The overall output noise power is,

$$\sigma_{err}^2 = \sigma_{o1}^2 + \sigma_{o2}^2$$

$$= \sigma_e^2(89.80 + 2.778) = 92.578\sigma_e^2$$

Substituting $\sigma_e^2 = \frac{2^{-2B}}{12}$, we get

$$\sigma_{err}^2 = 92.578 \times \frac{2^{-2B}}{12} = 7.715 \times 2^{-2B}$$

Two Mark questions and Answers:

1. What do finite word length effects mean?

The effects due to finite precision representation of numbers in a digital system are called finite word length effects.

2.what are the three-quantization errors to finite word length registers in digital filters?

1. Input quantization error
2. Coefficient quantization error
3. Product quantization error

3. What is input quantization error?.

The filter coefficients are computed to infinite precision in theory. But in digital computation the filter coefficients are represented in binary and are stored in registers. If a b bit register is used the filter coefficients must be rounded or truncated to b bits, which produces an error.

4 .What is product quantization error?.

The product quantization errors arise at the out put of the multiplier. Multiplication of a b bit data with a b bit coefficient results a product having 2b bits. Since a b bit register is used the multiplier output will be rounded or truncated to b bits which produces the error.

5.what are the different types of arithmetic in digital systems.?

There are three types of arithmetic used in digital systems. They are fixed point arithmetic, floating point ,block floating point arithmetic.

6.What do you understand by input quantization error?

In digital signal processing, the continuous time input signals are converted into digital using a b-bit ACD. The representation of continuous signal amplitude by a fixed digit produce an error, which is known as input quantization error.

7. List some of the finite word length effects in digital filters?

1. Errors due to quantization of input data.
2. Errors due to quantization of filter co-efficient
3. Errors due to rounding the product in multiplications

4. Limit cycles due to product quantization and overflow in addition.

8.what do you understand by a fixed-point number?

In fixed point arithmetic the position of the binary point is fixed. The bit to the right represents the fractional part of the number & those to the left represent the integer part. For example, the binary number 01.1100 has the value 1.75 in decimal.

9. What are the different formats of fixed-point representation?

a. Sign magnitude format
b. One's Complement format
c. Two's Complement format.

In all the three formats, the positive number is same but they differ only in representing negative numbers.

10. What is meant by sign magnitude representation?

For sign magnitude representation the leading binary digit is used to represent the sign. If it is equal to 1 the number is negative, otherwise it is positive.

11. What is meant by 1's complement form?

In 1,s complement form the positive number is represented as in the sign magnitude form. To obtain the negative of the positive number ,complement all the bits of the positive number.

12. What is meant by 2's complement form?

In 2's complement form the positive number is represented as in the sign magnitude form. To obtain the negative of the positive number ,complement all the bits of the positive number and add 1 to the LSB

13. Explain the floating-point representation of binary number.

The floating-point number will have a mantissa part. In a given word size the bits allotted for mantissa and exponent are fixed. The mantissa is used to represent a binary fraction number and the exponent is a positive or negative binary integer. The value of the exponent can be adjusted to move the position of binary point in mantissa. Hence this representation is called floating point.

14. What is meant by floating point representation?

In floating point form the positive number is represented as F =2CM,where is mantissa, is a fraction such that$1/2<M<1$and C the exponent can be either positive or negative.

15. What are the advantages of floating pint representation?

- Large dynamic range
- overflow is unlikely.

16.what are the advantages of floating point arithmetic?

1. Large dynamic range
2. Over flow in floating point representation is unlike.

17. What are the types of arithmetic used in digital computers?

The floating point arithmetic and two's complement arithmetic are the two types of arithmetic employed in digital systems.

18. What are the two types of quantization employed in digital system?

The two types of quantization in digital system are Truncation and Rounding.

19.what is the relationship between truncation error e and the bits b for representing a decimal into binary?

For a 2's complement representation, the error due to truncation for both positive and negative values of x is $0>x_T-x>-2^{-b}$.Where b is the number of bits and xt is the truncated value of x.The equation holds good for both sign magnitude, 1's complement if x>0. If x<0, then for sign magnitude and for 1's complement the truncation error satisfies.

20. What is truncation?

The truncation is the process of reducing the size of binary number by discarding all bits less significant than the least significant bit that is retained. In truncation of a binary number of b bits all the less significant bits beyond bth bit are discarded.

21. What is rounding?

Rounding is the process of reducing the size of a binary number to finite word sizes of b-bits such that, the rounded b-bit number is closest to the original unquantized number.

22. Explain the process of upward rounding?

In upward rounding of a number of b-bits, first the number is truncated to b-bits by retaining the most significant b-bits. If the bit next to the least significant bit that is retained is zero, then zero is added to the least significant bit of the truncated number. If the bit next to the least significant bit that is retained is one then one is added to the least significant bit of the truncated number.

23.what is meant by A/D conversion noise?

A DSP contains a device, A/D converter that operates on the analog input x(t) to produce xq(t) which is binary sequence of 0s and 1s.At first the signal x(t) is sampled at regular intervals to produce a sequence x(n) is of infinite precision. Each sample x(n) is expressed in terms of a finite number of bits given the sequence xq(n). The difference signal e(n)=xq(n)-x(n) is called A/D conversion noise.

24.what is the effect of quantization on pole location?

Quantization of coefficients in digital filters lead to slight changes in their value. These changes in value of filter coefficients modify the pole-zero locations. Sometimes the pole locations will be changed in such a way that the system may drive into instability.

25.which realization is less sensitive to the process of quantization?

Cascade form.

26.what is meant by quantization step size?

