I. PROBABILITY AND RANDOM VARIABLES

- **Conditional probability:**

$$P(A|B) = P(A \cap B)/P(B) = P(B|A)P(A)/P(B)$$

- **Random variable:** A mapping from sample space to real numbers.
- **Cumulative distribution function (CDF):**

$$F_X(x) = P(X \le x)$$

- **Probability density function (PDF):** $f_X(x) \ge 0 \; \forall x$ and $\int_{-\infty}^{\infty} f_X(x)dx = 1$.
- **Scalar Gaussian PDF:** $X \sim \mathcal{N}(\mu, \sigma^2)$, for $x \in \mathrm{R}$ $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$
- **Jointly Gaussian PDF:** $\mathbf{X} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, for $\mathbf{x} \in \mathrm{R}^n$ $f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{(2\pi)^{n/2}|\boldsymbol{\Sigma}|^{0.5}} \exp\left\{-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1}(\mathbf{x}-\boldsymbol{\mu})\right\}$
- **Q-function:** $Q(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-t^2/2} dt$ (i.e., probability that $\mathcal{N}(0,1)$ is larger than x)
- $E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x) dx$
- **Variance:** $Var(X) = E[X^2] - (E[X])^2$
- **Covariance:** $Cov(X,Y) = E[XY] - E[X]E[Y]$
- X and Y are *uncorrelated* if $E[XY] = E[X]E[Y]$
- X and Y are *independent* if $f_{X,Y}(x,y) = f_X(x) f_Y(y)$
- $f_{Y|X}(y|x) = f_{X,Y}(x,y)/f_X(x)$

II. RANDOM PROCESSES

An indexed family (ensemble) of random variables (equivalently, mapping from sample space to set of functions).

- **Mean (expectation) of a random process (r.p.):** $\mu_X(t) = E[X(t)] = \int_{-\infty}^{\infty} x f_{X(t)}(x) dx$
- **Autocorrelation function of a r.p.:** $R_X(t_1, t_2) = E[X(t_1) X^*(t_2)]$ $= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_1 x_2^* f_{X(t_1),X(t_1)}(x_1, x_2) dx_1 dx_2$
- **Autocovariance function of a r.p.:** $C_X(t_1,t_2) = E[(X(t_1) - \mu_X(t_1))(X(t_2) - \mu_X(t_2))^*]$ $= R_X(t_1,t_2) - \mu_X(t_1)\mu_X^*(t_2)$
- **Strict Sense Stationary (SSS) r.p.:** $X(t)$ is SSS if $f_{X(t_1+\tau),\ldots,X(t_k+\tau)}(x_1,\ldots,x_k) = f_{X(t_1),\ldots,X(t_k)}(x_1,\ldots,x_k)$ for all τ, k, $t_1,\ldots,t_k$.
- **Wide Sense Stationary (WSS) r.p.:** $X(t)$ is WSS if (i) $\mu_X(t) = \mu_X$ (i.e., constant) and (ii) $R_X(t_1,t_2) = R_X(t_1 - t_2)$.
- **Cyclostationary r.p.** if (i) $\mu_X(t) = \mu_X(t + T_0)$ and (ii) $R_X(t_1,t_2) = R_X(t_1 + T_0, t_2 + T_0)$. ($T_0$ is period.)
- For WSS $X(t)$, $R_X(\tau) = E[X(t+\tau)X^*(t)] = R_X^*(-\tau)$
- **Crosscorrelation function:** $R_{XY}(t_1,t_2) = E[X(t_1)Y^*(t_2)]$
- **Jointly WSS r.p.s:** $X(t)$ and $Y(t)$ are *jointly WSS* if (i) $X(t)$ is WSS, (ii) $Y(t)$ is WSS, and (iii) $R_{XY}(t_1,t_2) = R_{XY}(t_1 - t_2)$.
- $X(t)$ and $Y(t)$ are *independent* r.p.s if $(X(t_1),\ldots,X(t_k))$ and $(Y(u_1),\ldots,Y(u_l))$ are independent for all k, l, $(t_1,\ldots,t_k)$ and $(u_1,\ldots,u_l)$.
- $X(t)$ and $Y(t)$ are *uncorrelated* r.p.s if $X(t_1)$ and $Y(t_2)$ are uncorrelated r.v.s for all t_1 and t_2.
- A SSS r.p. is *ergodic* if time averages are equal to ensemble everages (expectations).
- **Filtering of a WSS r.p.:** If a WSS r.p. $X(t)$ passes through an LTI filter with impulse response $h(t)$, output $Y(t)$ is also WSS and $E[Y(t)] = H(0)E[X(t)]$, where $H(0) = \int_{-\infty}^{\infty} h(t)dt$.
- **Power Spectral Density (PSD):** Indicates distribution of average power among different frequencies. It is the Fourier transform (F.T.) of the autocorrelation function. $S_X(f) = \int_{-\infty}^{\infty} R_X(\tau) e^{-j2\pi f\tau} d\tau$ $R_X(\tau) = \int_{-\infty}^{\infty} S_X(f) e^{j2\pi f\tau} df$
- If a WSS r.p. $X(t)$ passes through an LTI filter with frequency response $H(f)$, output $Y(t)$ has the following PSD: $S_Y(f) = S_X(f)|H(f)|^2$
- $E[|X(t)|^2] = R_X(0) = \int_{-\infty}^{\infty} S_X(f) df$
- **Cross-Spectral Density:** $S_{XY}(f)$ is F.T. of $R_{XY}(\tau)$.
- **Gaussian r.p.:** $X(t)$ is Gaussian r.p. if $\int_0^T g(t)X(t)dt$ is a Gaussian r.v. for all $g(\cdot)$.
- **Gaussian r.p.:** $X(t)$ is Gaussian r.p. if $X(t_1),\ldots,X(t_n)$ are jointly Gaussian for all n, $t_1,\ldots,t_n$.
- If a Gaussian r.p. is WSS, it is also SSS.
- Linear (stable) filtering of a Gaussian r.p. leads to another Gaussian r.p.
- **White Noise:** Zero-mean WSS r.p. with $S_W(f) = N_0/2$ for all f (i.e., $R_W(\tau) = 0.5 N_0 \delta(\tau)$).
- **Baseband Representation of *Deterministic* Bandpass Signals:** $x(t) = Re\{\tilde{x}(t) e^{j2\pi f_c t}\}$ $x(t) = x_I(t)\cos(2\pi f_c t) - x_Q(t)\sin(2\pi f_c t)$
- One can also have baseband representation for bandpass random (noise) processes.

