

action taken, the solid lines indicate information flow, and the dotted lines indicate control flow. The intent is to use a DSP chip to implement all the boxes enclosed in the large dashed box. All functions inside the DSP, except the A/D, are to be performed in software. The A/D function, which is a hardware function, may have to be a separate chip, depending on the resolution required. The functions of the 11 blocks are described below.

Electroacoustic Block: The electroacoustic (microphone) block has two functions: To convert, in a linear fashion, acoustic airwaves to electric signals, and to low-pass filter the signal to prevent aliasing when the signal is sampled in the subsequent A/D block.

The first function is that of a reasonably sensitive microphone. The acoustic airwave input will have an intensity (measured relative to the hearing threshold of 10^{-12} W/m^2) in the range of 70 to 84 dB when the signal is present. It contains frequency components that may be as low as 70 and as high as 10,000 Hz, although the high-frequency components will be very weak. The information in the signal is at the fundamental frequency of each string, which means it is in the frequency range between 70 and 350 Hz. The fraction of the total signal power in this band will vary from string to string and also depend on where the string was plucked (plucking near the bridge emphasizes the harmonics). From looking at the input waveforms and their DFTs, it is concluded that this fraction ranges from 0.1 to 1.0.

The other function of the electroacoustic block is to remove the high-frequency components in the signal to prevent aliasing when the signal is sampled. The filter should pass components of the signals in the band from 70 to 350 Hz. The pass-band ripple is relatively unimportant since the signal of interest is a single-frequency sinusoid as opposed to a broad-band signal in which the relative amplitudes at different frequencies are important. However, ripple in the pass band causes the output to have a larger dynamic range than the input, which is undesirable. It is therefore prudent to keep the pass-band ripple relatively small. The pass-band ripple is set at ± 1.5 dB, as this is easily achieved and yet extends the dynamic range of the signal by only 3 dB. The phase performance is unimportant and therefore no phase specification is given. The stop band is 600 to 10,000 Hz. For the input signal in question it was determined that aliasing is negligible if the gain of the stop band is set 20 dB below the reference gain of the filter (i.e., the average gain of the pass band). This is 18.5 dB below the bottom rail of the pass band. The performance limits of this filter are specified by the template in Figure A.11.

The output of this block feeds an A/D converter. A common input voltage range for A/D converters, especially A/Ds that are inside DSP chips, is 0 to 5 volts. Therefore the output for the electroacoustic block is specified as a voltage between 0 and 5 volts. The acoustic signal from the guitar is an ac signal, which means a dc bias voltage will have to be inserted to center the ac signal in the voltage range specified. Obviously this bias voltage should be 2.5 volts. There must be some tolerance of error in this bias to allow for

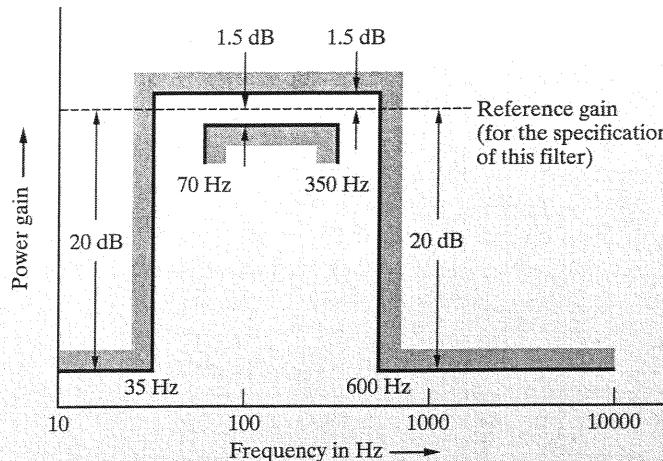


FIGURE A.11 The Template for the Anti-Aliasing Filter.

variations in component values in the bias circuit. The tolerance is set at $\pm 10\%$. This should allow a simple implementation of the bias circuit. If the systems analysis shows that the performance can be improved with a tighter tolerance, then this decision should be reevaluated.

