Digital Communication using GNU Radio

Prof. Kumar Appaiah
Department of Electrical Engineering
Indian Institute of Technology Bombay

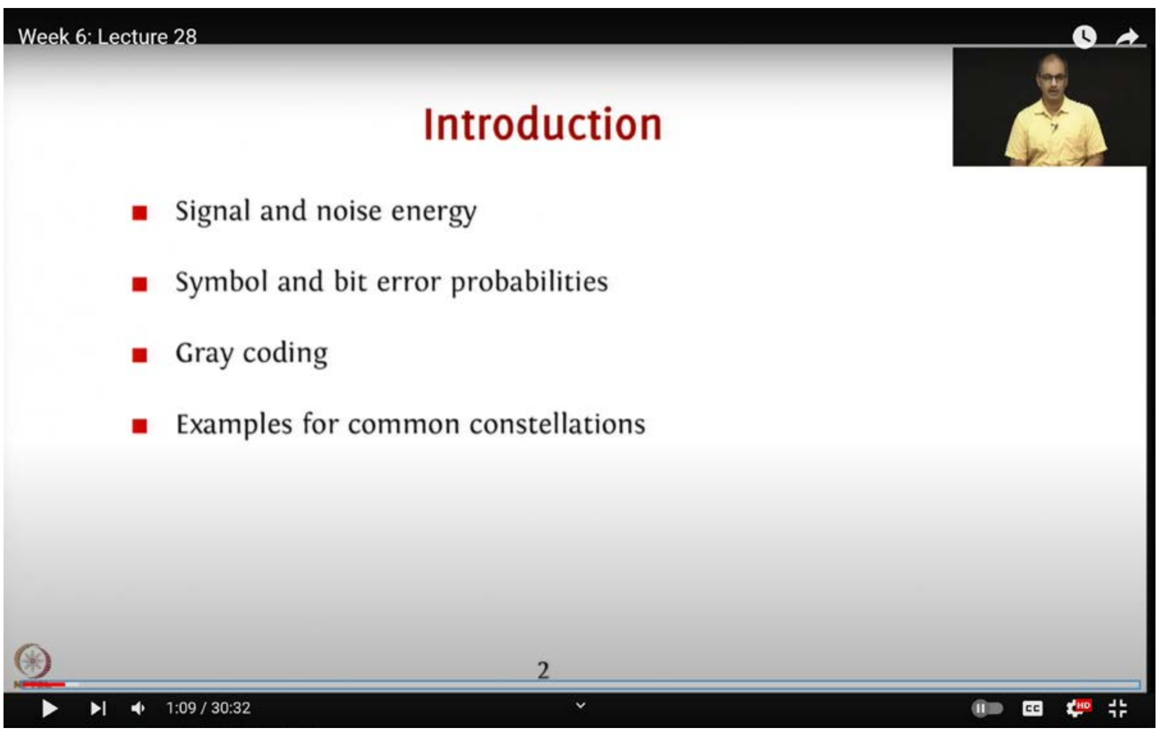
Week-06

Lecture-28

Signal-to-Noise Ratio and Symbol Error Probability - Part 1

Welcome to this lecture on Digital Communication using GNU Radio. I'm Kumar Appiah from the Department of Electrical Engineering at IIT Bombay. In this lecture, we will continue our exploration of demodulation, focusing on how to compute symbol and bit error rates for various modulation formats.

(Refer Slide Time: 01:09)



To set the stage, over this lecture and the next, we will cover several key topics:

1. Signal and Noise Energy: Understanding the relationship between these elements and their impact on communication performance.

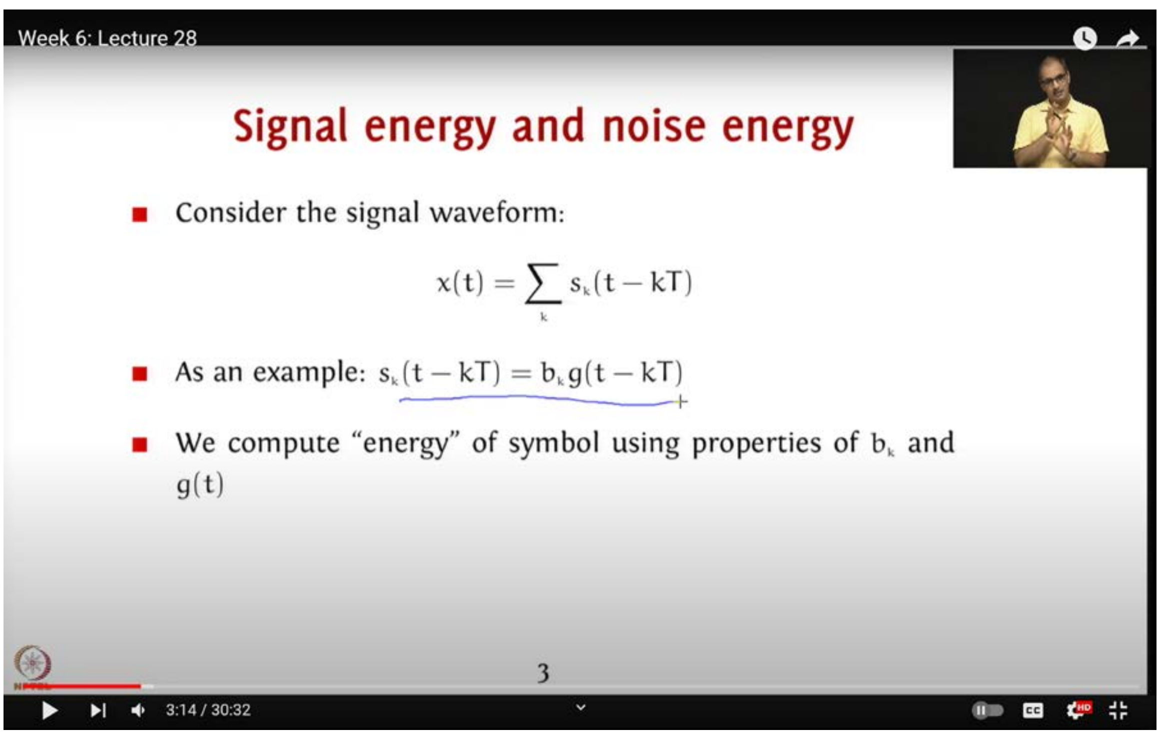
2. Symbol and Bit Error Probabilities: Analyzing how different modulation formats affect the likelihood of symbol errors.

3. Gray Coding: Delving into bit error rates and examining techniques like Gray coding to optimize performance and minimize error rates.

4. Examples for Common Constellations: Reviewing common constellations such as Pulse Amplitude Modulation (PAM) and Quadrature Amplitude Modulation (QAM) through practical examples.

An essential point we covered in previous lectures is that the performance of a communication system is influenced not just by the signal energy or noise energy individually, but by the signal-to-noise ratio (SNR), the ratio of signal power to noise power, or signal energy to noise energy. Hence, accurately computing and characterizing both signal and noise energy is crucial.

