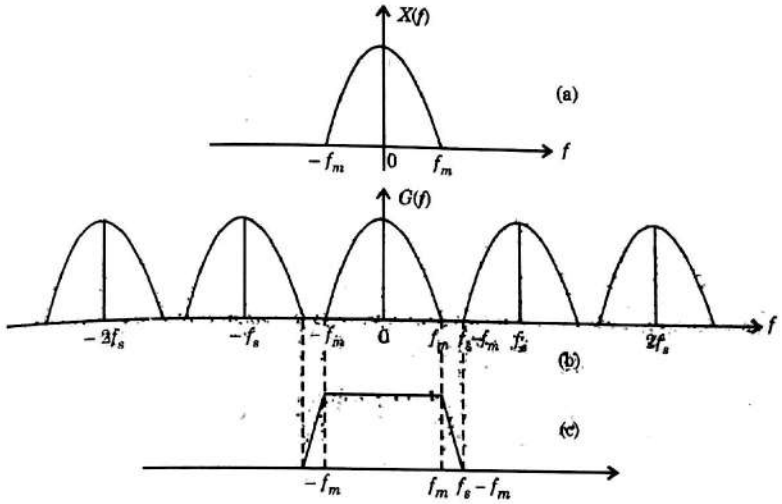


Fig. 9.5. *(a) Spectrum of original signal*
(b) Spectrum of sampled signal
(c) Amplitude response of practical low-pass filter.

9.6. Signal Reconstruction: The Interpolation Formula

The process of reconstructing a continuous-time signal $x(t)$ from its samples is called as **interpolation**.

As discussed earlier, a signal $x(t)$ bandlimited to f_m *Hz* can be reconstructed (interpolated) completely from its samples. This is achieved by passing the sampled signal through an ideal low-pass filter (LPF) of cut-off frequency f_m Hz.

The expression for sampled signal is written as

$$g(t) = x(t).\delta_{T_s}(t) \quad ...(9.10)$$

$$g(t) = \frac{1}{T_s}\,[x(t) + 2\,x(t) \cos w_s t + 2\,x(t) \cos 2w_s t + ...] \quad ...(9.11)$$

From above equation, it may be observed that the sampled signal contains a component $\frac{1}{T_s} \times x(t)$.

To recover $x(t)$ or $X(w)$, the sampled signal must be passed through an ideal low-pass filter of bandwidth of f_mHz and gain T_s.

Therefore, the reconstruction or interpolating filter transfer function may be expressed as

$$H(w) = T_s \times \text{rect}\left(\frac{w}{4\pi f_m}\right) \quad ...(9.12)$$

The impulse response $h(t)$ of this filter is the inverse Fourier transform of $H(w)$, *i.e.*

$$h(t) = F^{-1}[H(w)] = F^{-1}\left[T_s \text{ rect}\left(\frac{w}{4\pi f_m}\right)\right]$$

$$h(t) = 2\,f_m\,T_s \text{ sinc } (2\pi\,f_m t) \quad ...(9.13)$$

Hence if the bandwidth is $2f_m$, then the minimum sampling rate for bandpass signal must be $4\,f_m$ samples per second. Figure 9.8 shows the spectrum of an arbitrary bandpass signal.

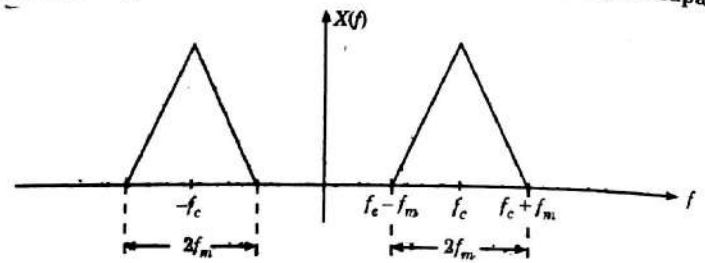


Fig. 9.8. *Spectrum of an orbitrary bandpass signal.*

The spectrum in figure 9.8 is centred around frequency f_c. The bandwidth is $2f_m$. Thus, the frequencies present in the bandpass signal are from $f_c - f_m$ to $f_c + f_m$. This means that the highest frequency present in the bandpass signal is $f_c + f_m$. Generally the centre frequency $f_c > f_m$.

This bandpass signal is first represented in terms of its inphase and quadrature components

Let $\quad x_I(t)$ = Inphase component of $x(t)$

and $\quad x_Q(t)$ = Quadrature component of $x(t)$

Thus, the signal $x(t)$ in terms of inphase and quadrature components will be expressed as

$$x(t) = x_I(t)\cos(2\pi f_c t) - x_Q(t)\sin(2\pi f_c t) \qquad ...(9.18)$$

The inphase and quadrature components are obtained by multiplying $x(t)$ by $\cos(2\pi f_c t)$ and $\sin(2\pi f_c t)$ and then suppressing the sum frequencies by means of low-pass filters. Thus, inphase $x_I(t)$ and quadrature $x_Q(t)$ components contain only low frequency components. The spectrum of these components is limited between $-f_m$ to $+f_m$. This is shown in figure 9.9.

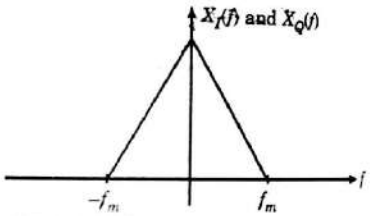


Fig. 9.9. *Spectrum of inphase and quadrature components of bandpass signal x(t).*

After few mathematical manipulation in equation (9.18), we obtain the reconstruction formula as

$$x(t) = \sum_{n=-\infty}^{\infty} x\left(\frac{n}{4f_m}\right)\sin c\left(2f_m t - \frac{n}{2}\right)\cos\left[2\pi f_c\left(t - \frac{n}{4f_n}\right)\right] \qquad ...(9.19)$$

Comparing this reconstruction formula with that of low-pass signals given in equation (9.17), we observe that $x(t)$ is replaced by $x\left(\frac{n}{4f_m}\right)$.

Here, $\quad x\left(\frac{n}{4f_m}\right) = x(nT_s)$ = Sampled version of bandpass signal

and $\quad T_s = \frac{1}{4f_m}$

Thus, if $4f_m$ samples per second are taken, then the bandpass signal of bandwidth $2f_m$ can be completely recovered from its samples.

Hence, for bandpass signals of bandwidth $2f_m$,

Minimum Sampling rate = Twice of bandwidth

= $4\,f_m$ samples per second.

Example 9.5. Show that a bandlimited signal of finite energy which has no frequency components higher than f_m Hz is completely described by specifying values of the signals at instants of time separated by $1/2f_m$ seconds and also show that if the instantaneous values of the signal are separated by intervals larger than $1/2\,f_m$ seconds, they fail to describe the signal. A bandpass signal has spectral range that extends from 20 to 82 kHz.

Find the acceptable range of sampling frequency f_s.

Solution: Let $x(t)$ be the bandlimited signal which has no frequency components higher than f_m Hz. Let this signal be sampled by a sampling function given as

$$\delta_{T_s}(t) = \sum_{n=-\infty}^{\infty}\delta(t - nT_s)$$

The sampling function is the train of impulses with T_s as distance between successive impulses. Let $x(nT_s)$ be the instantaneous amplitude of signal $x(t)$ at instant $t = T_s$. The sampled version of $x(t)$ may be given as multiplication of $x(nT_s)$ and $\delta_{T_s}(t)$ *i.e.*,

$$g(t) = \sum_{n=-\infty}^{\infty} x(nT_s)\,\delta(t - nT_s) \qquad ...(i)$$

Now, Fourier transform of this sampled signal may be obtained as

$$G(f) = FT\{g(t)\} = f_s \sum_{n=-\infty}^{\infty} X(f - nf_s)$$

Here, f_s is the sampling rate which is given as

$$f_s = \frac{1}{T_s}$$

and $\quad X(f - nf_s) = X(f)$ at $nf_s = 0, \pm f_s, \pm 2f_s \pm 3f_s + ...$

Hence, the same spectrum $X(f)$ appears at $f = 0$,

$$f = \pm f_s,\ f = \pm 2f_s \quad \text{etc.,}$$

This means that a periodic spectrum with period equal to f_s is generated in frequency domain because of sampling $x(t)$ in time-domain.

Thus, equation (*i*) may be written as

$$G(f) = f_s \times (f) + f_s \times (f \pm f_s) + f_s \times (f \pm 2f_s) + f_s \times (f \pm 3f_s) + f_s \times (f \pm 4f_s) + ... \qquad ...(ii)$$

or

$$G(f) = f_s \times (f) + \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} f_s \times (f - nf_s) \qquad ...(iii)$$

By definition of Fourier transform, we have

$$X(f) = \int_{-\infty}^{\infty} x(t)\,e^{-j2\pi ft}\,dt$$

For sampled version of $x(t)$, we have $t = nT_s$.

Then the equation (*iii*) becomes

$$G(f) = \sum_{n=-\infty}^{\infty} x(nT_s)\,e^{-j2\pi fnT_s} \qquad ...(iv)$$

Assuming that sampling is done at Nyquist rate, then

$$T_s = \frac{1}{2f_m}$$

So that $\quad 2 f_m T_s = 1$

Putting this value of $2 f_m T_s$ in equation (9.13), we have ...(9.13a)

$$h(t) = 1 \cdot \text{sinc}\,(2\pi f_m t) = \text{sinc}\,(2\pi f_m t) \quad ...(9.14)$$

Figure 9.6(*b*) shows the graph of $h(t)$.

From figure, it may be observed that $h(t) = 0$ at all Nyquist sampling instants $t = \pm\, n/2f_m$ except at $t = 0$.

Now, when the sampled signal $g(t)$ is applied at the input of this filter, the output will be $x(t)$.

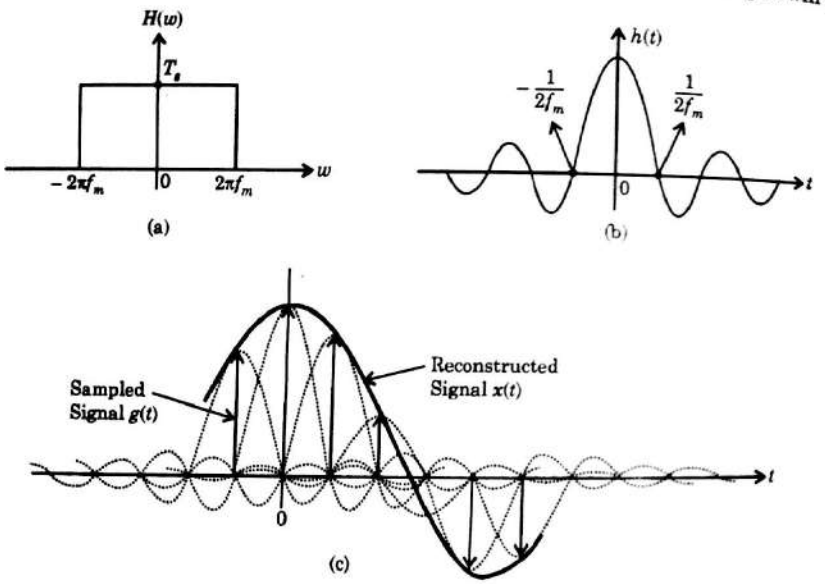


Fig. 9.6.

Each sample in $g(t)$, being an impulse, produces a sinc pulse of height equal to the strength of the sample.

Addition of the sinc pulses produced by all the samples results in $x(t)$.

For instant, the kth sample of the input $g(t)$ is the impulse $x(kT_s)\,\delta(t - kT_s)$.

The filter output of this impulse will be $x(kT_s)\,h(t - kT_s)$.

Therefore, the filter output to $g(t)$, which is $x(t)$, may be expressed as a sum

$$x(t) = \sum_k x(kT_s)\,h(t - kT_s) \quad ...(9.15)$$

$$= \sum_k x(kT_s)\,\text{sinc}\,[2\pi f_m (t - kT_s)] \quad ...(9.16)$$

$$x(t) = \sum_k x(kT_s)\,\text{sinc}\,(2\pi f_m - k\pi) \quad \left[\because T_s = \frac{1}{2f_m}\right] \quad ...(9.17)$$

Equation (9.17) is known as the **interpolation formula**, which provides values of $x(t)$ between samples as a weighted sum of all the sample values.

In the proof of sampling theorem, it is assumed that the signal $x(t)$ is strictly bandlimited. But, in general, an information signal may contain a wide range of frequencies and cannot be strictly bandlimited. This means that the maximum frequency f_m in the signal $x(t)$ cannot be predictable. Therefore, it is not possible to select suitable sampling frequency f_s.

9.7. Effect of Under Sampling: Aliasing

When a continuous-time bandlimited signal is sampled at a rate lower than Nyquist rate $f_s < 2 f_m$, then the successive cycles of the spectrum $G(w)$ of the sampled signal $g(t)$ overlap with each other as shown in figure 9.7.

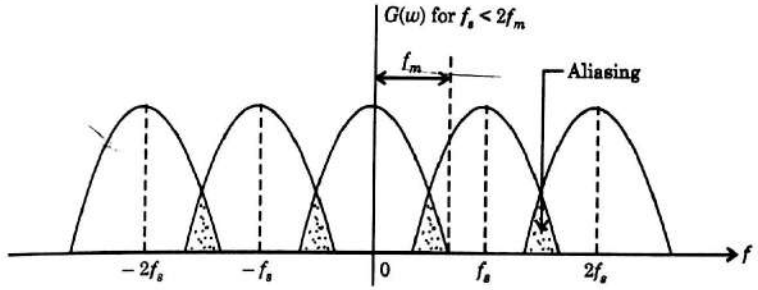


Fig. 9.7. *Spectruym of the sampled signal for the case $f_s < 2f_m$.*

Hence, the signal is under-slampled in this case ($f_s < 2 f_m$) and some amount of aliasing is produced in this under-sampling process. In fact, aliasing is the phenomenon in which a high frequency component in the frequency-spectrum of the signal takes identity of a lower-frequency component in the spectrum of the sampled signal.

From figure 9.7, it is obvious that because of the overlap due to aliasing phenomenon, it is not possible to recover original signal $x(t)$ from sampled signal $g(t)$ by low-pass filtering since the spectral components in the overlap regions add and hence the signal is distorted.

Since any information signal contains a large number of frequencies, so, to decide a sampling frequency is always a problem. Therefore, a signal is first passed through a low-pass filter. This low-pass filter blocks all the frequencies which are above f_m Hz. This process is known as band-limiting of the original signal $x(t)$. This low-pass filter is called **prelias filter** because it is used to prevent aliasing effect. After bandlimiting, it becomes easy to decide sampling frequency since the maximum frequency is fixed at f_m Hz.

In short, to avoid aliasing:

(*i*) Prealias filter must be used to limit band of frequencies of the signal to f_m Hz.

(*ii*) Sampling frequency 'f_s' must be selected such that

$$f_s > 2 f_m$$

9.8. Sampling of Bandpass Signals ✓

In previous sections, we discussed sampling theorem for low-pass signals. However, when the given signal is a bandpass signal, then a different criteria must be used to sample the signal. Therefore, the sampling theorem for bandpass signals may be expressed as under:

The bandpass signal $x(t)$ whose maximum bandwidth is $2f_m$ can be completely represented into and recovered from its samples if it is sampled at the minimum rate of twice the bandwidth. Here, f_m is the maximum frequency component present in the signal.

