Problem 9.1:

(a)

$$\begin{aligned} h(X) &= -\int_0^\infty \lambda e^{-\lambda x}[\ln\lambda - \lambda x]dx \text{ (in nats)} \\ &= -\ln\lambda \underbrace{\left[\int_0^\infty \lambda e^{-\lambda x}dx\right]}_{=1} + \lambda \underbrace{\left[\int_0^\infty x\lambda e^{-\lambda x}dx\right]}_{=E[X]=1/\lambda} \\ &= 1-\ln\lambda \text{ nats} \\ &= (1-\ln\lambda)/\ln 2 \text{ bits.} \end{aligned}$$

(b)

$$\begin{aligned} h(X) &= -\int_{-\infty}^\infty \frac{1}{2}\lambda e^{-\lambda|x|}\left[\ln\frac{\lambda}{2} - \lambda|x|\right]dx \\ &= -\ln(\lambda/2) + \lambda E[|X|] \\ &= -\ln(\lambda/2) + 2\lambda\int_0^\infty (x/2)\lambda e^{-\lambda x}dx \\ &= -\ln(\lambda/2) + (2\lambda)\frac{1}{2\lambda} \\ &= 1-\ln(\lambda/2) \text{ nats.} \end{aligned}$$

(c) $Y = X_1 + X_2$, where $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$.

$$\Longrightarrow Y \sim \mathcal{N}(\mu_1+\mu_2, \sigma_1^2+\sigma_2^2)$$

$$\Longrightarrow h(Y) = \frac{1}{2}\log_2(2\pi e(\sigma_1^2+\sigma_2^2))$$

Problem 9.3:

__For $\rho = 0$:__ X, Y are uncorrelated Gaussian $\Longrightarrow X, Y$ are independent $\Longrightarrow I(X;Y) = 0$.

__For general case ($-1 \le \rho \le 1$):__

$$I(X;Y) = h(X) + h(Y) - h(X,Y)$$

where

$$h(X) = h(Y) = \frac{1}{2}\log_2(2\pi e\sigma^2)$$

$$\begin{aligned} h(X,Y) &= \frac{1}{2}\log_2\left[(2\pi e)^2\begin{vmatrix}\sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2\end{vmatrix}\right] \quad \text{from (9.94)} \\ &= \frac{1}{2}\log_2\left((2\pi e)^2(1-\rho^2)\sigma^4\right) \end{aligned}$$

$$\Longrightarrow I(X;Y) = \frac{1}{2}\log_2\left(\frac{1}{1-\rho^2}\right)$$

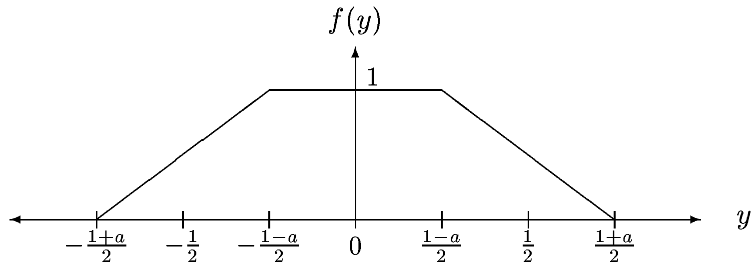
__For $\rho = \pm 1$:__ In this case, $I(X;Y) = +\infty$. This is because $X = \pm Y$ with probability one.

It is important to note that $I(X;Y)$ is independent of σ^2. Also note that there is a one-to-one correspondence between $|\rho|$ and $I(X;Y)$. The higher the (magnitude of) correlation, the higher the mutual information. Zero correlation implies zero mutual information.
The correlation and the mutual information are both measures of dependency.

Problem 9.4:
X is uniformly distributed on $[-1/2, +1/2]$. Z is uniformly distributed on $[-a/2, +a/2]$. $Y = X+Z$

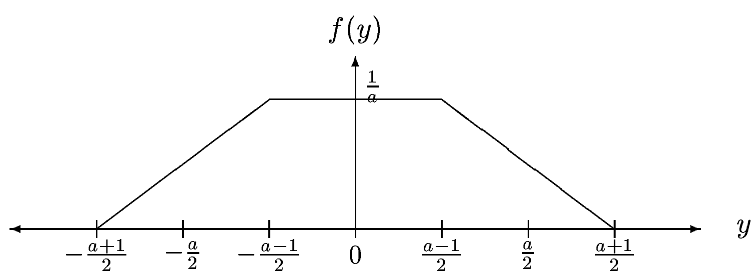
(a)