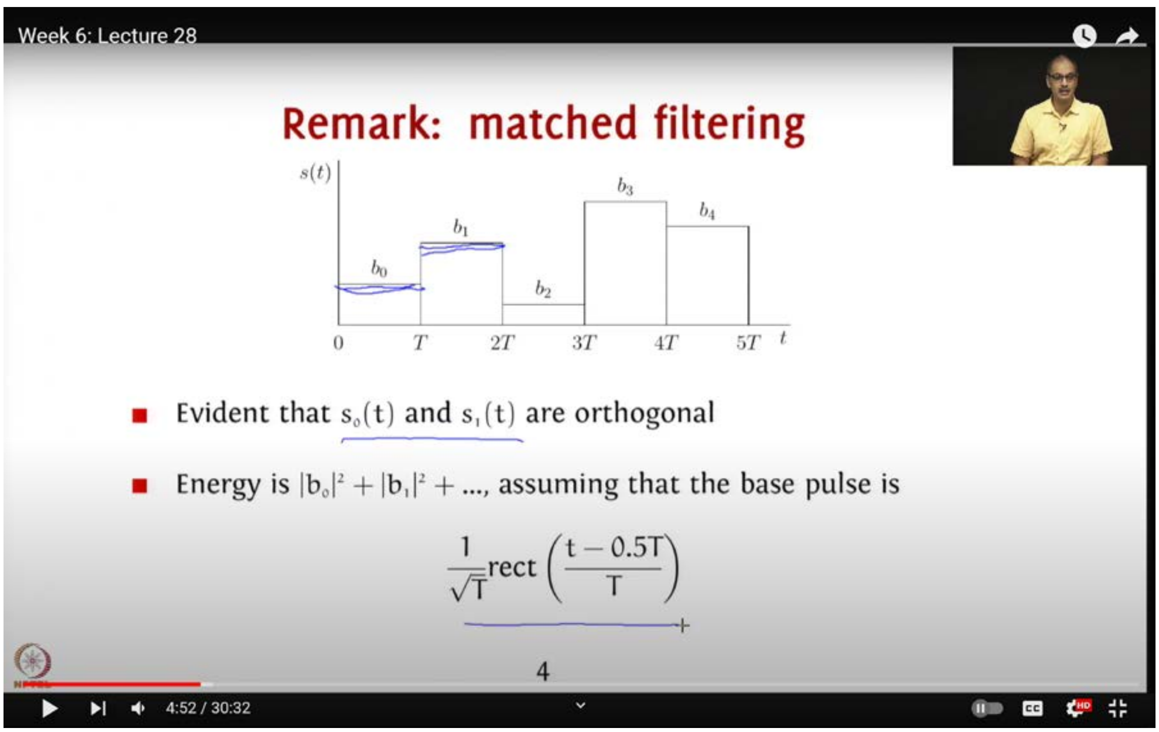
(Refer Slide Time: 03:14)



Regarding noise energy, we have previously explored this in the context of signal space and the projection of noise onto signal spaces and orthogonal basis signals. For signals, we apply a similar concept, but we will formalize the assumptions for this lecture.

Consider a signal waveform X(t), which can be expressed as $\sum s_k(t-kT)$. This waveform consists of a sequence of individual waveforms, each containing information about a symbol. Specifically, each waveform $s_k(t - kT)$ can be represented as $b_k\, g(t - kT)$, where g(t) is the template pulse used by the transmitter. For simplicity in this discussion, we will assume g(t) is our general pulse shape.

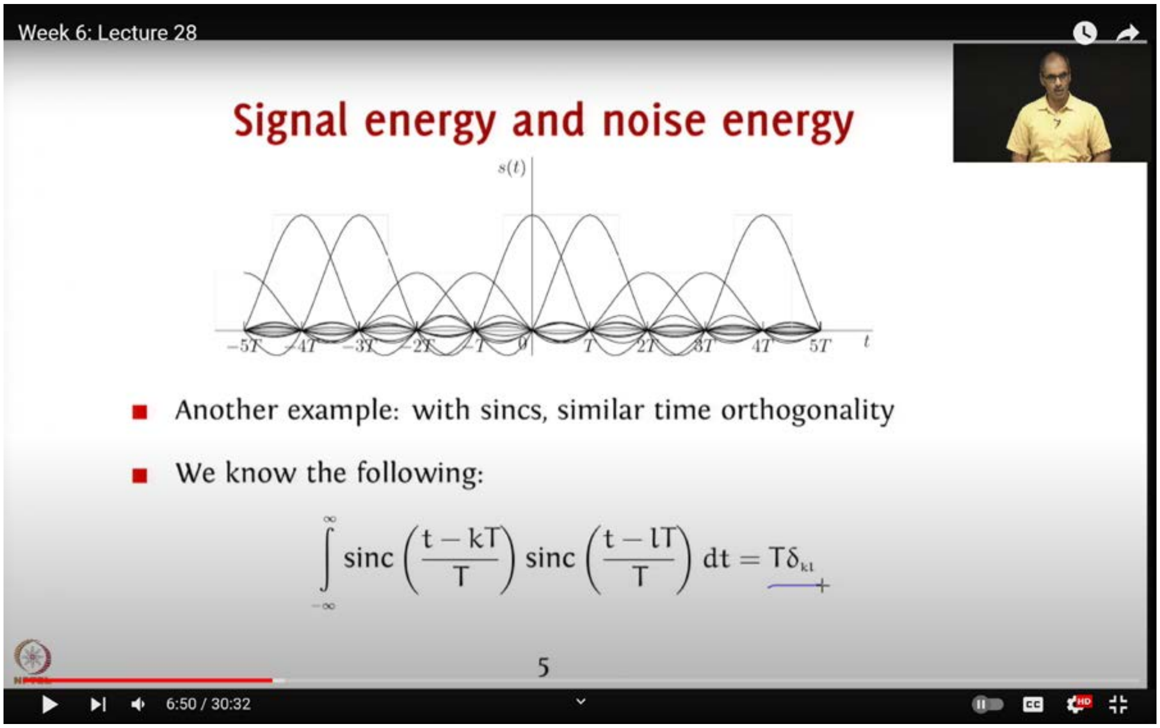
(Refer Slide Time: 04:52)



In digital communication, each symbol s_k conveys a different information value b_k. To compute the energy of a symbol, we need to consider the properties of b_k and g(t). Specifically, if $b_k \cdot g(t)$ represents the information contained in x(t) about the symbol, we need to account for the energy of this component.

Typically, symbols might overlap or interact, such as when using a sinc pulse, which spans across multiple symbols. However, for simplicity, we assume that the energy of each symbol is additive and independent of others. For example, in matched filtering scenarios, if s(t) is represented by a series of symbols like b_0, b_1, b_2, and b_3, and if you perform matched filtering, you're convolving the signal with a rectangular function.

(Refer Slide Time: 06:50)



To compute the energy of a symbol, we consider the energy contributions of each component:

- The energy of b_0 is $|b_0|^2 \cdot E_g$,
- The energy of b_1 is $|b_1|^2 \cdot E_g$,
- And so forth.