Now, given that the signal is bandlimited to f_m Hz and

$$T_s = \frac{1}{2f_m} \text{ seconds}$$

Therefore, $$f_s = \frac{1}{T_s} = 2f_m \qquad ...(v)$$

From equation (*ii*), it may be observed that $G(f)$ is periodic with a period f_s. Thus the spectrum $X(f)$ and $G(f)$ are shown in figure 9.10.

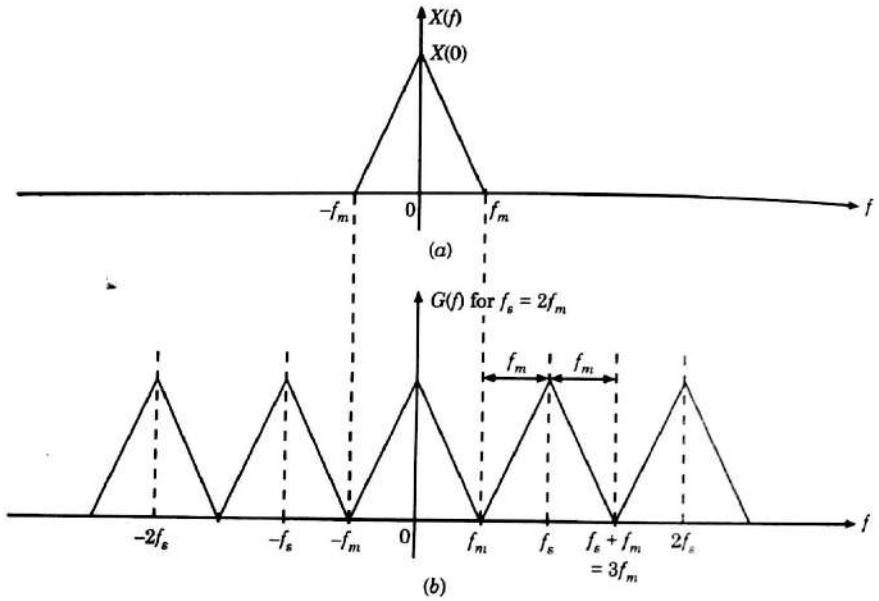


Fig. 9.10. *(a) Spectrum of x(t) (b) Spectrum of g(t) with $f_s = 2f_m$.*

Now, since $f_s = 2f_m$

therefore, $f_s - f_m = f_m$ and $f_s + f_m = 3f_m$

Hence, the periodic spectrums $X(f)$ just touch $\pm f_m, \pm 3f_m, \pm 5f_m$ ... etc.

Thus, there is no aliasing.

Using equation (*iii*), we write

$$X(f) = \frac{1}{f_s} G(f) - \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} X(f - nf_s) \qquad ...(vi)$$

Substituting $f_s = 2f_m$ in equation (*vi*), we get

$$X(f) = \frac{1}{2f_m} G(f) - \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} X(f - nf_s)$$

or $$X(f) = \frac{1}{2f_m} G(f) \quad \text{for } -f_m \le f \le f_m \qquad ...(vii)$$

Now putting the value of $G(f)$ from equation (*iv*) to (*vii*), we get

$$X(f) = \frac{1}{2f_m} \sum_{n=-\infty}^{\infty} x(nT_s)\, e^{-j2\pi f n.T_s}$$

Since $$T_s = \frac{1}{2f_m}$$

i.e., $$X(f) = \frac{1}{2f_m} \sum_{n=-\infty}^{\infty} x\left(\frac{n}{2f_m}\right) e^{-j\pi f n/f_m} \qquad ...(viii)$$

$x(t)$ may, be recovered from $X(f)$ by taking Inverse Fourier transform of last equation, *i.e.*,

$$x(t) = IFT\left\{\frac{1}{2f_m} \sum_{n=-\infty}^{\infty} x\left(\frac{n}{2f_m}\right) e^{-j\pi f_n/f_m}\right\} \qquad ...(ix)$$

This equation indicates that $x(t)$ is represented completely by its samples $x\left(\frac{n}{2f_m}\right)$ for $-\infty < n < \infty$. This means that the sequence $x\left(\frac{n}{2f_m}\right)$ has all the information contained in $x(t)$.

Reconstruction of Signal form Samples

Let us consider equation (*ix*) as

$$x(t) = IFT\left\{\frac{1}{2f_m} \sum_{n=-\infty}^{\infty} x\left(\frac{n}{2f_m}\right) e^{-j\pi f n/f_m}\right\} = \int_{-f_m}^{f_m} \frac{1}{2f_m} \sum_{n=-\infty}^{\infty} x\left(\frac{n}{2f_m}\right) e^{-j\pi f n/f_m} e^{j2\pi f t}\, dt$$

Interchanging the order of summation and integration, we get

$$x(t) = \sum_{n=-\infty}^{\infty} x\left(\frac{n}{2f_m}\right) \frac{1}{2f_m} \int_{-f_m}^{f_m} e^{j2\pi f\left(t - \frac{n}{2f_m}\right)} dt$$

or $$x(t) = \sum_{n=-\infty}^{\infty} x\left(\frac{n}{2f_m}\right) \frac{\sin(2\pi f_m t - n\pi)}{(2\pi f_m t - n\pi)} = \sum_{n=-\infty}^{\infty} x\left(\frac{n}{2f_m}\right) \frac{\sin \pi(2f_m t - n\pi)}{\pi(2f_m t - n)}$$

Since, $$\text{sinc}\, \theta = \frac{\sin \pi\theta}{\pi\theta}$$

Therefore, $$x(t) = \sum_{n=-\infty}^{\infty} x\left(\frac{n}{2f_m}\right) \text{sinc}\,(2f_m t - n) \quad -\infty < n < \infty$$

Hence, this is the interpolation formula to reconstruct $x(t)$ from its samples $x(nT_s)$. Therefore, from all above, it is clear that the signal may be completely represented into and recovered from its samples if the spacing between the successive samples is $\frac{1}{2f_m}$ seconds *i.e.*, $f_s = 2f_m$ samples per second.

Sampling Frequency for Bandpass Signal

Since the spectral range of the bandpass signal is 20 kHz to 82 kHz

Therefore

$$\text{Bandwidth} = 2f_m = 82 \text{ kHz} - 20 \text{ kHz} = 62 \text{ kHz}$$

Hence, Minimum Sampling rate = 2 × bandwidth = 2 × 62 = 124 kHz

Generally, the range of minimum sampling frequencies is specified for bandpass signals. It lies between $4f_m$ to $8f_m$ samples per second.

Therefore,

Range of minimum sampling frequencies

= (2 × bandwidth) to (4 × bandwidth)

= 2 × 62 kHz to 4 × 62 kHz = 124 kHz to 248 kHz **Ans.**

9.9. Sampling Techniques

In the last article, we discussed how sampling of a continuous-time signal is done. This sampling of a signal is done in several ways. Therefore, in this section, we shall discuss different types of sampling *i.e.*, sampling techniques.

Basically, there are three types of sampling techniques as under:

(*i*) Instantaneous sampling

(*ii*) Natural sampling

(*iii*) Flat-top sampling.

Out of these three, instantaneous sampling is called ideal sampling whereas natural sampling and flat-top sampling are called practical sampling methods. Now, let us discuss three different types of sampling techniques in detail.

9.9.1. Ideal Sampling or Instantaneous Sampling or Impulse Sampling

In the proof of sampling theorem, we used ideal or impulse sampling. In this type of sampling, the sampling function is a train of impulses. Figure 9.11(*b*) shows this sampling function.

$x(t)$ is the input signal (*i.e.*, signal to be sampled) as shown in figure 9.11(*a*).

Figure 9.11(*c*) shows a circuit to produce instantaneous or ideal sampling. This circuit is known as the **switching sampler**.

The working principle of this circuit is quite easy. The circuit simply consists of a switch. Now if we assume that the closing time '*t*' of the switch approaches zero, then the output $g(t)$ of this circuit will contain only instantaneous value of the input signal $x(t)$. Since the width of the pulse approaches zero, the instantaneous sampling gives a train of impulses of height equal to the instantaneous value of the input signal $x(t)$ at the sampling instant.

We know that the train of impulses may be represented as

$$\delta T_s(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT_S) \qquad ...(9.20)$$

This is known as sampling function and its waveform is shown in figure 9.11(*b*).

The sampled signal $g(t)$ is expressed as the multiplication of $x(t)$ ans $\delta_{Ts}(t)$.

Thus,

$$g(t) = x(t) \cdot \delta_{Ts}(t) \qquad ...(9.21)$$

$$= x(t) \cdot \sum_{n=-\infty}^{\infty} \delta(t - nT_s) \qquad ...(9.22)$$

or

$$g(t) = \sum_{n=-\infty}^{\infty} x(nT_s) \cdot \delta(t - nT_s) \qquad ...(9.23)$$

The Fourier transform of the ideally sampled signal given by above equation may be expressed as

$$G(f) = f_s \sum_{n=-\infty}^{\infty} X(f - nf_s) \qquad ...(9.24)$$

Note: *This equation gives the spectrum of ideally sampled signal. It shows that the spectrum $X(f)$ is periodic in f_s and weighted by f_s. However, it may be noted that ideal or instantaneous sampling is possible only in theory since it is impossible to have a pulse whose width approaches zero. Ideal sampling was used in last article to prove sampling theorem. Practically flat-top sampling and nature sampling are used.*

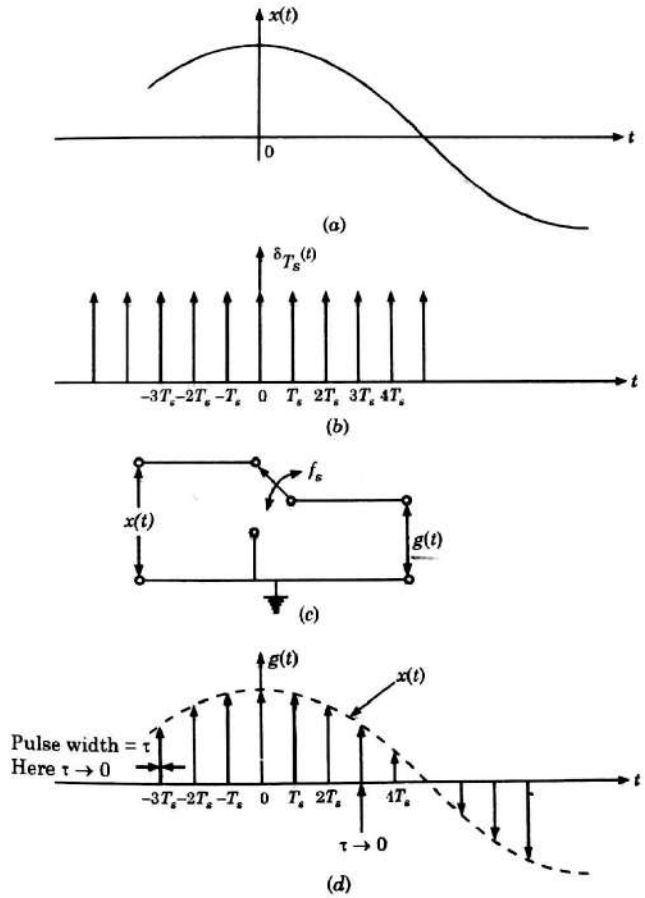


Fig. 9.11. *(a) Baseband signal (b) Impulse train (c) Functional diagram of a switching sampler (d) Sampled signal*

9.9.2. Natural Sampling

As discussed in last article, the instantaneous sampling results in the samples whose width τ approaches zero. Due to this, the power content in the instantaneously sampled pulse is negligible. Thus, this method is not suitable for transmission purpose. Natural sampling is a practical method and will be discussed in this section.

In natural sampling the pules has a finite width equal to τ.

Let us consider an analog continuous-time signal $x(t)$ to be sampled at the rate of f_s Hz. Here it is assumed that f_s is higher than Nyquist rate such that simpling theorem is satisfied.

Again, let us consider a sampling function $c(t)$ which is a train of periodic pulses of width τ and frequency equal to f_s Hz.

Figure 9.12 shows a functional diagram of a natural sampler. With the help of this natural sampler, a sampled signal $g(t)$ is obtained by multiplication of sampling function $c(t)$ and input signal $x(t)$.

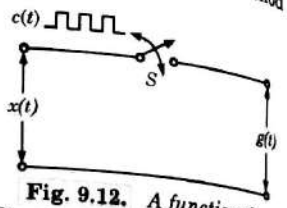


Fig. 9.12. *A functional diagram of a Natural Sampler.*

Now, according to figure 9.12, we have

when $c(t)$ goes high the switch 'S' is closed.

Therefore $\quad g(t) = x(t) \quad$ when $\quad c(t) = A$ ...(9.25)

and $\quad g(t) = 0 \quad$ when $\quad c(t) = 0$ ...(9.26)

where A is the amplitude of $c(t)$.

The waveforms of signals $x(t)$, $c(t)$ and $g(t)$ have been illustrated in figure 9.13(*a*), (*b*) and (*c*) respectively.

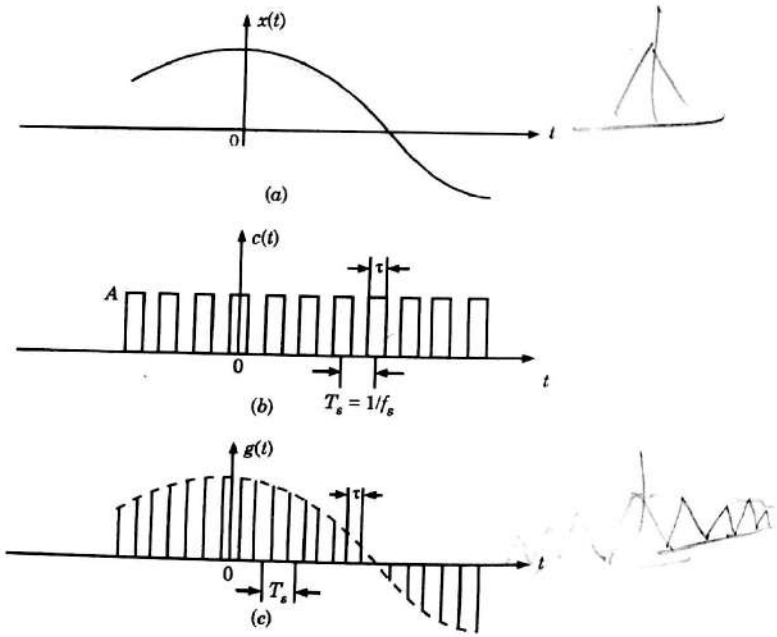


Fig. 9.13. *(a) Continuous time signal x(t).*
(b) Sampling function waveform i.e. periodic pulse train
(c) Naturally sampled signal waveform s(t).