The gain from input to output of the electroacoustic block must also be specified. This gain controls the signal level presented to the A/D block. It can be separated into two components, the reference (or average pass-band) gain and the frequency-dependent filter gain (or ripple). The reference gain should be such that the largest input signal does not saturate the output. This gain will be specified as a power gain and will be denoted G_r . Since the block in question converts acoustic intensity into an electric potential, this gain has units of volts squared per watts per meter squared (i.e., $V^2/(W/m^2)$). The reference gain, G_r , should be such that the maximum input sound level, which is 84 dB, does not saturate the output. The worst-case condition for this is for the signal to experience the maximum possible filter gain, which is 1.5 dB, and the maximum possible dc bias voltage, which is $2.5 + 10\% = 2.75$ volts. The other factor involved in the worst-case analysis is one that relates the peak signal power to average signal power. For a pure sinusoid this factor is 2. From observation of the guitar string data, a factor of 4 is more appropriate for this application. This peaking factor is denoted F_p and relates the average signal power to maximum instantaneous power (i.e., $P_{peak} = F_p P_{ave}$). The maximum possible reference gain, denoted G_{rmax} , can be found by equating the output voltage of the electroacoustic block to the maximum that the A/D block can handle. This equation is

$$G_{rmax} \times 10^{0.15} \times 10^{8.4} \times 10^{-12} W/m^2 = \frac{(5V - 2.75V)^2}{F_p}, \quad (A.1)$$

where $10^{0.15}$ is the ripple gain and the product $10^{8.4} \times 10^{-12}$ is the maximum possible signal power. Solving yields a maximum reference gain of

$$G_{max} = 3567 \frac{V^2}{W/m^2}. \quad (A.2)$$

The specification of the reference gain should make allowance for unit-to-unit variation in the "conversion gain" of the microphone element. Allowing for a unit-to-unit variation of $\pm 10\%$, the reference gain is specified at

$$G_r = G_{max}/1.1 = 3240 \frac{V^2}{W/m^2} \pm 10\%. \quad (A.3)$$

This specification could well be unrealistic. More information on this topic will surface when the electroacoustic conversion block is designed. The system specification may have to be revised at that time and another system analysis may need to be performed.

The output impedance is relatively unimportant, but, in combination with the input impedance of the next block, will act as a voltage divider and therefore must be specified. The output impedance is specified to be less than 10 ohms for frequencies from 0 to 10,000 Hz. The value of 10 ohms is sufficiently small that voltage division with the input impedance of the next stage should be negligible.

A/D Converter: The A/D block has four main parameters that must be specified. The first, which is not very important, is input impedance. The second is the input voltage range, which has been discussed in the electroacoustic block specification. The third is the sampling rate and the fourth is the quantizing resolution (i.e., the number of bits per sample). The latter two parameters are quite important, since they affect whether the A/D can be part of the DSP chip or must be a separate IC. They also have an impact on the processing power and memory requirements of the DSP chip. These parameters are discussed and specified below.

The input impedance of this block is not critical as long as it is large enough to ensure that the voltage-divider effect from the previous stage is negligible. It is specified as greater than 10 kohms for frequencies in the range of 35 to 600 Hz.

The input voltage range of the A/D has to match the output of the electroacoustic block. For that reason the input voltage range of the A/D was set in the specification of the electroacoustic block. It is specified to be 0 to 5 volts.

The sampling rate is an important parameter and care must be taken to specify it properly. The sampling rate must be greater than the Nyquist rate but not so large that the DSP chip does not have time to process a sample before the next one arrives. The Nyquist theorem dictates that the input to the A/D converter should be sampled at a rate at least twice the bandwidth of the anti-aliasing filter. The anti-aliasing filter rejects frequencies from 600 Hz

and up, which means the sampling rate should be at least 1200 Hz. Since the anti-aliasing filter is not a brick-wall filter, it would be safer to oversample. For this reason the sampling rate is specified at 2000 samples per second. This decision should be reviewed after the processing power of the DSP chip has been evaluated.