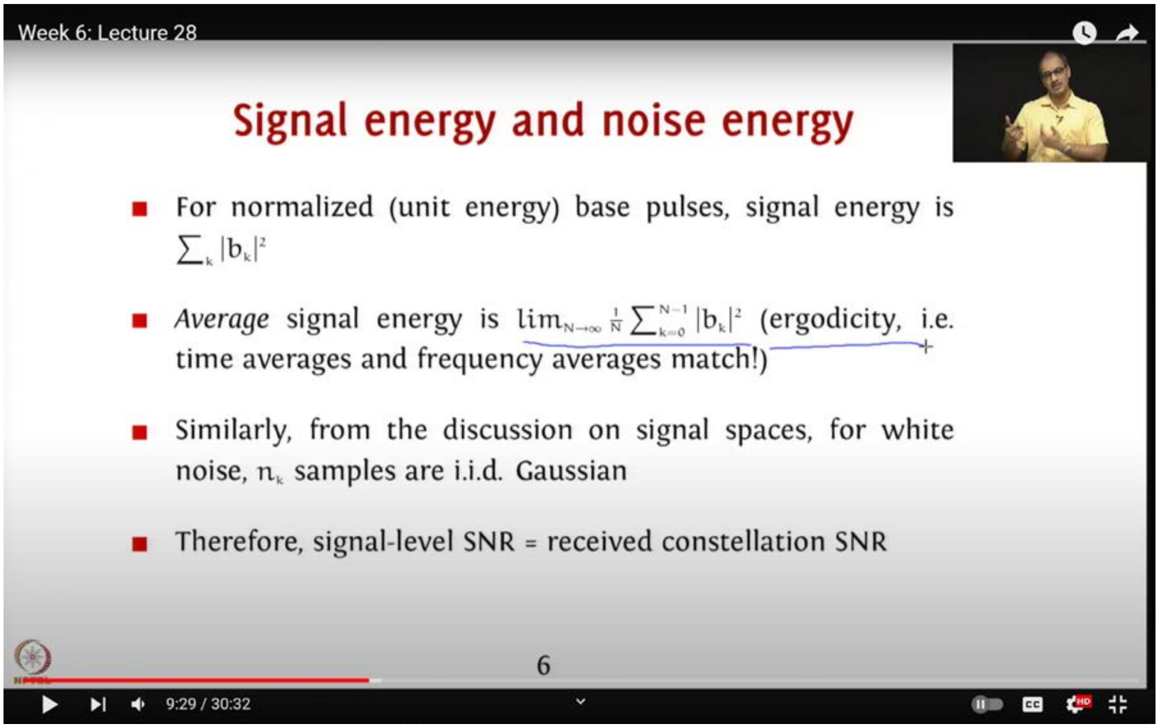
Here, E_g represents the energy of the pulse g(t). If we assume g(t) has unit energy, the energy of the signal is simply the sum of the squared magnitudes of the coefficients:

$$\text{Energy} = |b_0|^2 + |b_1|^2 + |b_2|^2 +$$

This assumes that the signals s_0 and s_1 are orthogonal. For instance, s_0 could be a rectangle pulse from 0 to T with amplitude b_0, while s_1 extends from T to 2T with amplitude b_1.

These signals are orthogonal by definition.

(Refer Slide Time: 09:29)



If we assume that the base pulse is a rectangular function with unit energy, the energy calculation holds true. To verify this, you can integrate the pulse energy to confirm.

In more complex scenarios, like using sinc functions, which carry information over an extended period, we can still apply the concept of orthogonality. Although sinc functions extend infinitely, they exhibit a form of orthogonality similar to rectangular pulses. If you integrate the product of a sinc function $\text{sinc}(t - kT)$ and another sinc function $\text{sinc}(t - lT)$, scaled appropriately, the result is $T \cdot \delta_{kl}$. This indicates that integrating a sinc function with itself gives a certain energy, while integrating two sinc functions that are offset by T results in zero. Thus, sinc functions demonstrate time orthogonality similar to rectangular pulses.

To accurately compute the energy of a signal, it's essential to understand the properties of the constellation, such as whether the symbols are BPSK, QAM-16, or another modulation format, as well as the base pulse g(t). In our example, we assumed that the base pulse has unit energy and scaled the sinc functions appropriately. This simplifies our calculation, so moving forward, we'll assume that the energy contained in our modulated signal is just the summation of $b_0^2,\ b_1^2$, and so forth.

You might wonder about the situation when dealing with passband signals. Fortunately, even in passband, as long as the scaling factors like $\sqrt{2}$ are applied correctly, the energy of the passband signal remains consistent. Therefore, there is no confusion or additional complexity here.

The energy of a signal, when normalized with a unit energy base pulse, is simply the sum of the squares of the constellation points b_k. To compute the energy of a signal, take each realized constellation point b_0, b_1, b_2, etc., square their magnitudes, and sum them up. This gives you the energy of the signal. However, as more symbols are added, the total energy will naturally increase. What we are more interested in is the average energy, or the power of the modulated signal, since this reflects the energy expenditure per symbol.

To determine the average energy, consider a large block of symbols. Compute the total energy of this block and divide by the number of symbols, n. For instance, if n is 10,000 or a million, you average the energy over these symbols. Mathematically, this is expressed as:

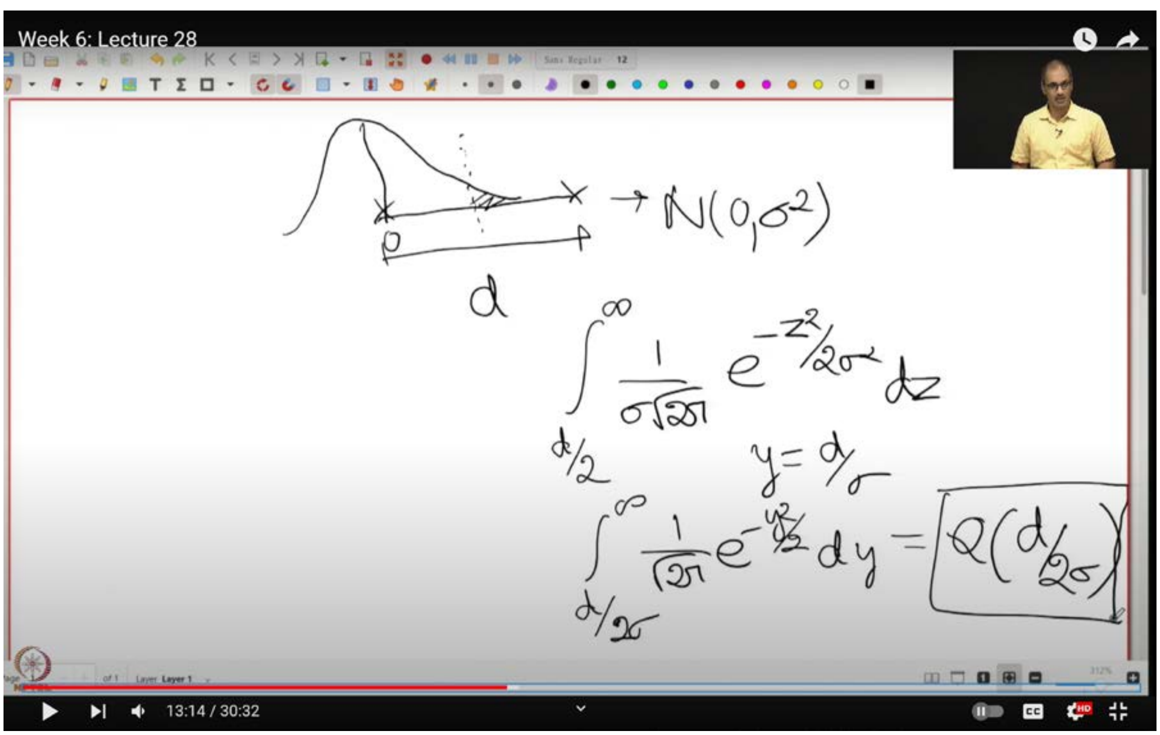
$$\text{Average Energy} = \lim_{n\to\infty} \frac{1}{n}\sum_{k=0}^{n-1} |b_k|^2$$