Now, the sampled signal $g(t)$ may also be described mathematically as

$$g(t) = c(t) \cdot x(t) \qquad ...(9.27)$$

Here, $c(t)$ is the periodic train of pulse of width τ and frequency f_s.

We know that the Exponential Fourier series for any periodic waveform is expressed as

$$s(t) = \sum_{n=-\infty}^{\infty} C_n e^{j2\pi nt/T_0} \qquad ...(9.28)$$

Also, for the periodic pulse train of $c(t)$, we have

$$T_0 = T_s = \frac{1}{f_s} = \text{period of } c(t)$$

or

$$f_0 = f_s = \frac{1}{T_0} = \frac{1}{T_s} = \text{frequency of } c(t)$$

Therefore according to equation (9.28) for periodic pulse train $c(t)$, we have

$$c(t) = \sum_{n=-\infty}^{\infty} C_n . e^{j2\pi f_s nt} \text{ with } \frac{1}{T_0} = f_s \qquad ...(9.29)$$

Now, it may be noted that since $c(t)$ is a rectangular pulse train, therefore C_n for this waveform will be expressed as

$$C_n = \frac{TA}{T_0} \sin c\,(f_n . T) \qquad ...(9.30)$$

here $\quad T$ = pulse width = τ

and $\quad f_n$ = harmonic frequency

But here, $\quad f_n = nf_s \quad$ or $\quad f_s = \dfrac{n}{T_0} = nf_0$

Hence,

$$C_n = \frac{\tau . A}{T_s} \cdot \sin c(f_n . \tau) \qquad ...(9.31)$$

Therefore, using equations (9.29) and (9.30) the Fourier series representation for $c(t)$ will be given as

$$c(t) = \sum_{n=-\infty}^{\infty} \frac{\tau . A}{T_s} \cdot \sin c(f_n . \tau)\, e^{j2\pi f_s \cdot nt} \qquad ...(9.32)$$

Now, substituting the value of $c(t)$ from equation (9.32) to equation (9.27), we get

$$g(t) = \frac{\tau A}{T_s} \cdot \sum_{n=-\infty}^{\infty} \sin c(f_n . \tau) . e^{j2\pi f_s nt} . x(t) \qquad ...(9.33)$$

This is required time-domain representation for naturally sampled signal $g(t)$.

Now, to get the frequency-domain representation of the naturally sampled signal $g(t)$, let us take its Fourier transform as

$$G(f) = FT[g(t)] = \frac{\tau A}{T_s} \sum_{n=-\infty}^{\infty} \text{si nc}\,(f_n . \tau)\, FT[e^{j2\pi f_s nt} . x(t)] \qquad ...(9.34)$$

Recall the frequency-shifting property of Fourier transform which states that

$$e^{j2\pi f_s n(t)} . x(t) \longleftrightarrow X(f - f_s . n) \qquad ...(9.35)$$

Therefore, $$G(f) = \frac{\tau A}{T_s} \cdot \sum_{n=-\infty}^{\infty} \text{sinc}\,(f_n.\tau)\, X(f - f_s n) \quad ...(9.36)$$

Now, since $f_n = f_s n$ = harmonic frequency
Therefore, equation (9.36) becomes

$$G(f) = \frac{\tau A}{T_s} \cdot \sum_{n=-\infty}^{\infty} \text{sinc}\,(nf_s.\tau)\, X\,(f - nf_s) \quad ...(9.37)$$

Hence, we write
Spectrum of Naturally Sampled signal:

$$G(f) = \frac{\tau A}{T_s} \cdot \sum_{n=-\infty}^{\infty} \text{sinc}\,(nf_s.\,\tau)\, X(f - nf_s) \quad ...(9.38)$$

This equation shows that the spectra of $x(t)$ *i.e.*, $X(f)$ are periodic in f_s and are weighed by the sin c function.

Figure 9.14, illustrates some arbitray spectra for $x(t)$ and corresponding spectrum $G(f)$.

Note: *Thus form figure 9.14, it may be noted that unlike the spectrum of instantaneously sampled signal shown in figure 9.1(f), the spectrum of a naturally sampled signal is weighted by a sin c function whereas the spectrum of an instantaneously sampled signal (figure 9.1(f)) remains constant throughout the frequency range.*

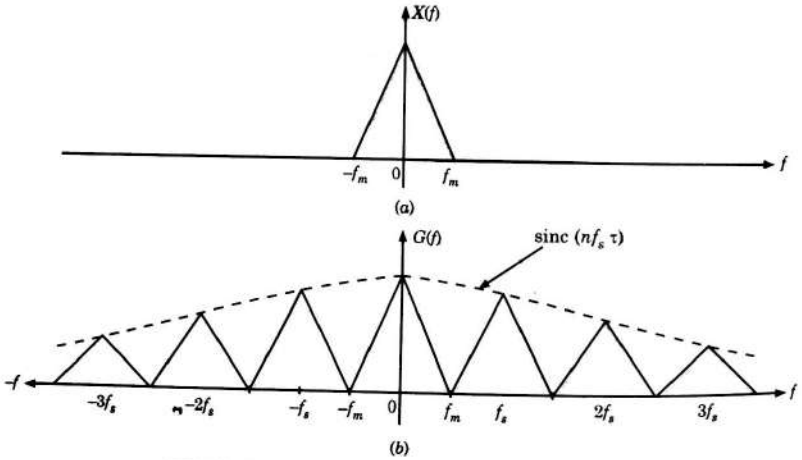


Fig. 9.14. *(a) Spectrum of continuous-time signal x(t),*
(b) Spectrum of naturally sampled signal

9.9.3. Flat Top Sampling or Rectangular Pulse Sampling

(U.P. Tech., Semester Examination, 2003-2004)

Flat top sampling like natural sampling is also a practically possible sampling method. But natural sampling is little complex whereas it is quite easy to get flat top samples.

In flat-top sampling or rectangular pulse sampling, the top of the samples remains constant and is equal to the instantaneous value of the baseband signal $x(t)$ at the start of sampling. The duration or width of each sample is τ and sampling rate is equal to $f_s = \frac{1}{T_s}$. Figure 9.15(*a*) shows

the functional diagram of a sample and hold circuit which is used to generate the flat top samples. Figure 9.15(*b*) shows the general waveform of flat top samples.

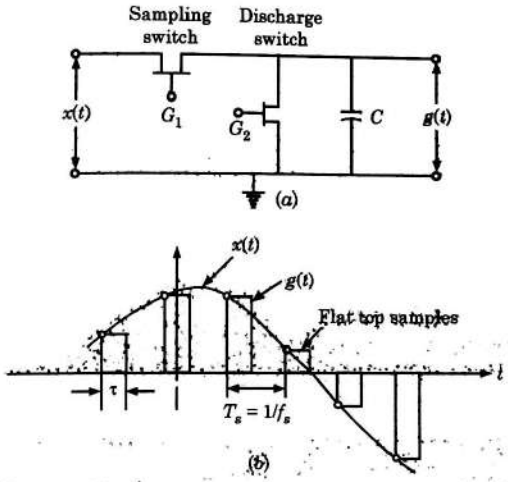


Fig. 9.15. *(a) A sample and hold circuit to generate flat top samples.*
(b) A general waveform of flat top sampling.

From figure 9.15(*b*), it may be noted that only starting edge of the pulse represents instantaneous value of the baseband signal $x(t)$. Also the flat top pulse of $g(t)$ is mathematically equivalent to the convolution of instantaneous sample and a pulse $h(t)$ as depicted in figure 9.16.

This means that the width of the pulse in $g(t)$ is determined by the width of $h(t)$ and the sampling instant is determined by delta function.

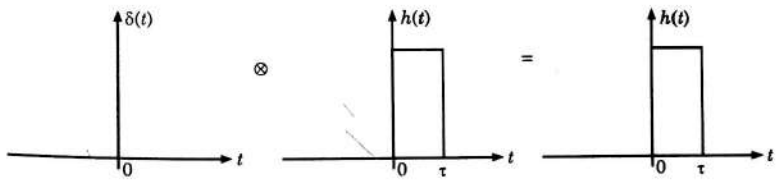


Fig. 9.16. *Convolution of any function with delta function is equal to that function.*

In figure 9.15(*b*), the starting edge of the pulse represents the point where baseband signal is sampled and width is determined by function $h(t)$. Therefore $g(t)$ will be expressed as

$$g(t) = s(t) \otimes h(t) \quad ...(9.39)$$

This equation has been explained in figure 9.17.

Now, from the property of delta function, we know that for any function $f(t)$

$$f(t) \otimes \delta(t) = f(t) \quad ...(9.40)$$

This property is used to obtain flat top samples. It may be noted that to obtain flat top sampling, we are not applying the equation (9.40) directly here *i.e.*, we are applying a modified form of equation (9.40). This modified equation is equation (9.39).

Thus, in this modified equation, we are taking $s(t)$ in place of delta function $\delta(t)$. Observe that $\delta(t)$ is a constant amplitude delta function whereas $s(t)$ is a varying amplitude train of impulses. This means that we are taking $s(t)$ which is an istantaneously sampled signal and this is convolved with function $h(t)$ as in equation (9.39).

Therefore, on convolution of $s(t)$ and $h(t)$, we get a pulse whose duration is equal to $h(t)$ only but amplitude is defined by $s(t)$.

Now, we know that the train of impulses may be represented mathematically as

$$\delta_{Ts}(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT_s) \quad ...(9.41)$$

The signal $s(t)$ is obtained by multiplication of baseband signal $x(t)$ and $\delta_{Ts}(t)$.

Thus,

$$s(t) = x(t) \cdot \delta_{Ts}(t) \quad ...(9.42)$$

$$= x(t) \sum_{n=-\infty}^{\infty} \delta(t - nTs) = \sum_{n=-\infty}^{\infty} x(nT_s) \cdot \delta(t - nT_s) \quad ...(9.43)$$

Now, sampled signal $g(t)$ is given as equation (9.39)

$$g(t) = s(t) \otimes h(t) \quad ...(9.44)$$

or

$$g(t) = \int_{-\infty}^{\infty} s(\tau) h(t - \tau)\, d\tau \quad ...(9.45)$$

or

$$g(t) = \int_{-\infty}^{\infty} \sum_{n=-\infty}^{\infty} x(nT_s)\, \delta(\tau - nT_s) h(t - \tau)\, d\tau = \sum_{n=-\infty}^{\infty} x(nT_s) \int_{-\infty}^{\infty} \delta(\tau - nT_s) h(t - \tau)\, d\tau \quad ...(9.46)$$

According to shifting property of delta function, we know that

$$\int_{-\infty}^{\infty} f(t)\, \delta(t - t_0) = f(t_0) \quad ...(9.47)$$

Using equations (9.46) and (9.47), we get

$$g(t) = \sum_{n=-\infty}^{\infty} x(nT_s)\, h(t - nT_s)$$

This equation represents value of $g(t)$ in terms of sampled value $x(nT_s)$ and function $h(t - nT_s)$ for flat top sampled signal.

Now, again from equation (9.39), we have

$$g(t) = s(t) \otimes h(t)$$

Taking Fourier transform of both sides of above equation, we get

$$G(f) = S(f)\, H(f) \quad ...(9.48)$$

We know that $S(f)$ is given as

$$S(f) = f_s \sum_{n=-\infty}^{\infty} X(f - nf_s) \quad ...(9.49)$$

Therefore, equation (9.48) becomes

$$G(f) = f_s \sum_{n=-\infty}^{\infty} X(f - nf_s) \cdot H(f) \quad ...(9.50)$$

Thus, spectrum of flat top sampled signal:

$$G(f) = f_s \sum_{n=-\infty}^{\infty} X(f - nf_s)\, H(f) \quad ...(9.51)$$

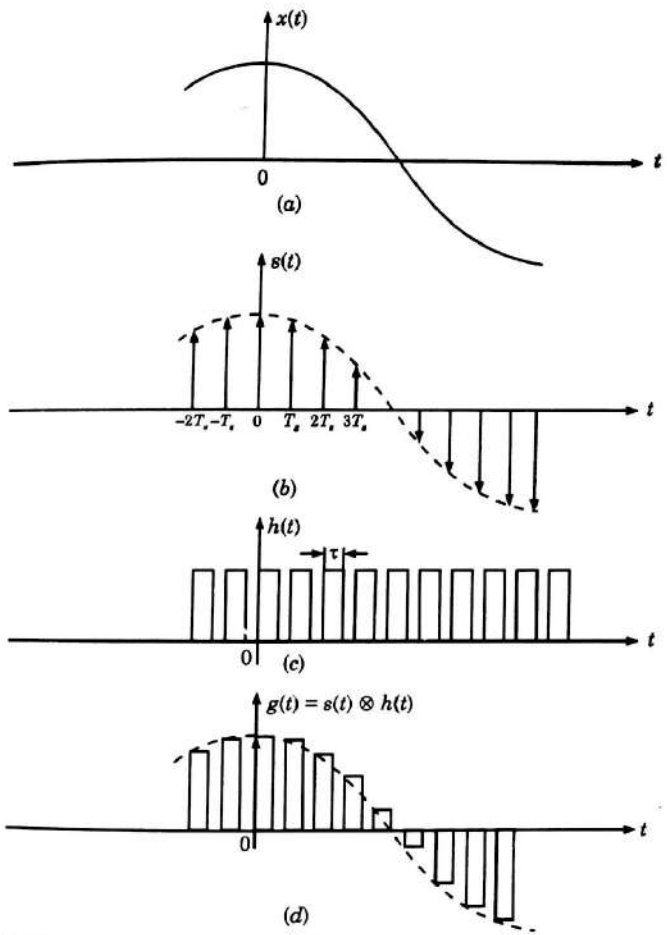


Fig. 9.17. *(a) Baseband signal x(t) (b) Instantaneously sample single s(t) (c) Constant pulse width function h(t) (d) Flat top sampled signal g(t) obtained through convolution of h(t) and s(t)*

9.10. Aperture Effect

The spectrum of flat top sampled signal is expressed as

$$G(f) = f_s \sum_{n=-\infty}^{\infty} X(f - nf_s) \cdot H(f) \quad ...(9.52)$$

This equation shows that the signal $g(t)$ is obtained by passing the signal $s(t)$ through a filter having transfer function $H(f)$. The corresponding impulse response $h(t)$ in time-domain has been

shown in figure 9.18(*a*). This is one pulse of rectangular pulse train shown in figure 9.17(c). Each sample of $x(t)$ [*i.e.*, $s(t)$] is convolved with this pulse. Equation (9.52) represents that the spectrum of this rectangular pulse is multiplied with that of $s(t)$.

Figure 9.18(*b*) shows the spectrum of one rectangular pulse of $h(t)$.