Case 1: $a \leq 1$**:** Here, the pdf of Y looks like:



$$\begin{aligned}
h(Y) &= -2\int_0^{(1-a)/2} 1\log_2 1 dy - 2\int_{(1-a)/2}^{(1+a)/2} -\frac{1}{a}\left(y-\frac{1+a}{2}\right)\log_2\left[-\frac{1}{a}\left(y-\frac{1+a}{2}\right)\right] dy \\
&= 0 + \frac{2}{a}\int_{-a}^{0} x\log_2\left[-\frac{x}{a}\right] dx \qquad \left\{x = y - \frac{1+a}{2}\right\} \\
&= -2a\int_0^1 t\log_2 t dt \qquad \left\{t = -\frac{x}{a};\, dx = -adt\right\} \\
&= \frac{-2a}{\ln 2}\left[\frac{t^2}{2}\ln t - \frac{t^2}{4}\right]_0^1 \\
&= \frac{a}{2\ln 2} \text{ bits} = \frac{a}{2} \text{ nats.}
\end{aligned}$$

Case 2: $a > 1$**:** Here, the pdf of Y looks like:



$$\begin{aligned}
h(Y) &= -2\int_0^{(a-1)/2} \frac{1}{a}\log_2\frac{1}{a} dy - 2\int_{(a-1)/2}^{(a+1)/2} -\frac{1}{a}\left(y-\frac{a+1}{2}\right)\log_2\left[-\frac{1}{a}\left(y-\frac{a+1}{2}\right)\right] dy \\
&= \ln a\left(\frac{a-1}{a}\right) - 2\int_0^1 \frac{x}{a}\log_2\left[\frac{x}{a}\right] dx \qquad \left\{x = -y + \frac{a+1}{2}\right\}
\end{aligned}$$

$$
\begin{aligned}
&= \left(1-\frac{1}{a}\right)\ln a - 2a\int_0^{\frac{1}{a}} t\log_2 t dt && \left\{t=\frac{x}{a}; dx = adt\right\} \\
&= \left(1-\frac{1}{a}\right)\ln a - 2a\left[\frac{t^2}{2}\ln t - \frac{t^2}{4}\right]_0^{\frac{1}{a}} \\
&= \ln a + \frac{1}{2a} \text{ nats.}
\end{aligned}
$$

Now, $h(Z) = \ln a$ nats. Thus

$$
I(X;Y) = h(Y) - h(Z) = \begin{cases} \frac{a}{2} - \ln a \text{ nats} & \text{if } 0 \le a \le 1 \\ \frac{a}{2} \text{ nats} & \text{if } a > 1 \end{cases}
$$

(b) If $a = 1$, then $X \in [-1/2, +1/2]$ and $Z \in [-1/2, +1/2]$. This implies that $Y \in [-1, +1]$. Since, $I(X;Y) = h(Y) - h(Z) = h(Y) - log_2(1) = h(Y)$,

$$
\max_{|X|\le 1/2} I(X;Y) = \max_{|X|\le 1/2} h(Y).
$$

The maximum value of $h(Y)$ is achieved when Y is uniformly distributed on $[-1, +1]$. This occurs if X is discrete and $\Pr\{X = -1/2\} = \Pr\{X = +1/2\} = \frac{1}{2}$. Thus, the channel capacity is

$$
C = \log_2 2 - \log_2 1 = 1 \text{ bit.}
$$

For this special channel, we can find a simple one-dimensional code with $R = C$ and λ_{max}=maximum probability of error=0. To do this, we send one of the two codewords: $x_1 = -1/2$ or $x_2 = +1/2$. (Note that these satisfied the *amplitude* constraint. When we receive y, we decode x_1 if $y < 0$ and we decode x_2 if $y >= 0$. This results in zero probability of error.

Problem 9.5:
The decay time is typically modeled as an exponential distribution, $f(x) = \lambda e^{-\lambda x}, \;\; x \ge 0$. From problem 9.1(a), we know $h(X) = \log_2(e/\lambda)$ bits.
Since the median is 80,

$$
\int_0^{80} \lambda e^{-\lambda x} dx = \frac{1}{2} \Leftrightarrow (1 - e^{-80\lambda}) = \frac{1}{2}.
$$

So $\lambda = (\ln 2)/80 = 0.00866$.
Three digits accuracy are equal to $\lceil \log_2 1000 \rceil = \lceil 9.97 \rceil = 10$ fractional bits.
Thus, to represent the r.v. to 10 fractional bits accuracy we need $\approx h(X) + 10 = \log_2 \frac{e}{\lambda} + 10 = 18.29$ bits.

Problem B:
The covariance matrix is:

$$
K = \begin{bmatrix} 1 & \rho & 0 \\ \rho & 1 & \rho \\ 0 & \rho & 1 \end{bmatrix}
$$

and $\det K = 1 - 2\rho^2$. Thus,

$$
\begin{aligned}
h(X,Y,Z) &= \frac{1}{2}\log(2\pi e)^3(\det K) \\
&= \frac{1}{2}\log(2\pi e)^3(1 - 2\rho^2) \\
&= \frac{3}{2}\log(2\pi e) + \frac{1}{2}\log(1 - 2\rho^2).
\end{aligned}
$$

Since (det K) must be positive, we need that $1 - 2\rho^2 > 0$ or $|\rho| < \sqrt{2}/2$.

Please let me know if you see a mistake in this solution.