Let as assume a sinusoidal signal varying between +1 and -1 having a dynamic range 2.If ADC is used to convert the sinusoidal signal employs (b+1) bits including sign bit, the number of levels available for quantizing x(n) is 2^{b+1}.Thus the interval between successive level is $q = \frac{2}{2^{b+1}} = 2^{-b}$ .Where q is known as quantization step size.

27. What are the errors generated by A/D process?

The A/D process generates two types of errors. They are quantization error and saturation error. The quantization error is due to representation of the sampled signal by a fixed number of digital levels. The saturation errors occur when the analog signal exceeds the dynamic range of A/D converter.

28. What is quantization step size?

In digital systems, the numbers are represented in binary. With b-bit binary wecan generate 2b different binary codes. Any range of analog value to be represented in binary should be divided into 2b levels with equal increment. The 2b levels are called quantization levels and the increment in each level is called quantization step size. If R is the range of analog signal then,Quantization step size, q = R/2b

29. Why errors are created in A/D process?

In A/D process the analog signals are sampled and converted to binary. The sampled analog signal will have infinite precision. In binary representation of bbits we have different values with finite precision. The binary values are called quantization levels. Hence the samples of analog are quantized in order to fit into any one of the quantized levels. This quantization process introduces errors in the signal.

30. What is steady state output noise power due to input quantization?

The input signal to digital system can be considered as a sum of unquantized signal and error signal due to input quantization. The response of the system can be expressed as a summation of response due to unquantized input and error signal. The response of the system due to error signal is given by convolution of error signal and impulse response. The variance of response of the system for error signal is called state output noise power.

31. What is meant by coefficient inaccuracy?

In digital computation the filter coefficients are represented in binary. With b-bit binary we can generate only 2b different binary numbers and they are called quantization levels. Any filter coefficient has to be fitted into any one of the quantizat6ion levels. Hence the filter coefficients are quantized to represent in binary and the quantization introduces errors in filter coefficients. Therefore the coefficients cannot be accurately represented in a digital system and this problem is referred to as coefficient inaccuracy.

32. How the digital filter is affected by quantization of filter coefficients?

The quantization of the filter coefficients will modify the value of poles & zeros and so the location of poles and zeros will be shifted from the desired location. This will create deviations in the frequency response of the system. Hence the resultant filter will have a frequency response different from that of the filter with unquantized coefficients.

33. How the sensitivity of frequency response to quantization of filter coefficients is minimized?

The sensitivity of the filter frequency response to quantization of the filter coefficients is minimized by realizing the filter having a large number of poles and zeros as an interconnection of second order sections. Hence the filter can be realized in cascade or parallel form, in which the basic buildings blocks are first order and second order sections.

34. Why rounding is preferred for quantizing the product?

In digital system rounding due to the following desirable characteristic of rounding performs the product quantization

1. The rounding error is independent of the type of arithmetic
2. The mean value of rounding error signal is zero.
3. The variance of the rounding error signal is least.

35. Define noise transfer function (NTF)?

The Noise Transfer Function is defined as the transfer function from the noise source to the filter output. The NTF depends on the structure of the digital networks.

36. What are the assumptions made regarding the statistical independence of the various noise sources in the digital filter?

The assumptions made regarding the statistical independence of the noise sources are,

1. Any two different samples from the same noise source are uncorrelated.
2. Any two different noise source, when considered, as random processes are uncorrelated.
3. Each noise source is uncorrelated with the input sequence.

37. What are limit cycles?

In recursive systems when the input is zero or some nonzero constant value, the nonlinearities die to finite precision arithmetic operations may cause periodic oscillations in the output. These oscillations are called limit cycles.

38. What are the two types of limit cycles?

The two types of limit cycles are zero input limit cycles and overflow limit cycles.

39. What is zero input limit cycles?

In recursive system, the product quantization may create periodic oscillations in the output. These oscillations are called limit cycles. If the system output enters a limit cycles, it will continue to remain in limit cycles even when the input is made zero. Hence these limit cycles are also called zero input limit cycles.

40.Determine "dead band" of the filter

The limit cycle occur as a result of quantization effect in multiplication. The amplitudes of the output during a limit cycle are confined to a range of values called the dead band of the filter.

41. How the system output cam be brought out of limit cycles?

The system output can be brought out of limit cycle by applying an input of largeV magnitude, which is sufficient to drive the system out of limit cycle.

42. What is saturation arithmetic?

In saturation arithmetic when the result of an arithmetic operation exceeds the dynamic range of number system, then the result is set to maximum or minimum possible value. If the upper limit is exceeded then the result is set to maximum possible value. If the lower limit is exceeded then the r4esult is set to minimum possible value.

43. What is overflow limit cycle?

In fixed point addition the overflow occurs when the sum exceeds the finite word length of the register used to store the sum. The overflow in addition may lead to oscillations in the output which is called overflow limit cycles.

44. How overflow limit cycles can be eliminated?

The overflow limit cycles can be eliminated either by using saturation arithmetic or by scaling the input signal to the adder.

45. What is the drawback in saturation arithmetic?

The saturation arithmetic introduces nonlinearity in the adder which creates signal distortion.

46.Explain briefly the need for scaling in the digital filter implementation.

To prevent overflow, the signal level at certain points in the digital filter must be scaled so that no overflow occurs in the adder.

5.8.Summary:

When working with any kind of digital electronics in which numbers are being represented, it is important to understand the different ways numbers are represented in these systems. Almost without fail, numbers are represented by two voltage levels which can represent a one or a zero. The number system based on ones and zeroes is called the binary system (because there are only two possible digits). Before discussing the binary system, a review of the decimal (ten possible digits) system is in order, because many of the concepts of the binary system will be easier to understand when introduced alongside their decimal counterpart.