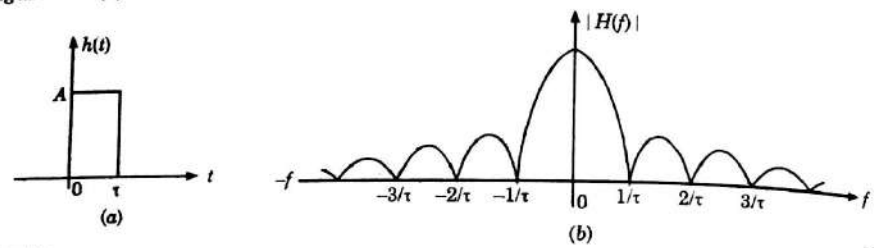


Fig. 9.18. *(a) One pulse of rectangular pulse train (b) Spectrum of the pulse shown in figure (a).*

We know that the spectrum of a rectangular pulse is expressed as

$$H(f) = \tau.\sin c(f \cdot t)\, e^{-j\pi f\tau} \quad [\because A = 1] \qquad ...(9.53)$$

Hence, from figure 9.18(*b*), it may be observed that by using flat top samples an amplitude distortion is introduced in the reconstructed signal $x(t)$ from $g(t)$. In fact, the high frequency roll-off of $H(f)$ acts like a low-pass filter and thus attenuates the upper portion of message signal spectrum. These high frequencies of $x(t)$ are affected. This type of effect is known as aperture effect.

Now, as the duration 'τ' of the pulse increases, the aperture effect is more prominent. Hence, during reconstruction an equalizer is needed to compensate for this effect. As depicted in figure 9.19, the receiver contains a low-pass reconstruction filter with cutoff frequency slightly higher than the maximum frequency present in the message signal. The equalizer compensates for the aperture effect. It also compensates for the attenuation by a low-pass reconstruction filter.

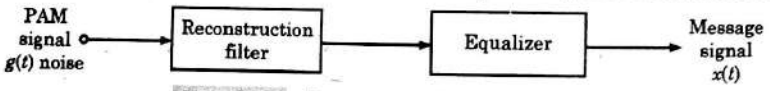


Fig. 9.19. *Recovering x(t) at receiver.*

From equation (9.53), it may be noted that the sample function $h(t)$ acts like a low-pass filter where Fourier transform as expressed as

$$H(f) = \tau.\ \text{sinc}\,(f.\tau)\, e^{-j\pi f\tau} \qquad ...(9.54)$$

This spectrum has been plotted in figure 9.18.

Equalizer used in casecade with the reconstruction filter has the effect of decreasing the inband loss of the reconstruction filter as the frequency increases in such a way as to compensate for the aperture effect.

Also, the transfer function of the equalizer is expressed as

$$H_{eq}(f) = \frac{K \cdot e^{-j2\pi f t_d}}{H(f)} \qquad ...(9.55)$$

Here 't_d' is known as the delay introduced by low-pass filter which is equal to $\tau/2$.

Therefore,
$$H_{eq}(f) = \frac{K \cdot e^{-j\pi f\tau}}{\tau\,\text{sinc}\,(f\tau)\, e^{-j\pi f\tau}} \qquad ...(9.56)$$

or
$$H_{eq}(f) = \frac{K}{\tau\,\text{sinc}\,(f\tau)} \qquad ...(9.57)$$

which is the transfer function of an equalizer.

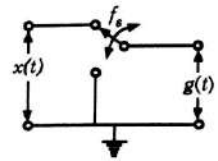
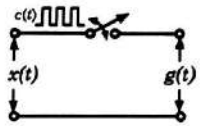
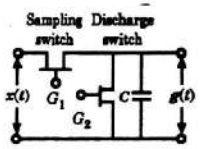
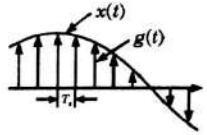
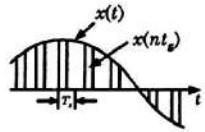
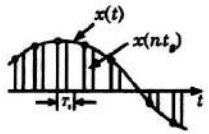
9.11. Comparison of Various Sampling Techniques

We can compare various sampling techniques on the basis of their method, noise interference and spectral properties etc.

The Table 9.1 lists some of the important points of comparison of generation.

Table 9.1. Comparison of three sampling techniques

(U.P. Tech., Semester, Examination, 2003-04)

S. No.	Parameter of comparison	Ideal or instantaneous sampling	Natural sampling	Flat top sampling
1.	Sampling principle	It uses multiplication	It uses chopping principle .	It uses sample and hold circuit
2.	Generation circuit			
3.	Waveforms involved			
4.	Feasibility	This is not a practically possible method	This method is used practically	This method is also used practically
5.	Sampling rate	Sampling rate tends to infinity	Sampling rate satisfies Nyquist criteria	Sampling rate satisfies Nyquist criteria
6.	Noise interference	Noise interference is maximum	Noise interference is minimum	Noise interference is maximum
7.	Time domain representation	$g(t) = \sum_{n=-\infty}^{\infty} x(nT_s)\delta(t - nT_s)$	$g(t) = \frac{\tau A}{T_s}\sum_{n=-\infty}^{\infty} x(t)\sin c(nf_s\tau)\, e^{j2\pi nf_s t}$	$g(t) = \sum_{n=-\infty}^{\infty} x(nT_s)h(t - nT_s)$
8.	Frequency domain representation	$G(f) = f_x \sum_{n=-\infty}^{\infty} X(f - nf_s)$	$G(f) = \frac{\tau A}{T_s}\sum_{n=-\infty}^{\infty} \sin c(nf_s\tau)X(f - nf_s)$	$G(f) = f_s \sum_{n=-\infty}^{\infty} X(f - nf_s)H(f)$

Example 9.6. Figure 9.20 shows the spectrum of an arbitary signal $x(t)$. This signal is sampled at the Nyquist rate with a periodic train of rectangular pulses of duration $\frac{50}{3}$ milliseconds.

Determine the spectrum of the sampled signal for frequencies upto 50 Hz giving relevant expression.

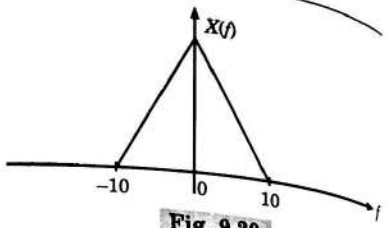


Fig. 9.20.

Solution: From figure 9.20, it may be observed that the signal is bandlimited to 10 Hz.

thus $f_m = 10$ Hz

So the Nyquist rate is $= 2 f_m$

$= 2 \times 10 = 20$ Hz

Since the signal is sampled at the Nyquist rate, the sampling frequency would be

$$f_s = 20 \text{ Hz}$$

Given that the rectangular pulses are used for sampling *i.e.*, flat top sampling is used. The spectrum of the flat top sampled signal is given as

$$G(f) = f_s \sum_{n=-\infty}^{\infty} X(f - nf_s) H(f) \quad \text{...(i)}$$

Value of $H(f)$ is expressed as

$$H(f) = \tau \operatorname{sinc}(f\tau)\, e^{-j\pi f\tau} \quad \text{...(ii)}$$

Here τ is the width of the rectangular pulse used for sampling.

The given value of rectangular sampling pulse duration is $\frac{50}{3}$ milliseconds *i.e.*,

$$\tau = \frac{50}{3} \times 10^{-3} = \frac{0.05}{3} \text{ sec}$$

Substituting the value of τ in equation (*ii*), we get

$$H(f) = \frac{0.05}{3} \operatorname{sinc}\left(\frac{0.05f}{3}\right) e^{-j0.05\pi f/3}$$

Again, putting this value of $H(f)$ and f_s in equation (*i*), we get

$$G(f) = 20 \sum_{n=-\infty}^{\infty} X(f - 20n) \frac{0.05}{3} \operatorname{sinc}\left(\frac{0.05f}{3}\right) e^{-j0.05\pi f/3} \quad (\because f_s = 20)$$

$$G(f) = \frac{1}{3} \sum_{n=-3}^{3} X(f - 20n) \operatorname{sinc}\left(\frac{0.05f}{3}\right) e^{-j0.05\pi f/s}$$

This expression gives the spectrum up to 60 Hz (since $n = \pm 3$) for the sampled signal. **Ans.**

Example 9.7. A flat top sampling system samples a signal of maximum frequency 1 Hz with 9.5 Hz sampling frequency. The duration of the pulse is 0.2 seconds. Compute the amplitude distortion due to aperture effect at the highest signal frequency. Also determine the equalization characteristic.

Solution: Given that sampling frequency

$$f_s = 2.5 \text{ Hz}$$

Maximum signal frequency,

$$f_{max} = 1 \text{ Hz}$$

and pulse width $\tau = 0.2$ seconds.

We know that the aperture effect is expressed by a transfer function $H(f)$ as

$$H(f) = \tau \operatorname{sin\,c}(f\tau)\, e^{-j\pi f\tau}$$

The magnitude of this equation will be

$$|H(f)| = \tau \operatorname{sinc}(f\tau)$$

$$|H(f)| = 0.2 \operatorname{sinc}(f \times 0.2) \quad \text{...(i)}$$

Now, aperture effect at the highest frequency will be obtained by putting $f = f_{max} = 1$ Hz in equation (*i*) *i.e.*,

$$|H(1)| = 0.2 \operatorname{sinc}(0.2) = 0.18709$$

or $|H(1)| = 18.70\%$ **Ans.**

Also, the equalizer characteristic is expressed as

$$H_{eq}(f) = \frac{K}{\tau \operatorname{sin\,c}(f\tau)}$$

Substituting, $\tau = 0.2$ second and assuming $K = 1$, the last equation becomes

$$H_{eq}(f) = \frac{1}{0.2 \operatorname{sin\,c}(0.2f)}$$

This equation is the plot of $H_{eq}(f)$ versus f and it represents the equalization characteristic to overcome aperture effect.

9.12. Analog Pulse Modulation Methods

We know that in analog modulation systems, some parameter of a sinusoidal carrier is varied according to the instantaneous value of the modulating signal. In pulse modulation methods, the carrier is no longer a continuous signal but consists of a pulse train. Some parameter of which is varied according to the instantaneous value of the modulating signal. There are two types of pulse modulation systems as under:

(*i*) Pulse Amplitude Modulation (PAM)

(*ii*) Pulse Time Modulation (PTM)

In pulse amplitude modulation (PAM), the amplitude of the pulses of the carrier pulse train is varied in accordance with the modulating signal whereas in pulse time modulation (PTM), the timing of the pulses of the carrier pulse train is varied.

They are two types of PTM as under:

(*i*) Pulse width modulation (PWM)*

(*ii*) Pulse position modulation (PPM)

In Pulse width modulation, the width of the pulses of the carrier pulse train is varied in accordance with the modulating signal whereas in Pulse position modulation (PPM), the position of pulses of the carrier pulse train is varied. Figure 9.21 shows three types of pulse analog modulation methods.

According to the sampling theorem, if a modulating signal is bandlimited to f_m Hz,** the sampling frequency must be at least $2f_m$ Hz and, hence the frequency of the carrier pulse train must also be least $2f_m$ Hz.

At this point, it may be noted that all the above pulse modulation methods (*i.e.*, PAM, PWM and PPM) are called analog Pulse modulation methods because the modulating signal is analog in nature in PAM, PWM and PPM.

* Pulse width modulation is also known as Pulse duration modulation (PDM).

** If a signal is said to be bandlimited to f_m Hz, then it means that the maximum frequency component in this signal is f_m Hz.

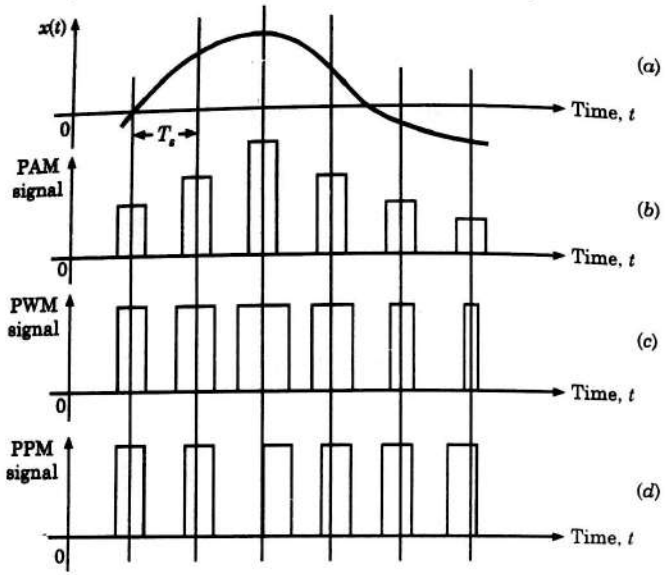


Fig. 9.21. *Different types of pulse analog modulation methods.*

9.13. Pulse Amplitude Modulation (PAM)

Pulse amplitude modulation may be defined as that type of modulation in which the amplitudes of regularly spaced rectangular pulses vary according to instantaneous value of the modulating or message signal. In fact, the pulses in a PAM signal may be of flat top type or natural type or ideal type. Actually all the sampling methods which have been discussed in last sections are basically pulse amplitude modulation methods.

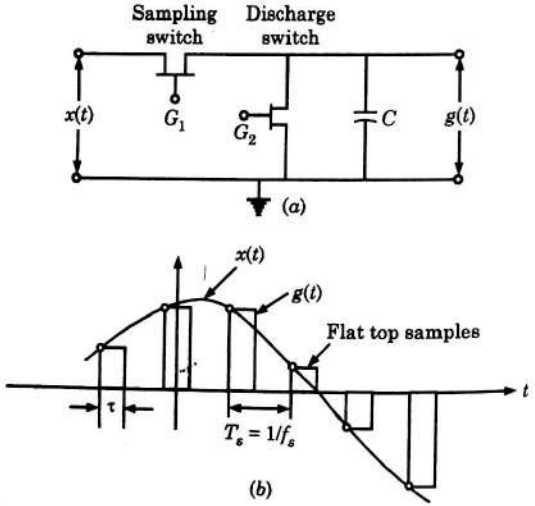


Fig. 9.22. *(a) Sample and hold circuit generating flat top sampled PAM (b) Waveforms of flat top sampled PAM*

Out of these three pulse amplitude modulation methods, the Flat top PAM is most popular and is widely used. The reason for using flat top PAM is that during the transmission, the noise interferes with the top of the transmitted pulses and this noise can be easily removed if the PAM pulse has flat top.

However, in case of natural samples PAM signal, the pulse has varying top in accordance with the signal variation. Now, when such type of pulse is received at the receiver, it is always contaminated by noise. Then it becomes quite difficult to determine the shape of the top of the pulse and thus amplitude detection of the pulse is not exact. Due to this, errors are introduced in the received signal.

Therefore, flat top sampled PAM is widely used.

Figure 9.22 shows the sample and hold circuit to produce flat top sampled PAM and the waveform for flat top sampled PAM.