III. DIGITAL MODULATION AND DEMODULATION

- **M-ary communication system:** We have M different messages, $m_1, \ldots, m_M$. Transmitter modulates them as $s_1(t), \ldots, s_M(t)$ and receiver observes $r(t) = s_i(t) + n(t)$ for $t \in [0, T)$, where T is symbol duration and $n(t)$ is zero-mean additive white Gaussian noise (AWGN) with PSD of $S_n(f) = N_0/2$ for all f.
- **Basis function representation** of $s_1(t), \ldots, s_M(t)$: $s_i(t) = \sum_{j=1}^{N} s_{ij}\psi_j(t)$ with $s_{ij} = \int_0^T s_i(t)\psi_j(t)dt$, where $\psi_1(t), \ldots, \psi_N(t)$ are orthonormal basis functions. Then, $\boldsymbol{s}_i = [s_{i1} \cdots s_{iN}]^T$ is vector representation of $s_i(t)$.
- **Properties:** $\int_0^T s_i^2(t)dt = \|\boldsymbol{s}_i\|^2$, $\int_0^T s_i(t)s_j(t)dt = \boldsymbol{s}_i^T\boldsymbol{s}_j$.
- Two ways to find orthonormal basis functions: (1) Gram-Schmidt, (2) Intuition, trial and error.
- **Optimal Receiver for AWGN Channels:** Perform correlations of $r(t)$ with each of the orthonormal basis functions, call the resulting correlator outputs as $r_1, \ldots, r_N$. Defining $\boldsymbol{r} = [r_1 \cdots r_N]^T$, find the symbol which has the minimum distance to $\boldsymbol{r}$; that is, find $\underset{i\in\{1,\ldots,M\}}{\arg\min} \|\boldsymbol{r} - \boldsymbol{s}_i\|$.
- Given m_i (or, $\boldsymbol{s}_i$), $\boldsymbol{r}$ is a jointly Gaussian random vector with mean $\boldsymbol{s}_i$ and covariance matrix $0.5N_0\mathbf{I}$, where $\mathbf{I}$ is $N \times N$ identity matrix. (Note that $\boldsymbol{r} = \boldsymbol{s}_i + \boldsymbol{n}$.)
- **MAP Decision Rule:** $\hat{m} = \underset{i\in\{1,\ldots,M\}}{\arg\max}\, P(m_i \mid \boldsymbol{r})$
  Equivalently, $\hat{m} = \underset{i\in\{1,\ldots,M\}}{\arg\max}\, P(m_i)p(\boldsymbol{r} \mid m_i)$.
- **ML Decision Rule:** $\hat{m} = \underset{i\in\{1,\ldots,M\}}{\arg\max}\, p(\boldsymbol{r} \mid m_i)$
- **Union Bound:** $P_{e,i} \le \sum_{k=1,k\neq i}^{M} Q(\|\boldsymbol{s}_i - \boldsymbol{s}_k\|/\sqrt{2N_0})$
  $P_e \le \frac{1}{M}\sum_{i=1}^{M}\sum_{k=1,k\neq i}^{M} Q(\|\boldsymbol{s}_i - \boldsymbol{s}_k\|/\sqrt{2N_0})$
- **Loose Upper Bound:** $P_e \le (M-1)Q(d_{\min}/\sqrt{2N_0})$
- **M-ary PAM:** $s_i(t) = A_i p(t)$, $i = 1, \ldots, M$, $t \in [0, T)$. 1-dimensional signal space.