An important concept here is ergodicity. Even if you don't observe a vast sequence of constellation points, you can use ergodicity principles to compute the average energy. This means that time averages and frequency averages are equivalent. Therefore, to find the average energy of a constellation, you don't need to observe a massive sequence of points. Instead, you can probabilistically compute it directly from the constellation. Since these random processes generated by the symbols are ergodic, the average energy of your modulated signal will match the average energy calculated from the constellation itself.

In this context, it is often much simpler to examine the constellation and determine the amount of energy it contains. Similarly, from our previous discussion on signal spaces, we know that n_k samples are also i.i.d. Gaussian, thanks to the orthogonality principles we discussed. Therefore, we will assume that the signal-to-noise ratio (SNR) at the signal level is equivalent to the received constellation SNR. This is because the received constellation, which includes all the noise, is processed through operations like matched filtering and sampling. These operations do not introduce additional noise, so the SNR present before these operations will remain the same when you receive and decode the sampled points to determine which constellation symbol was transmitted.

Now, let's move on to the basic calculation of symbol error rate. We will use our standard approach, which we discussed previously in the context of modulation and symbol errors. Suppose we have two constellation points with a distance d between them, and the noise has a mean of 0 and a variance of σ^2. In this setup, the optimal decision region is determined by the distance from the received point to each of the constellation points. When dealing with Gaussian noise, the decision-making process involves minimizing the squared error between the received signal and each possible constellation point.

(Refer Slide Time: 13:14)



If the received point falls within the decision region closer to one constellation point, it is chosen as the transmitted symbol. However, if the Gaussian noise causes the received signal to fall into the decision region of the other point, an error occurs. The probability of making such an error is given by the function $Q\left(\frac{d}{2\sigma}\right)$, where Q is the Q-function, representing the probability of error for both constellation points, assuming they are equally probable.

To recall, the Q-function represents the tail probability of the Gaussian distribution. If we assume a Gaussian noise with mean 0, the error occurs if the realization of this Gaussian noise exceeds $\frac{d}{2}$. To compute this probability, we need to evaluate the integral:

$$\int_{\frac{d}{2}}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{z^2}{2\sigma^2}\right) dz$$

To simplify this, we can use the substitution $y = \frac{d}{\sigma}$, leading to the integral:

$$\int_{\frac{d}{2\sigma}}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{y^2}{2}\right) dy$$

This integral evaluates the tail probability of the Gaussian distribution, which provides the symbol error rate for this modulation scheme.

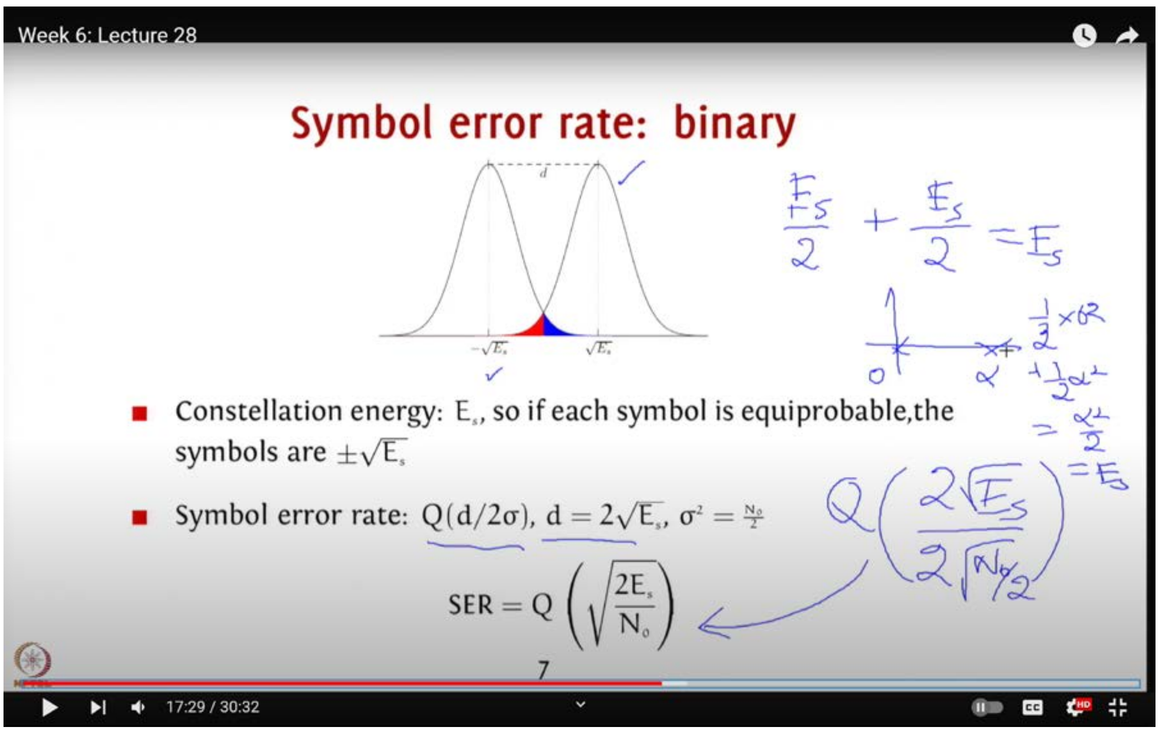
Let me just grab my eraser. I'll erase this part and then grab my pen. The integral we are discussing is:

$$\int_{\frac{d}{2}}^{\infty} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{y^2}{2\sigma^2}\right) dy$$

This represents the tail integral of a standard Gaussian distribution with unit variance and zero mean, and is known as the Q-function, denoted as $Q\left(\frac{d}{2\sigma}\right)$. The Q-function is a crucial formula we will use whenever dealing with a pair of constellation points separated by a distance d. The Q-function $Q\left(\frac{d}{2\sigma}\right)$ gives us the probability of making an error when the two points are equally likely. For the purposes of our discussion, we will assume the points are equally likely.