Working Principle

A sample and hold circuit shown in figure 9.22 is used to produce flat top sampled PAM. The working principle of this circuit is quite easy. The sample and hold (S/H) circuit consists of two field

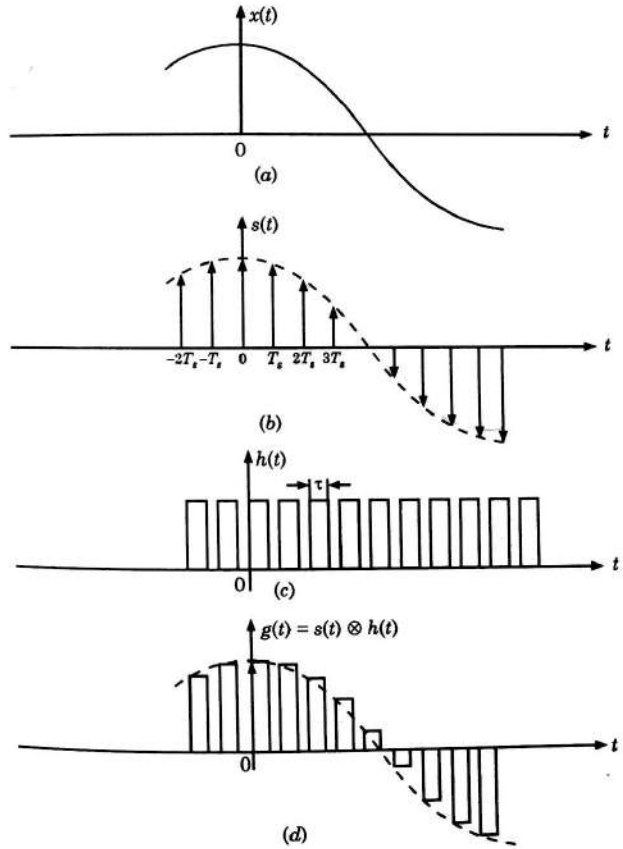


Fig. 9.23. *(a) Baseband signal x(t) (b) Instantaneously sample single s(t) (c) Constant pulse width function h(t) (d) Flat top sampled PAM signal g(t) obtained through convolution of h(t) and s(t).*

effect transistors (FET) switches and a capacitor. The sampling switch is closed for a short duration by a short pulse applied to the gate G_1 of the transistor. During this period, the capacitor 'C' is

quickly charged upto a voltage equal to the instantaneous sample value of the incoming signal $x(t)$. Now, the sampling switch is opened and the capacitor 'C' holds the charge. The discharge switch is then closed by a pulse applied to gate G_2 of the other transistor. Due to this, the capacitor 'C' is discharged to zero volts. The discharges switch is then opened and thus capacitor has no voltage.

Hence, the output of the sample and hold circuit consists of a sequence of Flat top samples as shown in figure 9.23.

Mathematical Analysis

In a Flat top PAM, the top of the samples remains constant and is equal to the instantaneous value of the baseband signal $x(t)$ at the start of sampling. The duration or width of each sample is τ and sampling rate is equal to $f_s = \frac{1}{T_s}$. From figure 9.22 (*b*), it may be noted that only starting edge of the pulse represents instantaneous value of the baseband signal $x(t)$. Also, the flat top pulse of $g(t)$ is mathematically equivalent to the convolution of instantaneous sample and a pulse $h(t)$ as depicted in figure 9.24.

This means that the width of the pulse in $g(t)$ is determined by the width of $h(t)$ and the sampling instant is determined by the delta function.

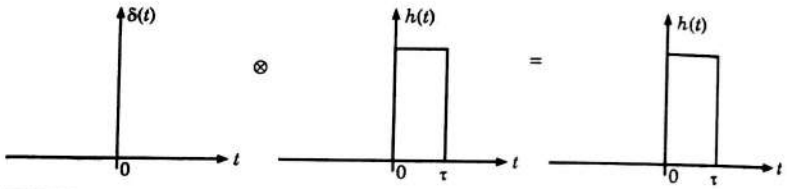


Fig. 9.24. *Convolution of any function with delta function is equal to that function.*

In figure 9.22 (*b*), the starting edge of the pulse represents the point where baseband signal is sampled and width is determined by function $h(t)$.

Therefore, $g(t)$ will be expressed as

$$g(t) = s(t) \otimes h(t) \qquad ...(9.58)$$

This equation has been explained in figure 9.23.

Now, from the property of delta function, we know that for any function $f(t)$

$$f(t) \otimes s(t) = f(t) \qquad ...(9.59)$$

This property is used to obtain flat top samples. It may be noted that to obtain flat top sampling we are not applying the equation (9.59) directly here *i.e.*, we are applying a modified from of equation (9.59). This modified equation is equation (9.58).

Thus, in this modified equation, we are taking $s(t)$ in place of delta functions $\delta(t)$. Observe that $\delta(t)$ is a constant amplitude delta function whereas $s(t)$ is a varying amplitude train of impulses. This means that we are taking $s(t)$ which is an instantaneously sampled signal and this is convolved with function $h(t)$ as in equation (9.58).

Therefore, on convolution of $s(t)$ and $h(t)$, we get a pulse whose duration is equal to $h(t)$ only but amplitude is defined by $s(t)$.

Now, we know that the train of impulses may be represented mathematically as

$$\delta_{T_s}(t) = \sum_{n=-\infty}^{\infty} \delta(t - nT_s) \qquad ...(9.60)$$

The signal $s(t)$ is obtained by multiplication of baseband signal $x(t)$ and $\delta_{T_s}(t)$.

Thus,

$$s(t) = x(t) \cdot \delta_{T_s}(t) \qquad ...(9.61)$$

or

$$s(t) = x(t) \sum_{n=-\infty}^{\infty} \delta(t - nT_s) = \sum_{n=-\infty}^{\infty} x(nT_s) \cdot \delta(t - nT_s) \qquad ...(9.62)$$

Now, sampled signal $g(t)$ is given as equation (9.58)

$$g(t) = s(t) \otimes h(t) \qquad ...(9.63)$$

or

$$g(t) = \int_{-\infty}^{\infty} s(\tau)\, h(t - \tau)\, d\tau \qquad ...(9.64)$$

or

$$g(t) = \int_{-\infty}^{\infty} \sum_{n=-\infty}^{\infty} x(nT_s)\, \delta(\tau - nT_s)\, h(t - \tau)\, d\tau$$

or

$$g(t) = \sum_{n=-\infty}^{\infty} x(nT_s) \int_{-\infty}^{\infty} \delta(\tau - nT_s)\, h(t - \tau)\, d\tau \qquad ...(9.65)$$

According to shifting property of delta function we know that

$$\int_{-\infty}^{\infty} f(t)\, \delta(t - t_0) = f(t_0) \qquad ...(9.66)$$

Using equations (9.65) and (9.66), we get

$$g(t) = \sum_{n=-\infty}^{\infty} x(nT_s)\, h(t - nT_s)$$

This equation represents value of $g(t)$ in terms of sampled value $x(nT_s)$ and function $h(t - nT_s)$ for flat top sampled signal.

Now, again from equation (9.58), we have

$$g(t) = s(t) \otimes h(t)$$

Taking Fourier transform of both sides of above equation, we get

$$G(f) = S(f)\, H(f) \qquad ...(9.67)$$

We know that $S(f)$ is given as

$$S(f) = f_s \sum_{n=-\infty}^{\infty} X(f - nf_s) \qquad ...(9.68)$$

Therefore, equation (9.67) will become

$$G(f) = f_s \sum_{n=-\infty}^{\infty} X(f - nf_s)\, H(f) \qquad ...(9.69)$$

Thus, spectrum of flat top PAM signal:

$$G(f) = f_s \sum_{n=-\infty}^{\infty} X(f - nf_s)\, H(f) \qquad ...(9.70)$$

Here, $H(f)$ is the Fourier transform of the rectangular pulse. The spectrum of this rectangular pulse is shown in figure 9.18(*b*). Let the spectrum of $s(t)$ be the rectangular pulse train as shown in figure 9.25(*a*) and the spectrum of $h(t)$ *i.e.*, $H(f)$ is shown in figure 9.25(*b*).

By equation (9.67), we know that

$$G(f) = S(f) \cdot H(f)$$

Thus, according to above equation, we can plot the spectrum $G(f)$ as shown in figure 9.25(b).

Note: *It may be observed in figure 9.25(b) that higher frequencies in $S(f)$ are attenuated due to roll-off characteristics of the 'sinc' pulse. This effect is popularly known as* ***aperture effect.***

An equalizer is needed to overcame this effect.

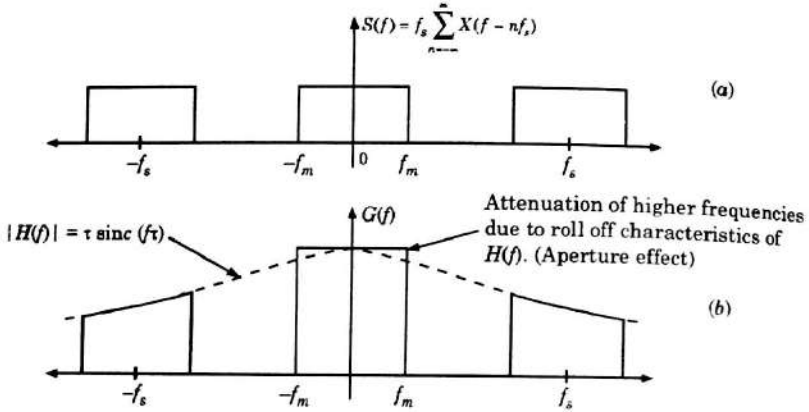


Fig. 9.25. *(a) Spectrum of some arbitrary signal. The signal is sampled at f_s and maximum frequency in the signal is f_m.*
(b) Spectrum of flat top signal. The dotted curve is $H(f) = \tau \sin c\,(f\tau)$

9.13.1. Naturally Sampled Pulse Amplitude Modulated (PAM) Signal

We have discussed natural sampling in article 9.9. This natural sampling is basically pulse amplitude modulation (PAM). Therefore, it is called naturally sampled PAM signal.

Thus, time-domain representation of a naturally-sampled PAM signal will be given as

$$g(f) = \frac{\tau A}{T_s} \sum_{n=-\infty}^{\infty} x(t), \sin c(f_n \tau)\, e^{j2\pi n f_s t} \qquad ...(9.71)$$

and the frequency-domain representation, *i.e.* frequency-spectrum of a naturally-sampled PAM signal will be given as

$$G(f) = \frac{\tau A}{T_s} \sum_{n=-\infty}^{\infty} \sin c\,(nf_s\tau) \cdot X(f - nf_s) \qquad ...(9.72)$$

9.13.2. Instantaneous or Idealy Sampled Pulse Amplitude Modulated (PAM) Signal

We have discussed ideal or instantaneous sampling in article 9.9. This instantaneous sampling is basically pulse amplitude modulation (PAM). Therefore, it is called ideally or instantaneously sampled PAM signal.

Thus, time-domain representation of a ideally or instantaneously sampled PAM signal will be given as

$$g(t) = \sum_{n=-\infty}^{\infty} x(nT_s) \cdot \delta(t - nT_s) \qquad ...(9.73)$$

and the frequency-domain representation *i.e.*, frequency-spectrum of a ideally or instantaneously sampled PAM sinal will be given as

$$G(f) = f_s \sum_{n=-\infty}^{\infty} X(f - nf_s) \qquad ...(9.74)$$

9.13.3. Transmission Bandwidth in Pulse Amplitude Modulation (PAM)

In a pulse amplitude modulated (PAM) signal the pulse duration 'τ' is considered to be vary very small in comparison to time period (*i.e.*, sampling period) T_s between any two samples *i.e.*,

$$\tau << T_s \qquad ...(9.75)$$

Now, if the maximum frequency in the modulating signal $x(t)$ is f_m, then according to sampling theorem, the sampling frequency f_s must be equal to or higher than the Nyquist rate, *i.e.*

or $$f_s \geq 2f_m \qquad ...(9.76)$$

or $$\frac{1}{T_s} \geq 2 f_m \qquad (\because f_s = \frac{1}{T_s})$$

or $$T_s \leq \frac{1}{2f_m}$$

But according to equation (9.75), we have

$$\tau << T_s$$

therefore $$\tau << T_s \leq \frac{1}{2f_m} \qquad ...(9.77)$$

Now, if the 'ON' and 'OFF' time of the pulse amplitude modulated (PAM) pulse is same as shown in figure 9.26(*a*) then maximum frequency of the PAM pulse will be

$$f_{max} = \frac{1}{\tau + \tau} = \frac{1}{2\tau} \qquad ...(9.78)$$

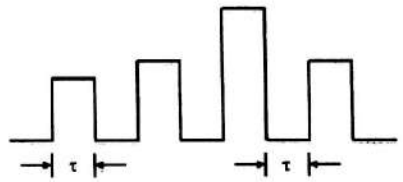


Fig. 9.26. *(a) Illustration of maximum frequency in PAM signal.*

Therefore, the bandwidth required for the transmission of a PAM signal would be equal to the maximum frequency f_{max} given by the equation (9.78)

Thus, we have

Transmission bandwidth

$$BW \geq f_{max} \qquad ...(9.79)$$

But $$f_{max} = \frac{1}{2\tau}$$

Hence $$BW \geq \frac{1}{2\tau}$$

Again, since $$\tau << \frac{1}{2f_m}$$

Therefore $$BW \geq \frac{1}{2\tau} >> f_m \qquad ...(9.80)$$

or $$BW >> f_m \qquad ...(9.81)$$

www.dahiyastech.blogspot.com

9.13.4. Demodulation of PAM Signals

As discussed earlier, demodulation is the reverse process of modulation in which the modulating signal is recovered back from a modulated signal. For pulse-amplitude modulated (PAM) signals, the demodulation is done using a holding circuit. Figure 9.26(*b*) shows the block diagram of a PAM demodulator.

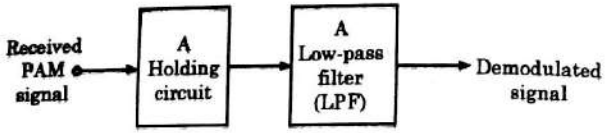


Fig. 9.26. *(b) A block diagram of PAM demodulator*

In this method, the received PAM signal is allowed to pass through a holding circuit and a low pass filter (LPF) as shown in above figure. Now, figure 9.27(*a*) illustrates a very simple holding circuit. Here the switch '*S*' is closed after the arrival of the pulse and it is opened at the end of the pulse. In this way, the capacitor *C* is charged to the pulse amplitude value and it holds this value during the interval between the two pulses. Hence, the sampled values are held as shown in figure 9.27(*c*). After this the holding circuit output is smothened in Low Pass filter as shown in figure 9.27(*c*). It may be observed that some kind of distortion is introduced due to the holding circuit. In fact the circuit of figure 9.27(*b*) is known as **zero-order holding circuit.** This zero-order holding circuit considers only the previous sample to decide the value between the two pulses.