- **M-ary PSK:** $s_i(t) = A\,p(t)\cos(2\pi f_c t + \phi_i)$
  $i = 1, \ldots, M$, $t \in [0, T)$.
  In general, 2-dimensional signal space for $M > 2$.
- **M-ary QAM:** $s_i(t) = A_i p(t)\cos(2\pi f_c t + \theta_i)$
  $i = 1, \ldots, M$, $t \in [0, T)$.
  In general, 2-dimensional signal space for $M > 2$.
  Square-constellation M-QAM is equivalent to two $\sqrt{M}$-PAM's, each along one basis function.
- **M-ary FSK:** $s_i(t) = A\cos(2\pi f_i t + \phi_i)$
  $i = 1, \ldots, M$, $t \in [0, T)$.
  M-dimensional signal space assuming $f_i T$ is a distinct integer for each i.
- **Coherent FSK Receiver:** Receiver estimates phases (ϕ_i's) and uses them in the demodulator. Received signal $r(t)$ is correlated with $\sqrt{2/T}\cos(2\pi f_i t + \phi_i)$ for $i = 1, \ldots, M$, and the symbol corresponding to maximum correlator output is selected.
- **Non-coherent FSK Receiver:** Receiver does not estimate phases (ϕ_i's). Received signal $r(t)$ is correlated with both $2\cos(2\pi f_i t)$ and $2\sin(2\pi f_i t)$, and the squares of the correlator outputs are added for $i = 1, \ldots, M$. The symbol corresponding to maximum output is selected.
- **M-ary Biorthogonal Signals:** $M/2$ orthogonal signals and their negatives.
- **M-ary Simplex Signals:** Find the average of M orthogonal, equal-energy signals and subtract that average from each signal. (Reduces average signal power.)
- **Comparison:** PAM, PSK, QAM are spectrally efficient but not power efficient. FSK (orthogonal) modulation is power efficient but not spectrally efficient.

IV. MISCELLANEOUS FORMULAS

- $\mathrm{sinc}(x) = \sin(\pi x)/(\pi x)$
- F.T. of $\mathrm{sinc}(t)$ is a rectangular pulse of amplitude 1 between -0.5 and 0.5.
- F.T. of $\mathrm{sinc}^2(t)$ is a triangular pulse between -1 and 1 with maximum amplitude of 1 at zero.
- $\sin(2x) = 2\sin(x)\cos(x)$
- $\cos(2x) = 1 - 2\sin^2(x) = 2\cos^2(x) - 1$
- $\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y)$
- $\sin(x+y) = \sin(x)\cos(y) + \cos(x)\sin(y)$
- $\sin(x)\cos(y) = 0.5\sin(x+y) + 0.5\sin(x-y)$
- $\cos(x)\cos(y) = 0.5\cos(x+y) + 0.5\cos(x-y)$
- $\sin(x)\sin(y) = 0.5\cos(x-y) - 0.5\cos(x+y)$
- $\cos(\pi/3) = \sin(\pi/6) = 1/2$
- $\cos(\pi/6) = \sin(\pi/3) = \sqrt{3}/2$
- $\cos(\pi/4) = \sin(\pi/4) = 1/\sqrt{2}$