Now, let's examine a specific constellation, known as Binary Phase Shift Keying (BPSK). This constellation involves binary signaling with equal energy pulses. In this case, we choose the constellation points to be $\pm\sqrt{E_s}$, where E_s represents the average energy per symbol. The reason for selecting $\pm\sqrt{E_s}$ is that, if the symbols are either +a or -a, and both appear with equal probability, then choosing $\pm\sqrt{E_s}$ ensures that the average energy per symbol is E_s.

(Refer Slide Time: 17:29)



If you recall, our base pulse had unit energy, so the signal energy when sending +a is a^2, and similarly, the signal energy when sending -a is also a^2. Therefore, choosing the constellation points as $\pm\sqrt{E_s}$ means that each symbol has an average energy of E_s. This choice ensures that the average energy for the transmitted symbols, whether $+\sqrt{E_s}$ or $-\sqrt{E_s}$, is E_s, and each is sent with a probability of 0.5.

For BPSK, the distance d between the constellation points is $2\sqrt{E_s}$. Since the noise variance along each dimension is $\sigma^2 = \frac{N_0}{2}$, the formula for the probability of error using the Q-function becomes:

$$Q\left(\frac{2\sqrt{E_s}}{2\sigma}\right) = Q\left(\frac{\sqrt{E_s}}{\sigma}\right)$$

Here, σ^2 is the noise variance, which is $\frac{N_0}{2}$ for each dimension.

So, we refer to the noise variance as $\frac{N_0}{2}$. Plugging this into our Q-function formula, where d is $2\sqrt{E_s}$ and σ is $\sqrt{\frac{N_0}{2}}$, we get:

$$Q\left(\frac{2\sqrt{E_s}}{2\sqrt{\frac{N_0}{2}}}\right) = Q\left(\frac{\sqrt{2E_s}}{\sqrt{N_0}}\right)$$

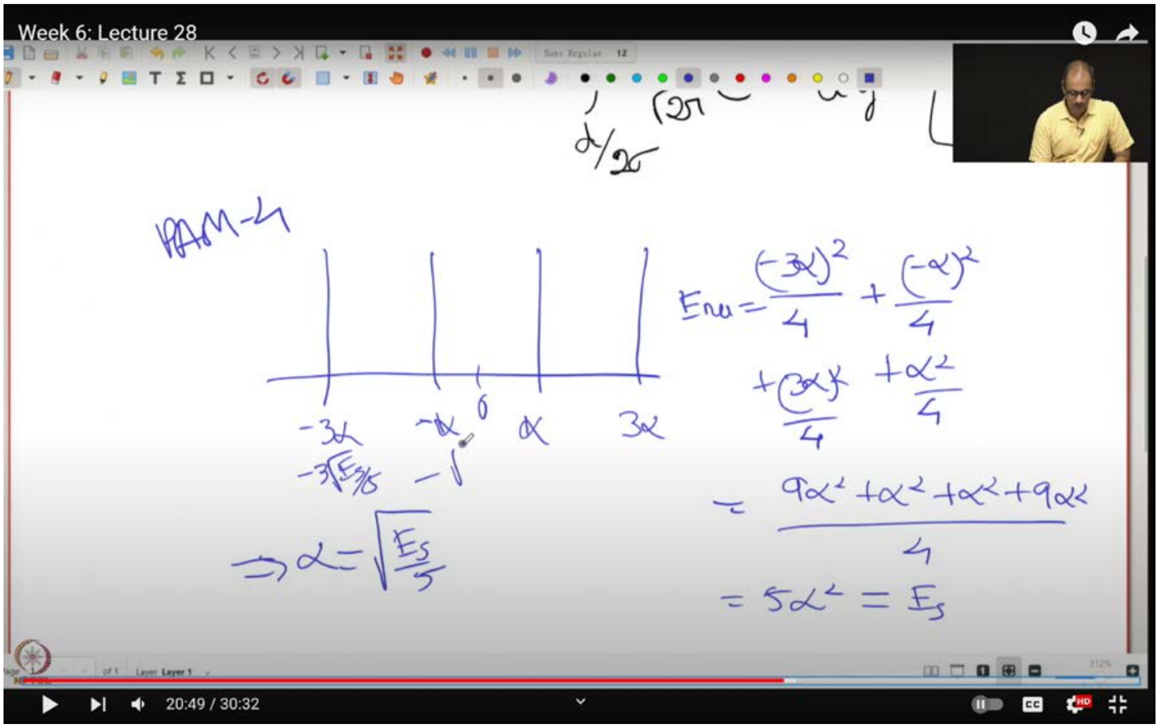
Thus, the symbol error rate for binary phase shift keying (BPSK) is:

$$Q\left(\frac{\sqrt{2E_s}}{N_0}\right)$$

It's worth noting that we chose our constellation points to be $+\sqrt{E_s}$ and $-\sqrt{E_s}$. If we had used a different constellation where one symbol is 0 and the other is $\sqrt{2E_s}$, the energy calculations would be different. Specifically, if you use symbols 0 and $\sqrt{2E_s}$, the average energy calculation would yield a different result. For instance, if α is the symbol and you want the average energy to be E_s, you would need to set $\alpha = \sqrt{2E_s}$. The distance d between the symbols in this case is $\sqrt{2E_s}$.

Choosing a constellation with a symbol at 0 and another at $\sqrt{2E_s}$ will result in slightly poorer performance, with a higher symbol error rate compared to using $\pm\sqrt{E_s}$. Intuitively, given an energy budget, your goal should be to maximize the separation between constellation points.

(Refer Slide Time: 20:49)



Now, let's examine PAM-4 (Pulse Amplitude Modulation with 4 levels). For PAM-4, the constellation typically looks like this:

$$-3\alpha, -\alpha, \alpha, 3\alpha$$

These points are equi-probable, and we want to find α such that the constellation has an energy E_s.