Note: *It may be noted the first order hold circuit considers the previous two samples whereas a second order holding circuit considers the previous three samples and so on. However, as the order of the holding circuit increases, the distortion decreases at the cost of the circuit complexity. In fact, the amount of permissible distortion decides the order of the holding circuit.*

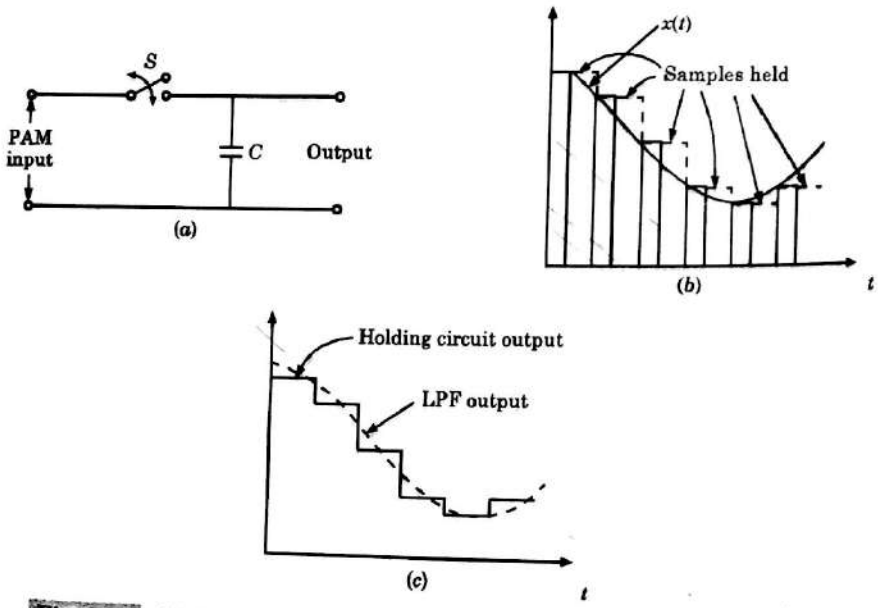


Fig. 9.27. *(a) A zero-order holding circuit (b) The output of holding circuit (c) The output of a low pass filter (LPF)*

9.13.5. Transmission of PAM Signals

If the PAM signals are to be transmitted directly *i.e.*, over a pair of wires then no further signal processing is necessary. However, if they are to be transmitted through the space using an antenna, they must first be amplitude or frequency or phase modulated by a high frequency carrier and only then they can be transmitted. Thus, the overall system will be then known as PAM-AM or PAM-FM or PAM-PM respectively. At the receiving end, AM or FM or PM detection is first employed to get the PAM signal and then the message signal is recovered from it.

Example 9.8. For a pulse-amplitude modulated (PAM) transmission of voice signal having maximum frequency equal to f_m = 3 kHz, calculate the transmission bandwidth. It is given that the sampling frequency f_s = 8 kHz and the pulse duration $\tau = 0.1T_s$.

Solution: We know that the sampling period T_s is expressed as

$$T_s = \frac{1}{f_s} = \frac{1}{8 \times 10^3} \text{ seconds}$$

$$T_s = 0.125 \times 10^{-3} \text{ seconds} = 125\ \mu \text{ seconds} \qquad ...(i)$$

Also, τ is given that

$$\tau = 0.1\ T_s$$

Using (*i*), we get

$$\tau = 0.1 \times 125 = 12.5\ \mu \text{ seconds} \qquad ...(ii)$$

Now, we know that the transmission bandwidth for PAM signal is expressed as

$$BW \geq \frac{1}{2\tau}$$

Using equation (*ii*), we get

$$BW \geq \frac{1}{2 \times 12.5 \times 10^{-6}} \geq \frac{1 \times 10^6}{25}$$

$$BW \geq 40 \text{ kHz} \qquad \textbf{Ans.}$$

9.13.6. Drawbacks of Pulse-Amplitude Modulated (PAM) Signal

Following are the drawbacks of a PAM signal:

(*i*) The bandwidth required for the transmission of a PAM signal is very very large in comparison to the maximum frequency present in the modulating signal.

(*ii*) Since the amplitude of the PAM pulses varies in accordance with the modulating signal therefore the interference of noise is maximum in a PAM signal. This noise cannot be removed easily.

(*iii*) Since the amplitude of the PAM signal varies, therefore, this also varies the peak power required by the transmitter with modulating signal.

9.14. Pulse Time Modulation

In pulse time modulation, the signal to be transmitted is sampled as in pulse amplitude modulation (PAM). In pulse time modulation, amplitude of pulse is held constant, whereas position of pulse or width of pulse is made proportional to the amplitude of signal at the sampling instant. There are two types of pulse time modulation, viz. Pulse Width Modulation (PWM) and Pulse Position Modulation (PPM). Because in both PWM and PPM, amplitude is held constant and does not carry any information, therefore amplitude limiters can be used. The amplitude limiters, similar to those used in FM, will clip off the portion of the signal corrupted by noise and hence provide a good degree of noise immunity.

9.14.1. Pulse Width Modulation

Let us first discuss Pulse Width Modulation (PWM). This is also known as Pulse Duration Modulation (PDM). Three variations of pulse width modulation are possible. In one variation, the leading edge of the pulse is held constant and change in pulse width with signal is measured with respect to the leading edge. In other variation, the tail edge is held constant and with respect to it, pulse width is measured. In the third variation, centre of the pulse is held constant and pulse width changes on either side of the centre of the pulse. This has been illustrated in figure 9.28.

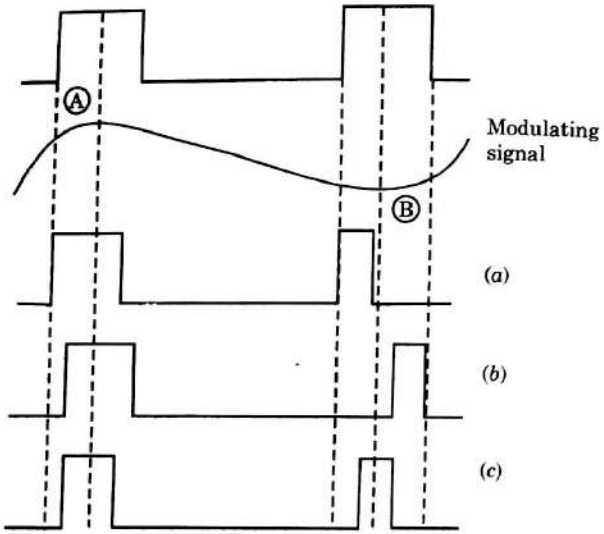


Fig. 9.28. *PWM waveforms*

The modulating signal is at its positive peak at point (A) and at its negative peak at (B). In figure 9.28 (*a*), the leading edge of pulse is kept constant and pulse width is measured from the lead edge. As shown, pulse width is maximum corresponding to point (A), while it is minimum at point (B).

In figure 9.28 (*b*), the tail edge of the pulse is kept constant and pulse width is measured from the tail end of the pulse. As before, pulse width is maximum corresponding to positive peak of the modulating signal and minimum at the negative peak.

As shown in figure 9.28 (*c*), the center of the pulse is kept constant and pulse extends on either side of the center of the pulse, depending upon the modulating signal.

9.14.2. Frequency Spectrum for PWM Wave

With a sinusoidal modulating signal at frequency f_m, the spectrum of PWM signal consists the modulating signal frequency f_m along with several harmonics. This is shown in figure 9.29.

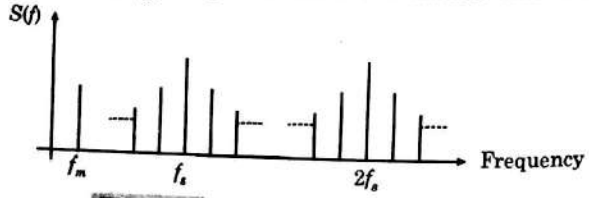
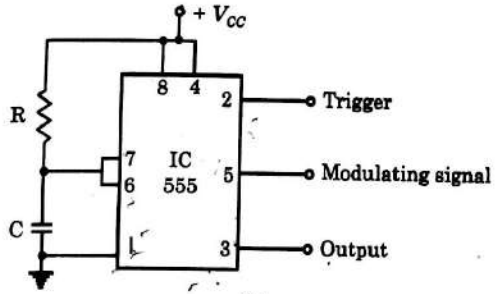


Fig. 9.29. *Spectrum of PWM signal*

To have a better separation with respect to frequency, between highest frequency of baseband signal [in Fig. 9.29, f_m] and lower sidebands of f_s (sampling frequency), a higher sampling frequency which is more than Nyquist rate is used; and pulse width deviation is kept small.

9.14.3. Modulation of PWM Signal or PWM Generation

Figure 9.30 shows pulse width modulator. It is basically a monostable multivibrator with a modulating input signal applied at the control voltage input. Internally, the control voltage is adjusted to the 2/3 V_{CC}. Externally applied modulating signal changes the control voltage, and hence the threshold voltage level. As a result, the time period required to charge the capacitor up to threshold voltage level changes, giving pulse modulated signal at the output, as shown in the figure 9.30 (*b*).



(*a*)

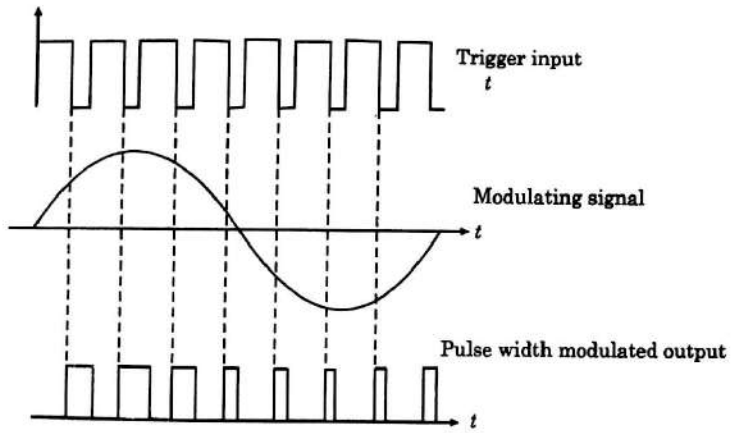


Fig. 9.30. (*b*)

Figure 9.31 illustrates another monostable multivibrator circuit to generate pulse width modulation (PWM).

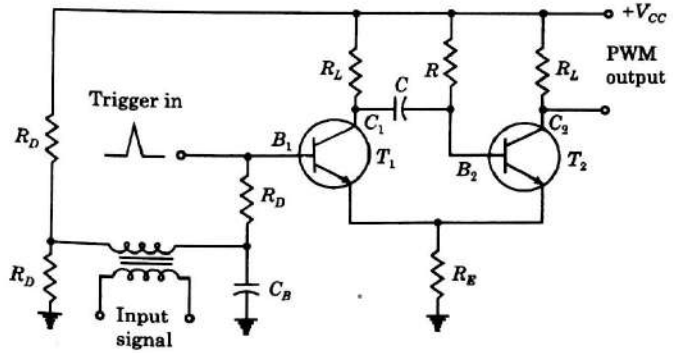


Fig. 9.31. *Monostable multivibrator generating pulse width modulation (PWM).*

The stable state for above circuit is achieved when T_1 is OFF and T_2 is ON. The positive going trigger pulse at B_1 switches T_1 ON. Because of this, the voltage at C_1 falls as T_1 now begins to draw the collector current. As a result, voltage at B_2 also falls and T_2 is switched OFF, C begins to charge up to the collector supply voltage (V_{CC}) through resistor R. After a time determined by the supply voltage and the RC time constant of the charging network, the base of the T_2 becomes sufficiently positive to switch T_2 ON. The transistor T_1 is simultaneously switched OFF by regenerative action and stays OFF until the arrival of the next trigger pulse. To make T_2 ON, the base of the T_2 must be slightly more positive than the voltage across resistor R_E. This voltage depends on the emitter curent I_E which is controlled by the signal voltage applied at the base of transistor T_1. Therefore, the changing voltage necessary to turn OFF transistor T_2 is decided by the signal voltage. If signal voltage is maximum, the voltage that capacitor should charge to turn ON T_2 is also maximum. Therefore, at maximum signal voltage, capacitor has to charge to maximum voltage requiring maximum time to charge. This gives us maximum pulse width at maximum input signal voltage. At minimum signal voltage, capacitor has to charge for minimum voltage and we get minimum pulse width at the output. With this discussion, it can be noted that pulse width is controlled by the input signal voltage, and we get pulse width modulated waveform at the output.

9.14.4. Demodulation of PWM Signal

Figure 9.32 (*a*) shows the block diagram of PWM detector. As shown in the figure 9.32 (*a*), the received PWM signal is applied to the Schmitt trigger circuit. This Schmitt trigger circuit removes the noise in the PWM waveform. The regenerated PWM is then applied to the ramp generator and the synchronization pulse detector. The ramp generator produces ramps for the duration of pulses such that height of ramps are proportional to the widths of PWM pulses. The maximum ramp voltage is retained till the next pulse. On the other hand, synchronous pulse detector produces reference pulses with constant amplitude and pulse width. These pulses are delayed by specific amount of delay as shown in the figure 9.32 (*b*). The delayed reference pulses and the output of ramp generator is added with the help of adder. The output of adder is given to the level shifter. Here, negative offset shifts the waveform as shown in the figure 9.32 (*b*) ⑥. Then the negative part of the waveform is clipped by rectifier. Finally, the output of rectifier is passed through low-pass filter to recover the modulating signal, as shown in the figure 9.32 (*b*).

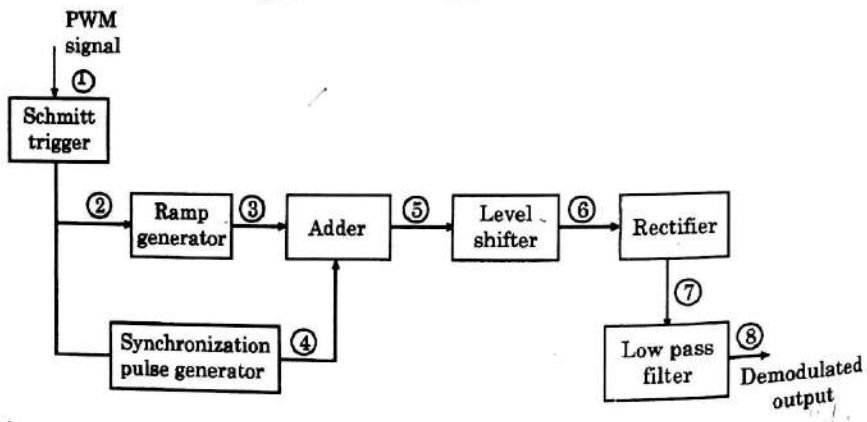


Fig. 9.32. *(a) PWM detector*

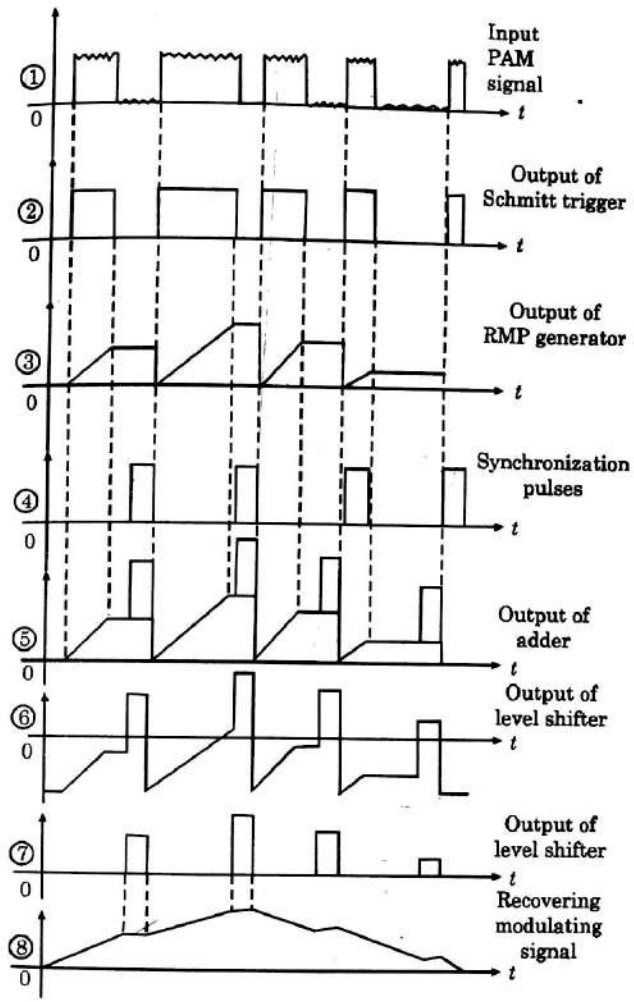


Fig. 9.32. *(b) Waveforms for PWM detection circuit.*

9.14.5 Advantages of PWM

(*i*) Unlike, PAM, noise is less, since in PWM, amplitude is held constant.

(*ii*) Signal and noise separation is very easy, as shown in figure 9.32 (*b*) ②.

(*iii*) PWM communication does not require synchronization between transmitter and receiver.

9.14.6. Disadvantages of PWM

(*i*) In PWM, pulses are varying in width and therefore their power contents are variable. This requires that the transmitter must be able to handle the power contents of the pulse having maximum pulse width.

(*ii*) Large bandwidth is required for the PWM communication as compared to PAM.

9.14.7. Pulse Position Modulation

In this system, the amplitude and width of the pulses are kept constant, while the position of each pulse, with reference to the position of a reference pulse, is changed according to the

instantaneous sampled value of the modulating signal. Thus, the transmitter has to send synchronizing pulses to keep the transmitter and receiver in synchronism. As the amplitude and width of the pulses are constant, the transmitter handles constant power output, a definite advantage over the PWM. But the disadvantage of the PPM system is the need for transmitter-receiver synchronization. Pulse position modulation is obtained from pulse width modulation, shown in the figure 9.33. Each trailing edge of the PWM pulse is a starting point of the pulse in the PPM. Therefore, position of the pulse is 1:1 proportional to the width of pulse in PWM and hence it is proportional to the instantaneous amplitude of the sampled modulating signal.

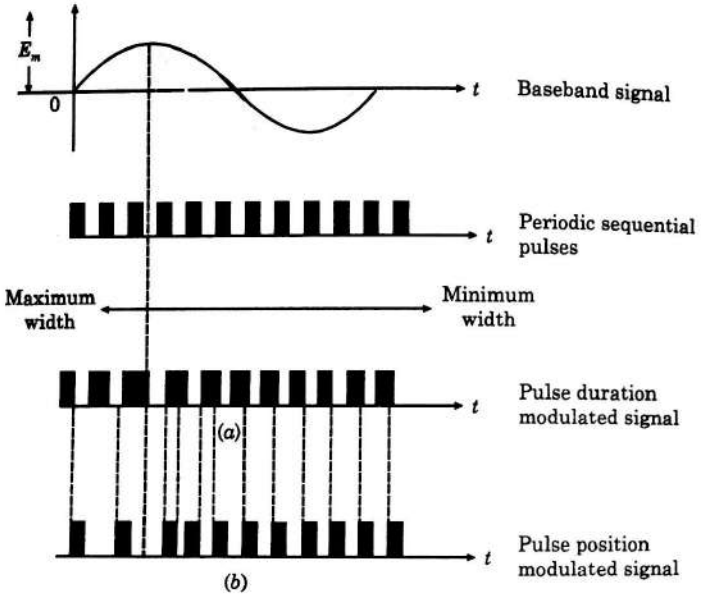


Fig. 9.33. *PPM generator*

9.14.8. Generation of PPM Signal

Figure 9.34 (*a*) shows the PPM generator. It consists of differentiator and a monostable multivibrator. The input to the differentiator is a PWM waveform. The differentiator generates positive and negative

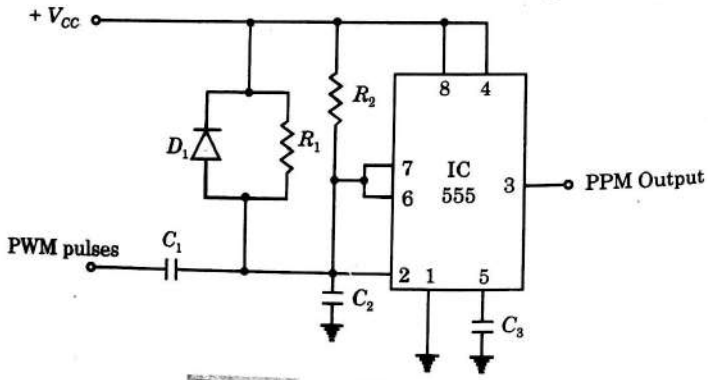


Fig. 9.34. *(a) PPM generator*

spikes corresponding to leading and trailing edges of the PWM waveform. Diode D1 is used to bypass the positive spikes. The negative spikes are used to the trigger monostable multivibrator. The monostable multivibrator then generates the pulses of same width and amplitude with reference to trigger to give pulse position modulated waveform, as shown in figure 9.34 (*b*).

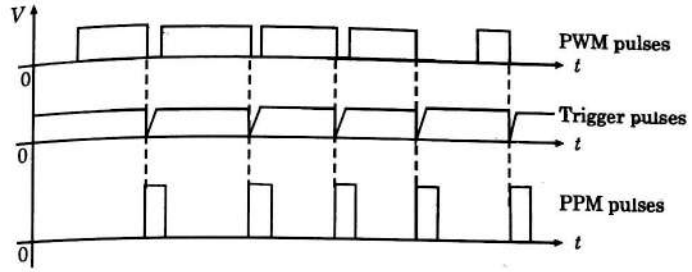


Fig. 9.34. *(b) Waveforms of PPM generator*

9.14.9. Demodulation of PPM

In case of pulse-position modulation, it is customary to convert the received pulses that vary in position to pulses that vary in length. One way to achieve this is illustrated in figure 9.35.

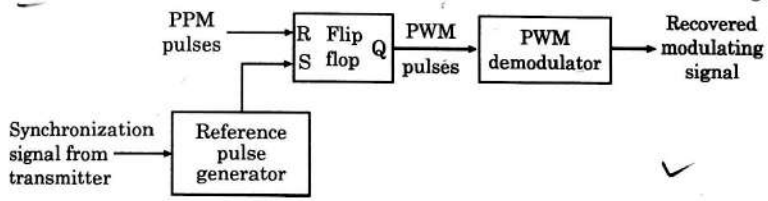


Fig. 9.35. *PPM demodulator*

As shown in figure 9.35, flip-flop circuit is set or turned 'ON' (giving high output) when the reference pulse arrives. This reference pulse is generated by reference pulse generator of the receiver with the synchronization signal from the transmitter. The flip-flop circuit is reset or turned 'OFF' (giving low output) at the leading edge of the position modulated pulse. This repeats and we get PWM pulses at the output of the flip-flop.

The PWM pulses are then demodulated by PWM demodulator to get orginal modulating signal.

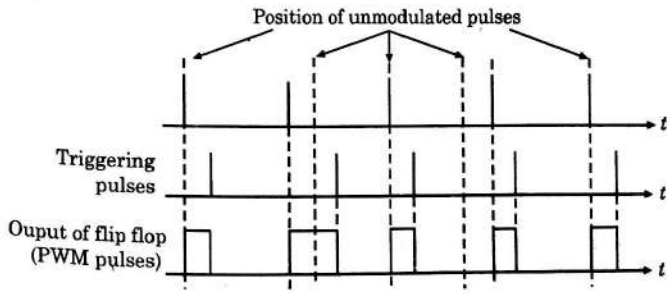


Fig. 9.36. *Demodulation waveform for PPM*

9.14.10. Advantages of PPM

(i) Like PWM, in PPM, amplitude is held constant thus less noise interference.

(ii) Like PPM, signal and noise separation is very easy.

(*iii*) Because of constant pulse widths and amplitudes, transmission power for each pulse is same.

9.14.11. Disadvantages of PPM

(*i*) Synchronization between transmitter and receiver is required.

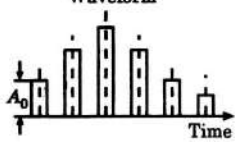
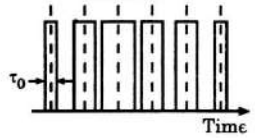
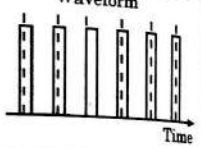
(*ii*) Large bandwidth is required as compared to PAM.

9.15. Comparison of Various Pulse Analog Modulation Methods

(U.P. Tech, Semester Examination, 2003-04)

In this section, let us compare PAM, PWM and PPM in the form of a Table 9.2.

Table 9.2. Comparison of PAM, PPM and PDM

S. No.	*Pulse Amplitude Modulation (PAM)*	*Pulse Width/Duration Modulation (PWM) or (PDM)*	*Pulse Position Modulation (PPM)*
1.	Waveform 	Waveform 	Waveform 
2.	Amplitude of the pulse is proportional to amplitude of modulating signal.	Width of the pulse is proportional to amplitude of modulating signal.	The relative position of the pulse is proprotional to the amplitude of modulating signal.
3.	The bandwidth of the transmission channel depends on width of the pulse.	Bandwidth of transmission channel depends on rise time of the pulse.	Bandwidth of transmission channel depends on rising time of the pulse.
4.	The instantaneous power of the transmitter varies.	The instantaneous power of the transmitter varies.	The instantaneous power of the transmitter remains constant.
5.	Noise interference is high System is complex.	Noise, interference is minimum.	Noise, interference is minimum.
6.	Similar to amplitude modulation.	Simple to implement similar to frequency modulation.	Simple to implement similar to phase modulation.

SUMMARY

1. There are two types of signals, continuous time signal and discrete-time signals.
2. Due to some recent advance development in digital technology over the past few decades, the inexpensive, light weight, programmable and easily reproducible discrete-time systems are available. Therefore, the processing of discrete-time signals is more flexible and is also preferable to processing of continuous-time signals.
3. The sampling theorem is extremely important and useful in signal processing.
4. With the help of sampling theorem, a continuous-time signal may be completely represented and recovered from the knowledge of samples taken uniformly.
5. The concept of sampling provides a widely used method for using discrete-time system technology to implement continuous-time systems and process the continuous-time signals.

6. Sampling of the signals is the fundamental operation in signal-processing. A continuous time signal is first converted to discrete-time signal by sampling process.
7. The sufficient number of samples of the signal must be taken so that the original signal is represented in its samples completely. Also, it should be possible to recover or reconstruct the original signal completely from its samples. The number of samples to be taken depends on maximum signal frequency present in the signal.
8. A continuous-time signal may be completely represented in its samples and recovered back if the sampling frequency is $f_s \geq 2 f_m$. Here, f_s is the sampling frequency and f_m is the maximum frequency present in the 'signal.
9. When the sampling rate becomes exactly equal to $2 f_m$ samples per second, then it is called Nyquist rate. Nyquist rate is also called the *minimum* sampling rate. It is given by
$$f_s = 2 f_m$$
10. Maximum sampling interval is called *Nyquist interval*. It is given by

Nyquist Interval $\quad T_s = \dfrac{1}{2 f_m}$ seconds.

11. The low pass filter is used to recover original signal from its samples. This is also known as interpolation filter.
12. A low-pass filter is that type of filter which passes only low-frequencies upto a specified cut-off frequency and rejects all other frequenices above cut-off frequency.
13. The process of reconstructing a continuous-time signal $x(t)$ from its samples is called as interpolation.
14. Aliasing is the phenomenon in which a high frequency component in the frequency-spectrum of the signal takes identity of a lower-frequency component in the spectrum of the sampled signal.
15. Because of the overlap due to aliasing phenomenon, it is not possible to recover original signal $x(t)$ from sampled signal $g(t)$ by low-pass filtering since the spectral components in the overlap regions add and hence the signal is distorted.
16. Since any information signal contains a large number of frequencies, so, to decide a sampling frequency is always a problem. Therefore, a signal is first passed through a low-pass filter. This low-pass filter blocks all the frequencies which are above f_mHz. This process is known as band limiting of the original signal $x(t)$. This low-pass filter is called prelias filter because it is used to prevent aliasing effect.
17. To avoid aliasing, we must have:
    (*i*) Preaias filter must be used to limit band of frequencies of the signal to f_m Hz.
    (*ii*) Sampling frequency 'f_s' must be selected such that
$$f_s > 2 f_m$$
18. The bandpass signal $x(t)$ whose maximum bandwidth is $2f_m$ can be completely represented into and recovered from its samples if it is sampled at the minimum rate of twice the bandwidth. Here, f_m is the maximum frequency component present in the signal.
19. Basically, there are three types of sampling techniques as under:
    (*i*) Instantaneous sampling
    (*ii*) Natural sampling
    (*iii*) Flat top sampling.
20. Natural sampling is a practical method and will be discussed in this section.
21. Spectrum of Naturally Sampled signal:
$$G(f) = \frac{\tau A}{T_s} \cdot \sum_{n=-\infty}^{\infty} \text{sinc}\,(n f_s \cdot \tau)\, X(f - n f_s)$$
22. Flot top sampling like natural sampling is also a practically possible sampling method. But natural sampling is little complex whereas it is quite easy to get flat top samples.

23. Spectrum of flat top sampled signal:

$$G(f) = f_s \sum_{n=-\infty}^{\infty} X(f - nf_s)\, H(f)$$

24. In analog modulation systems, some parameter of a sinusoidal carrier is varied according to the instantaneous value of the modulating signal. In pulse modulation methods, the carrier is no longer a continuous signal but consists of a pulse train. Some parameter of which is varied according to the instantaneous value of the modulating signal.
25. There are two types of pulse modulation systems as under:
    (*i*) Pulse Amplitude Modulation (PAM)
    (*ii*) Pulse Time Modulatin (PTM)
26. In pulse amplitude modulation (PAM), the amplitude of the pulses of the carrier pulse train is varied in accordance with the modulating signal whereas in Pulse time modulation (PTM), the timing of the pulses of the carrier pulse train is varied.
27. They are two types of PTM as under:
    (*i*) Pulse width modulation (PWM)
    (*ii*) Pulse position modulation (PPM).
28. In Pulse width modulation, the width of the pulses of the carrier pulse train is varied in accordance with the modulating signal whereas in Pulse position modulation (PPM), the position of pulses of the carrier pulse train is varied.
29. It may be noted that all the above pulse modulation methods (*i.e.*, PAM, PWM and PPM) are called analog Pulse modulation methods because the modulating signal is analog in nature in PAM, PWM and PPM.
30. Pulse amplitude modulation may be defined as that type of modulation in which the amplitudes of regularly spaced rectangular pulses vary according to instantaneous value of the modulating or message signal. In fact, the pulses in a PAM signal may be of flat top type or natural type or ideal type. Actually all the sampling methods which have been discussed in last sections are basically pulse amplitude modulation methods.
31. BW of PAM signal is given by

$$BW \geq f_{max}$$

But $$f_{max} = \frac{1}{2\tau}$$

Hence $$BW \geq \frac{1}{2\tau}$$

Again, since $$\tau << \frac{1}{2f_m}$$

Therefore $$BW \geq \frac{1}{2\tau} >> f_m.$$

32. If the PAM signals are to be transmitted directly *i.e.*, over a pair of wires then no further signal processing is necessary. However, if they are to be transmitted through the space using an antenna, they must first be amplitude or frequency or phase modulated by a high frequency carrier and only then they can be transmitted.
33. The bandwidth required for the transmission of a PAM signal is very very large in comparison to the maximum frequency present in the modulating signal.
34. Since the amplitude of the PAM pulses varies in accordance with the modulating signal therefore the interference of noise is maximum in a PAM signal. This noise cannot be removed easily.
35. Since the amplitude of the PAM signal varies, therefore, this also varies the peak power required by the transmitter with modulating signal.