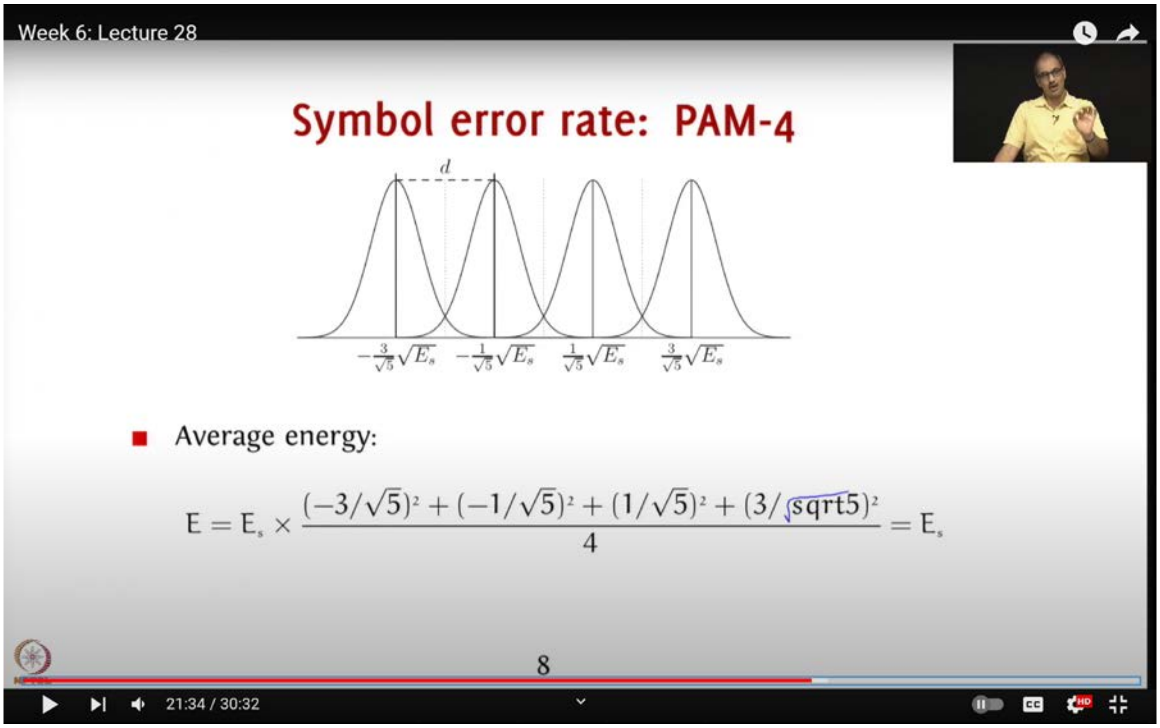
To do this, let's calculate the average energy for the PAM-4 constellation. The energy for each symbol is:

- For -3α: $(-3\alpha)^2 = 9\alpha^2$
- For -α: $(-\alpha)^2 = \alpha^2$
- For α: α^2
- For 3α: $(3\alpha)^2 = 9\alpha^2$

Since each symbol is equally probable, the average energy is:

$$\frac{1}{4}(9\alpha^2 + \alpha^2 + \alpha^2 + 9\alpha^2) = \frac{20\alpha^2}{4} = 5\alpha^2$$

(Refer Slide Time: 21:34)



Setting this equal to E_s, we get:

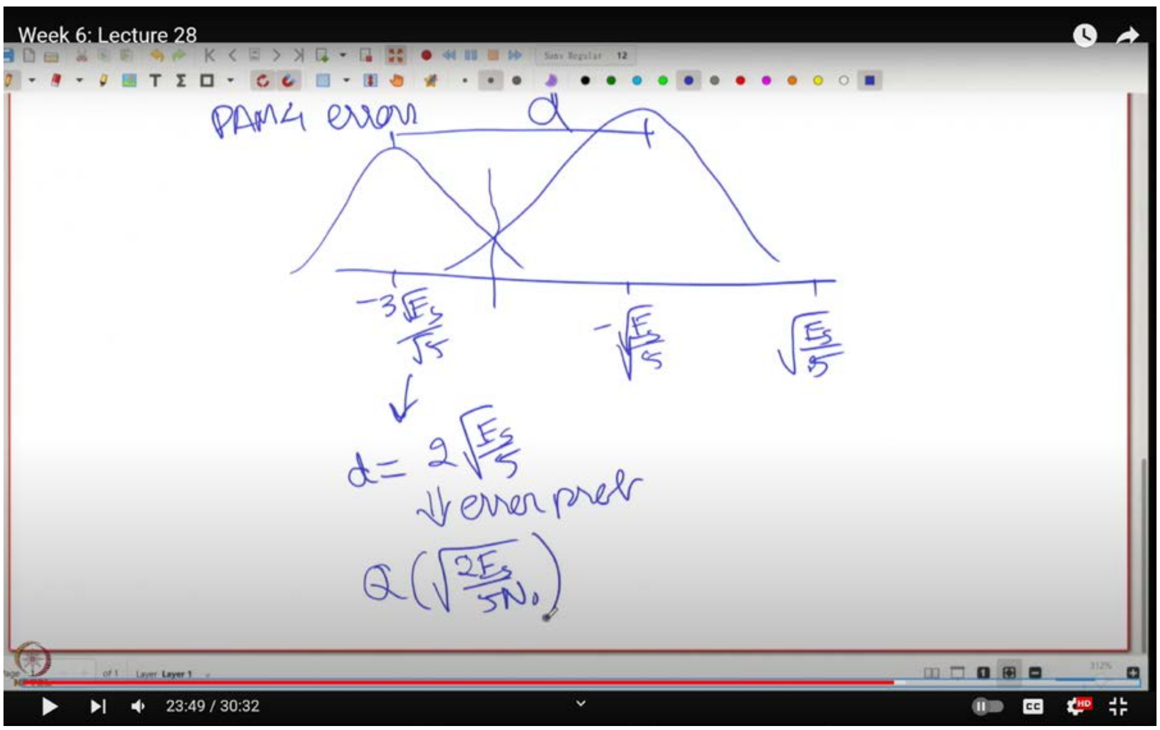
$5\alpha^2 = E_s$ imples $\alpha = \sqrt{\frac{E_s}{5}}$

So, substituting $\alpha = \sqrt{\frac{E_s}{5}}$ gives us a PAM-4 constellation with an energy of E_s. This calculation demonstrates that by choosing equi-probable symbols and spacing them symmetrically around zero, you can achieve a PAM-4 constellation with the desired energy. Any deviation from this configuration could result in inefficient power usage and a higher symbol error rate for the same E_s.

If you're given an E_s and need to design a PAM-4 constellation to minimize the symbol error rate under equi-probable signaling, here's how you should arrange it. First, the value for α is $\sqrt{\frac{E_s}{5}}$. Therefore, your four constellation points should be:

$$-3\sqrt{\frac{E_s}{5}}, -\sqrt{\frac{E_s}{5}}, \sqrt{\frac{E_s}{5}}, 3\sqrt{\frac{E_s}{5}}$$

(Refer Slide Time: 23:49)



I've written these points as:

$$-3\sqrt{\frac{E_s}{5}}, -\sqrt{\frac{E_s}{5}}, \sqrt{\frac{E_s}{5}}, 3\sqrt{\frac{E_s}{5}}$$

To verify that the average energy is correct, you square each point, sum them up, and then divide by 4. This is the same calculation we did with $\sqrt{5}$. For these points, the average energy calculation is:

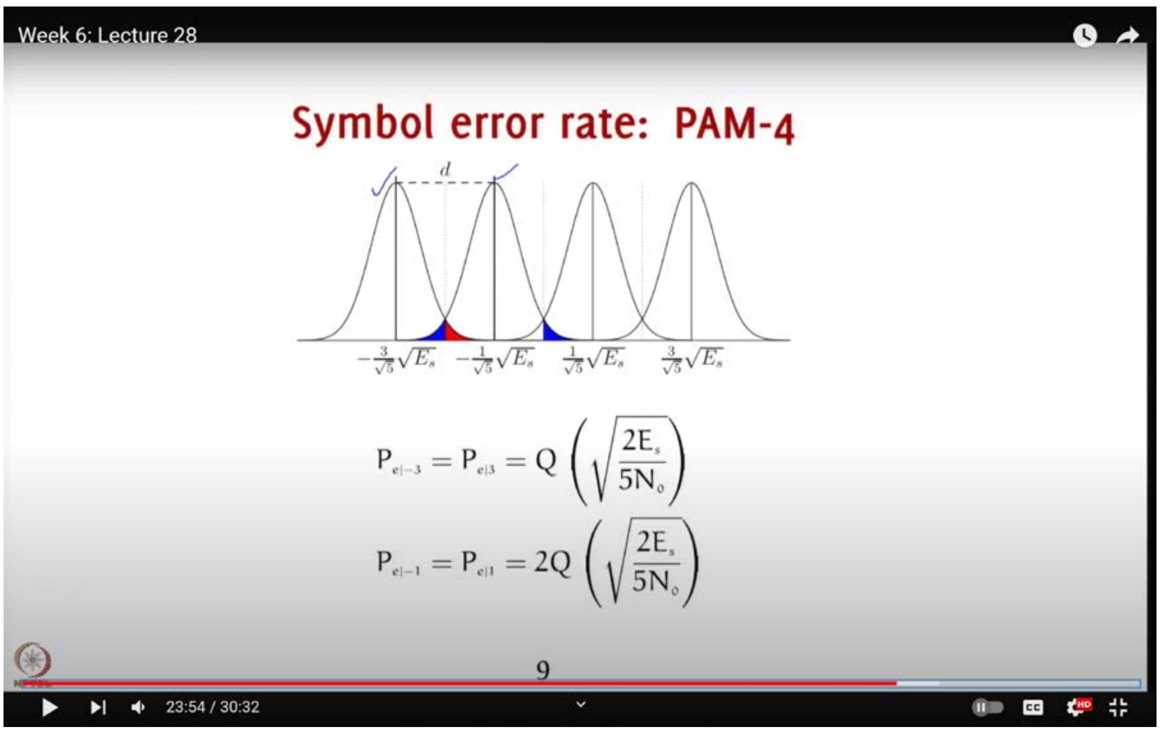
$$\frac{1}{4}[(-3\alpha)^2 + (-\alpha)^2 + \alpha^2 + (3\alpha)^2]$$

Substituting $\alpha = \sqrt{\frac{E_s}{5}}$, you get:

$$\frac{1}{4}[9\alpha^2 + \alpha^2 + \alpha^2 + 9\alpha^2] = \frac{20\alpha^2}{4} = 5\alpha^2$$

Since $\alpha^2 = \frac{E_s}{5}$, this simplifies to E_s, confirming that the constellation points are laid out correctly to meet the energy requirement.

(Refer Slide Time: 23:54)

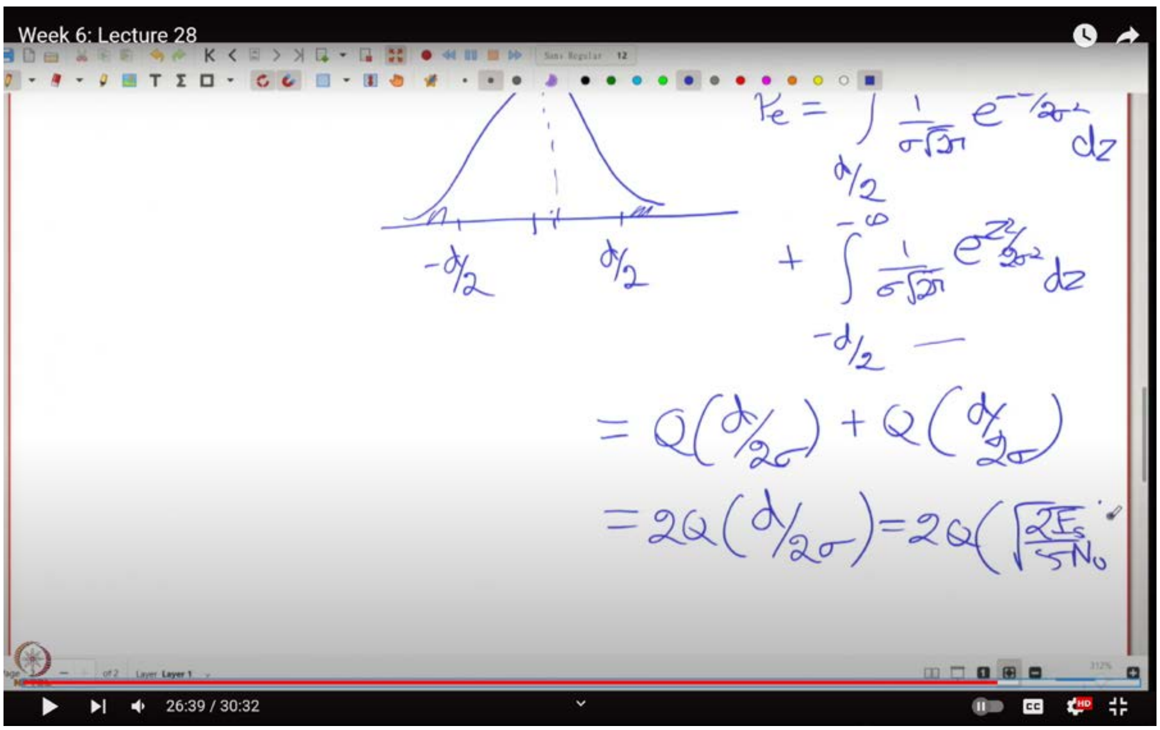


Next, let's calculate the bit error rate. The edge symbol, $-3\sqrt{\frac{E_s}{5}}$, encounters a symbol error when it crosses the decision boundary, which is straightforward. However, for the second and third symbols, each can cross two decision boundaries: one to the left and one to the right, making the calculation trickier.

For this PAM-4 constellation, let's consider the symbol $-3\sqrt{\frac{E_s}{5}}$ and $-\sqrt{\frac{E_s}{5}}$. The distance d between these symbols is:

$$d = 2 \times \sqrt{\frac{E_s}{5}}$$

(Refer Slide Time: 26:39)



To find the symbol error probability, use the Q-function:

$$P_e = Q\left(\frac{d}{2\sigma}\right)$$

Here, $\sigma^2 = \frac{N_0}{2}$, so:

$$P_e = Q\left(\frac{\sqrt{\frac{2E_s}{5}}}{\sqrt{N_0}}\right) = Q\left(\frac{\sqrt{2E_s}}{\sqrt{5N_0}}\right)$$

This calculation is straightforward for the edge symbol. For symbols near $-\sqrt{\frac{E_s}{5}}$ and $\sqrt{\frac{E_s}{5}}$, you need to account for errors crossing two decision boundaries. Although I've drawn the decision regions unsymmetrically for clarity, they are symmetric in practice. To handle these errors, consider the two different probabilities of error for each scenario.

When calculating the probability of error for a PAM-4 constellation, you must account for the different probabilities of error occurring due to crossing decision boundaries. These errors are mutually exclusive, so you can simply add the probabilities together. Here's how you can approach this problem from first principles:

Let's set up the problem: Suppose you have a Gaussian distribution centered around 0, and you're considering two key decision boundaries: $-\frac{d}{2}$ and $\frac{d}{2}$. To compute the symbol error probability, you need to evaluate the probability that the Gaussian noise causes a symbol to cross these boundaries.