SHORT QUESTIONS WITH ANSWERS

Q.1. What is sampling?

Ans. The process of converting an analog signal into a discrete signal (or) making an analog (or) continuous signal to occur at a particular interval of time is known as sampling.

Q.2. What do you mean by sampling period (T_s) and sampling rate (f_s)?

Ans. During sampling, the time elapsed between sample to sample (or) time taken by the next sample to occur is known as sampling period. It is denoted as T_s.

Sampling rate is reciprocal of sampling period. It is denoted as $f_s = 1/T_s$.

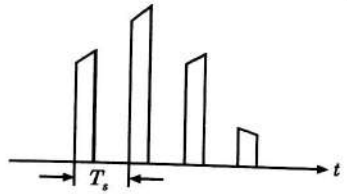


Fig. 9.37.

Q.3. What is Guard band?

Ans. When the sampling rate is chosen much higher than the Nyquist rate then a small space occurs between the samples. This space is said to be **Guard band**. This is the desired one for sampling the signals.

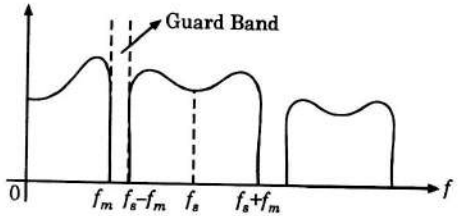


Fig. 9.38.

Q.4. What is Sampling Theorem?

Ans. Let $m(t)$ be a signal which is bandlimited such that its highest frequency spectral component is f_M. Let the values of $m(t)$ be determined at regular intervals separated by times $T_s \leq 1/2f_M$, *i.e.*, the signal is periodically sampled every T_s seconds. Then, these samples $m(nT_s)$ where n is an integer, uniquely determine the signal, and the signal may be reconstructed from these samples with no distortion at the receiver end.

Q.5. What do you mean by aperture effect?

Ans. During flat top sampling, to convert varying amplitudes of pulses to flat top pulses we use a sinc function. Because of this, there would be decrease in the amplitude. This distortion is named as Aperture effect.

This may be eliminated by using an equalizer in cascade with the output low pass filter.

Q.6. What is Natural Sampling?

Ans. It is similar to that of Instantaneous Sampling.

Let us consider a baseband signal $m(t)$ which is to be sampled. To sample $m(t)$, a sampling function $s(t)$ consisting of a train of pulses having duration T and separated by the sampling time T_s is given to one of the input of a product modulator and $m(t)$ to other input. Hence, the output consists of pulses of duration T but with a varying amplitude. This is natural sampling.

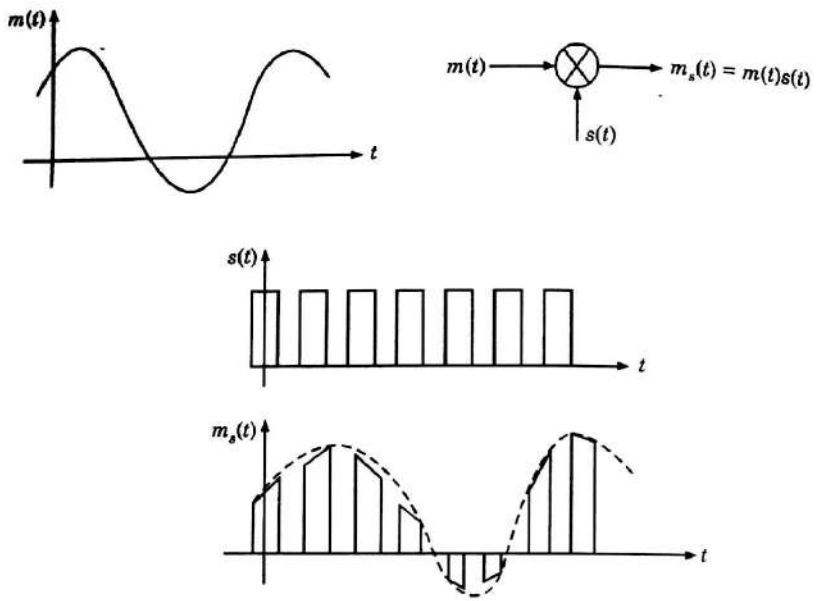


Fig. 9.39.

Q.7. What is Pulse Amplitude Modulation (PAM)?

Ans. The process in which amplitudes of regularly spaced rectangular pulses vary with the instantaneous sample values of a continuous message signal in a one to one fashion is known as pulse amplitude modulation.

Q.8. What do you mean by synchronization in PAM systems?

Ans. In general, most pulse systems require synchronization of the receiver to the transmitter. There are many ways to provide synchronization to the systems.

One of the method is a start-stop method of synchronization. This method involves transmitting some information, in addition to the message-bearing pulses, to serve as a time mark within each frame interval so that the gates in the receiver are made to open and close at the appropriate instants of time.

In some cases, the necessary time mark is established by transmitting a distinctive marker pulse per frame, whereas in some other cases, it is established by omitting a pulse in a particular time slot. One thing to be noted is that when markers are used, they should differ from the message-bearing pulses in some recognizable fashion.

In a PAM system, the marker pulse may be identified by making its amplitude exceed that of all possible message pulses.

Let us consider the transmission of PAM pluses with marker pulses. This has been shown in figure 9.40.

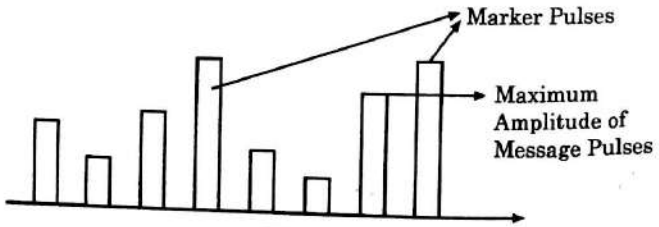


Fig. 9.40.

The PAM signal which is transmitted along with the marker pulses are detected at the receiver by giving these received pulses to a slicer.

The slicer is considered with a slicing level that is just in excess of the maximum amplitude of the message pulses, so that these pulses produces zero output.

An ideal slicer has the property that its output is zero whenever its input is below its slicing level and is constant whenever the input exceeds this level.

Q.9. What do you mean by Pulse-Time Modulation?

Ans. The modulation technique in which the time (or) duration of the pulses is varied in accordance with the amplitude of the message signal keeping the amplitude of the pluses constant is referred to as pulse-time modulation.

Q.10. What are the different types of PTM systems?

Ans. There are two kinds of Pulse-time modulation schemes. They are:

(i) Pulse duration (or) pulse width (or) pulse length modulation (PDM (or) PWM (or) PLM).

(ii) Pulse position modulation.

Q.11. What is pulse duration modulation (PDM)?

Ans. The method in which the samples of the message signal are used to vary the duration (or) width of the individual pluses. This is referred to as pulse duration modulation.

Q.12. What is pulse position modulation (PPM)?

Ans. In PPM, the position of a pulse relative to its unmodulated time of occurrence is varied in accordance with the message signal.

Q.13. Describe the spectral representation of PDM and PPM waves.

Ans. The spectral analysis of a PDM and PPM wave is complicated.

Let us consider the qualitative description of the spectra of PDM and PPM waves.

Let L_s denotes the time separation between the leading edges of duration modulated pluses obtained by natural sampling.

f_m is sinusoidal modulatin wave of frequency.

Thus the spectrum of a naturally sampled PDM wave consists of the following components:

- Sinusoidal components of frequencies equal to the integer multiplies of $1/T_S$, corresponding to spectral lines at $\pm n/T_S$, where $n = 1, 2, 3,...$ These sinusoidal components as well as the DC components are contributed by the unmodulated pulse train which is regarded as the carrier of the PDM wave.
- A sinusoidal component of frequency f_m and in phase with modulating wave, corresponding to spectral lines at $\pm f_m$.
- Sinusoidal components of frequencies equal to pairs of side-frequencies centred around each spectral line of the unmodulated pulse train, expect the dc component.

It may be noted that these dc components represent the cross-modulation products between the sinusoidal modulation and sampling frequencies.

REVIEW QUESTIONS

1. State and prove sampling theorem in time domain.
2. What is Nyquist rate and Nyquist interval?
3. A bandlimited signal $x(t)$ is sampled by a train of rectangular pulses of width τ and period T.
   (i) Find an expression for the sampled signal.
   (ii) Determine the spectrum of the sampled signal and sketch it.
4. What is aliasing and how it is reduced?

NUMERICAL PROBLEMS

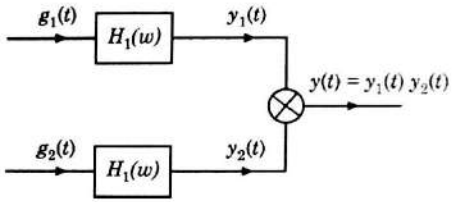
1. Determine the Nyquist sampling rate and the Nyquist sampling interval for the following singnals:
   (*a*) sinc (100 πt)
   (*b*) sinc^2 (100πt)
   (*c*) sinc (100πt) + sinc (50πt)
   (*d*) sinc (100πt) + 3 sinc^2 (60πt)
   (*e*) sinc (100πt) sinc (100πt)
2. A signal $g(t)$ bandlimited to BHz is sampling by a periodic pulse train $pT_s(t)$ made up of a rectangular pulse of with $\frac{1}{8}B$ seconds (centred at the origin) repeating at the Nyquist rate (2B pulses per second). Show that the sampled signal $\overline{g}(t)$ is given by

$$\overline{g}(t) = \frac{1}{4}g(t) + \sum_{n=1}^{\infty}\frac{2}{n\pi}\sin\left(\frac{n\pi}{4}\right)g(t)\cos nu_s t$$

$$\omega_s = 4\pi B$$

Show that the signal $g(t)$ can recovered by passing $\overline{g}(t)$ through an ideal low-pass filter of bandwidth BHz and a gain of 4.

3. Signal $g_1(t)$ = 104 rect (104t) and $g_2(t) = \delta(t)$ are applied at the inputs of ideal low-pass filters $H_1(\omega)$ = rect (ω/40, 000π) and $H_2(\omega)$ = rect (ω/20, 000π) (Figure 9.41). The outputs $y_1(t)$ and $y_2(t)$ of these filters are multiplied to obtain the signal $y(t) = y_1(t)\, y_2(t)$. Find the Nyquist rate of $y_1(t)$, $y_2(t)$, and $y(t)$. *(GATE Examination-1998)*



@sTc

Fig. 9.41.

4. A signal $g(t)$ sinc^2 (5πt) is sampled (using uniformly spaced impulses) at a rate of : (*i*) 5 Hz; (*ii*) 10 Hz; (*iii*) 20 Hz. Now, for each of the three cases:
   (*a*) Sketch the sampled signal.
   (*b*) Sketch the spectrum of the sampled signal.
   (*c*) Explain whether you can recover the signal $g(t)$ from the sampled signal.
   (*d*) If the sampled signal is passed through an ideal low-pass filter of bandwidth 5 Hz, sketch the spectrum of the output signal.
5. A zero-order hold circuit (Figure 9.42) is often used to reconstruct a signal $g(t)$ from its samples

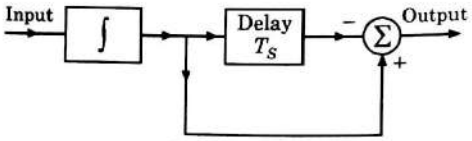


Fig. 9.42.

   (*a*) Find the unit impulse response of this circuit.
   (*b*) Find the transfer function $H(\omega)$ and sketch $|H(\omega)|$.
   (*c*) Show that when a sampled signal $\overline{g}(t)$ is sampled at the input of this circuit, the output is a staircase approximation of $g(t)$. The sampling interval is T_s. *(Pune University, 1998)*

OBJECTIVE TYPE QUESTIONS

Fill up the Blanks and Multiple Choice Questions

1. _____ is the process in which the analog signal is converted into a corresponding sequence of samples that the are uniformly spaced in time.
2. In the process of uniformly sampling, a signal in the time domain results in a periodic spectrum is the frequency domain with a period equal to the _____ .
3. The sampling rate (f_s) of value 2W samples per second for a signal bandwidth of W Hz is often referred to as _____ .
4. Prior to sampling a _____ is used to attenuate the high frequency components of the signal that lie outside the band of interest.
5. The bandpass signal is represented as a combination_____ and _____ components.
6. The aperture effect in flat top pulses is reduced by using an
   (*a*) Predictor (*b*) Integrator
   (*c*) Equalizer (*d*) Compander
7. The narrow samples produced at the pulse demodulator output are distributed to appropriate low pass reconstruction filter by means of
   (*a*) Commutator (*b*) Multiplexer
   (*c*) Decommutator (*d*) none of the above
8. Which of the following method is employed in telephony
   (*a*) Time division multiplexing (*b*) Frequency division multiplexing
   (*c*) both (*a*) and (*b*) (*d*) only (*a*)
9. In _____ the amplitudes of regularly spaced rectangular pulses vary with the instantaneous sample values of a continuous message signal in a one-to-one fashion.
10. The figure of merit of a PPM system is proportional to the square of _____.
11. PPM and PDM systems suffer from a _____ similar to that experienced in FM systems.
12. The sampled wave in practical system consists of _____ and _____ rather then impulses.
13. The term _____ refers to the signal from an adjacent channel spilling over into a desired time slot.
14. _____ is the process which involves sampling of the lowpass in-phase and quadrature components of the bandpass signal in accordance with the sampling theorem.
15. In _____ the samples of the message signal are used to vary the duration of the individual pulses.
16. Pulse duration modulation is also known as _____ (or) _____ .
17. In _____ the position of a pulse relative to its unmodulated time of occurrence is varied in accordance with the message signal.

ANSWERS

1. Sampling
2. Sampling rate
3. Nyquist rate
4. a lowpass pre-alias filter
5. In-phase, quadrature
6. (*c*)
7. (*c*)
8. (*c*)
9. Pulse amplitude modulation
10. normalized transmission bandwidth (B_T/W)
11. Threshold effect
12. Finite amplitude, finite duration pulses.
13. Crosstalk
14. Quadrature sampling
15. Pulse duration modulation
16. Pulse width (or) pulse length modulation
17. Pulse position modulation

□□□