The probability of symbol error in this scenario is given by:

$$\text{Probability of Error} = \int_{\frac{d}{2}}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z^2}{2\sigma^2}}\, dz + \int_{-\infty}^{-\frac{d}{2}} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z^2}{2\sigma^2}}\, dz$$

Here, the first integral covers the probability of crossing $\frac{d}{2}$, and the second integral covers the probability of crossing $-\frac{d}{2}$. Combining these, you get:

$$\text{Probability of Error} = 2\int_{\frac{d}{2}}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{z^2}{2\sigma^2}}\, dz$$

Recognize that this integral is just the Q-function, so:

$$\text{Probability of Error} = 2 \cdot Q\left(\frac{d}{2\sigma}\right)$$

Given that the distance d between constellation points is $\sqrt{\frac{2E_s}{5}}$ and the noise variance σ^2 is

$\frac{N_0}{2}$, you substitute these values into the Q-function formula:

$$\text{Probability of Error} = 2 \cdot Q\left(\frac{\sqrt{\frac{2E_s}{5}}}{\sqrt{\frac{N_0}{2}}}\right)$$

Thus:

$$\text{Probability of Error} = 2 \cdot Q\left(\frac{\sqrt{2E_s}}{5\sqrt{N_0}}\right)$$

Now, for the specific symbols $-3\sqrt{\frac{E_s}{5}}$ and $3\sqrt{\frac{E_s}{5}}$, each can cross one decision boundary, yielding an error probability of:

$$Q\left(\frac{\sqrt{2E_s}}{5\sqrt{N_0}}\right)$$

For the symbols $-\sqrt{\frac{E_s}{5}}$ and $\sqrt{\frac{E_s}{5}}$, there are two potential errors because they can cross two decision boundaries. Both scenarios have the same error probability:

$$Q\left(\frac{\sqrt{2E_s}}{5\sqrt{N_0}}\right)$$

Thus, the total symbol error probability for this constellation is:

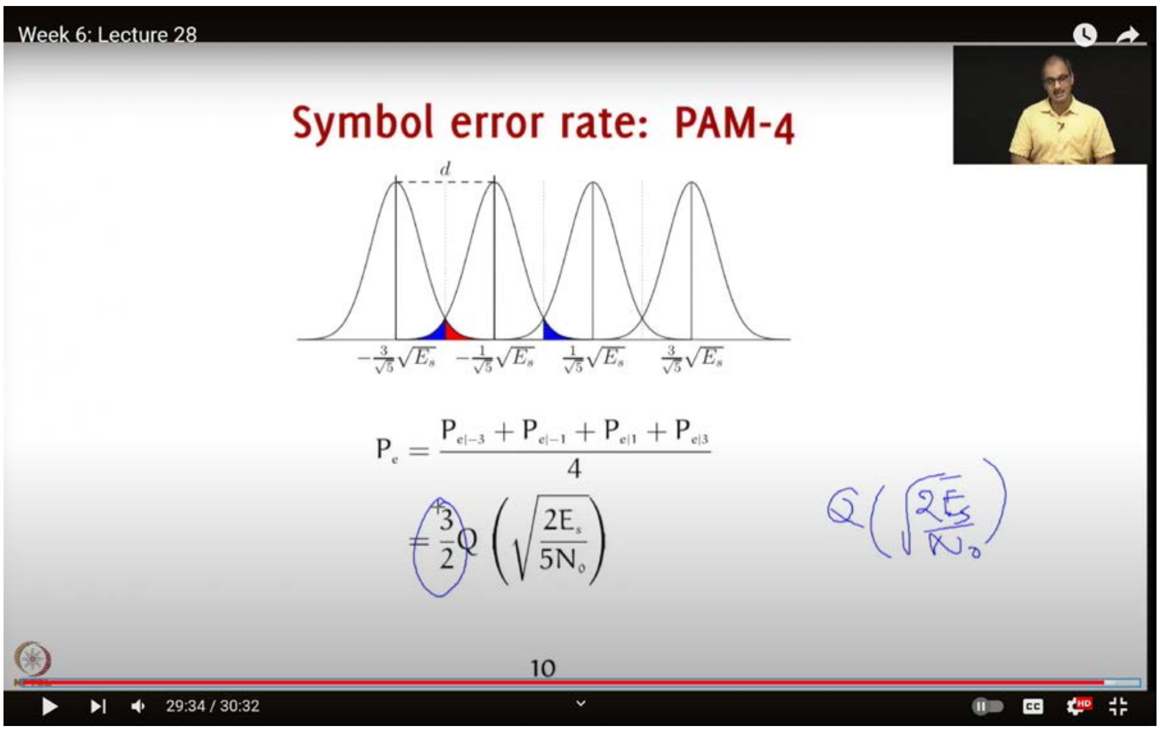
$$\text{Probability of Error} = 2 \cdot Q\left(\frac{\sqrt{2E_s}}{5\sqrt{N_0}}\right)$$

Therefore, the probability of symbol error is not uniform across all constellation points. Some points may have higher or lower symbol error rates depending on their positions and the distance between them.

To determine the average symbol error rate, you need to multiply the probability of error for each constellation point by the probability of that constellation point occurring. In our case, since all constellation points are equally likely, the average symbol error rate can be calculated as follows:

$$\text{Average Symbol Error Rate} = \frac{1}{4}Q\left(\frac{\sqrt{2E_s}}{\sqrt{5N_0}}\right) + \frac{1}{4}Q\left(\frac{\sqrt{2E_s}}{\sqrt{5N_0}}\right) + \frac{1}{2}Q\left(\frac{\sqrt{2E_s}}{\sqrt{5N_0}}\right)$$

(Refer Slide Time: 29:34)



After performing the computation, you will find that the average symbol error rate is:

$$\frac{3}{2}Q\left(\frac{\sqrt{2E_s}}{\sqrt{5N_0}}\right)$$

It’s not surprising to see that the probability of symbol error for PAM-4, given the same energy E_s and noise variance N_0, is significantly higher compared to BPSK. This higher error rate is due to the fact that, within the same energy budget, PAM-4 uses more symbols. Thus, there is a greater chance of error when the signal-to-noise ratio (SNR) is the same. For BPSK, the SNR is $\frac{\sqrt{2E_s}}{N_0}$, and with PAM-4, the increased number of symbols (specifically, a factor of $\frac{2}{5}$) leads to a larger integral and a higher error rate. Consequently, the symbol error rate for PAM-4 is notably higher than for BPSK.

As you move to more complex constellations like PAM-8 or PAM-16, the symbol error rates will further increase, and such constellations will only perform well under very high SNR conditions.

In this lecture, we've examined the symbol error rate probabilities for BPSK and PAM-4, computed the energy of constellations, and understood how to draw decision boundaries correctly for accurate symbol error rate calculations. In the next lecture, we will extend these concepts to quadrature amplitude modulation (QAM) and other constellations, and explore bit error rates as well. Thank you.