The A/D block samples and quantizes the analog signal. This produces a digital signal represented by a sequence of numbers. The resolution of the A/D is measured by the number of bits used to represent a sample and affects the type and size of the data structures used in the signal-processing algorithms. It also affects the cost of the A/D converter. At this point one could specify the resolution without any calculation. In essence one could just take a guess and choose 12 bits, say, for the resolution. Whether or not 12 bits is a good choice will be revealed later, in the system analysis. Since it is clear that the resolution of the A/D converter has a significant impact on the system, the resolution analysis is performed at this point.

The resolution required of the A/D converter can be roughly established with the criterion that the quantization noise be less than the main source of noise, which is the background room noise. If the quantizer noise power is at least 3 dB less than the background noise power, it will contribute at most 1.8 dB to the total noise, which is not significant. Thus the resolution of the quantizer can be roughly established by setting the quantizer noise power to half the value of the room noise power.

According to the requirements specification, the background acoustic noise can be as high as 60 dB. This is 24 dB lower than the maximum possible signal power. The noise could, with bad luck, experience the maximum possible filter gain, which is 1.5 dB. Since this was the filter gain used in the worst-case analysis for maximum possible signal power, the background noise power at the output of the electroacoustic block will be 24 dB below the maximum possible signal power. Therefore the maximum possible background noise power is

$$P_{nmax} = \frac{(5V - 2.75V)^2}{F_p} \times 10^{-2.4} = 0.0050 \text{ V}^2. \quad (\text{A.4})$$

The quantization noise power is equal to the square of the resolution divided by 12. The resolution is the size of analog intervals that map to the same number. For an N -bit quantizer the resolution, which is also called step size is $5/2^N$ V. The quantizer noise power is therefore

$$P_q = \frac{1}{12} \times \left(\frac{5}{2^N} \right)^2 \text{ V}^2. \quad (\text{A.5})$$

Setting $P_q = P_{nmax}/2$ and solving for N yields $N = 4.85$, which must be rounded to the nearest integer. This results in $N = 5$.

The resolution of the A/D converter can now be specified. Since DSP chips are byte-oriented and many come with built-in 8-bit A/D converters, the

resolution of the A/D is specified to be 8 bits. This gives a 9 dB safety margin in the resolution of the A/D converter.

Power-Threshold Detector: This block is invoked on interrupt from the A/D converter when the interrupt is enabled. The interrupt is enabled by a power-up reset and also by either the fundamental-detector block or the LED-sequence-generator block.

The function of this block is to calculate the power in the received data stream by using a sliding average over a short period of time. If this power exceeds the minimum possible signal power, the power-threshold detector disables its interrupt and enables the store-data block interrupt. This passes control to the data-collection process.

There are two parameters that have to be specified: the length of the moving average window and the threshold for the power of a plucked string. The window length for the moving-average power calculation must be only a fraction of the time a plucked string generates significant sound. This is because the decision from the threshold detector must be made and control passed to the store-data block while the sound is still present. As seen from the graphs in Section A.6, persistence depends on the string plucked. All notes persist for at least 1/2 second. Obviously, the specification of the window length is subjective. An optimum value could only be found experimentally, and that would take considerable time. Instead, a window length of 100 milliseconds is chosen based on engineering judgment. This is long enough to calculate a reasonable average for the power, yet short enough to leave at least 1/2 second of signal for the store-data block to collect.

The specification of the power threshold, which is the lowest possible power generated by plucking a string, is less subjective. It can be calculated from signal levels given in the requirements specification. From the requirements specification it is known that the total acoustic input could be as low as 70 dB. It is also known that as little as 0.1 of this power could be in the frequency component of interest. Since we can be sure only that the fundamental frequency will pass the anti-aliasing filter, the worst-case power level is $70 - 10 = 60$ dB, which is 24 dB below the maximum input power level. The lowest possible output power occurs if the reference gain is at its lower limit and the signal experiences the minimum pass-band filter gain. Since the lower limit of the pass-band gain is 3 dB below the upper limit and since the lowest possible reference gain is 0.83 dB lower than the maximum possible reference gain, the smallest possible power is $24 + 3 + .83$ dB down from the largest possible signal. Thus the output power level for the weakest possible signal is

$$P_{\text{threshold}} = (5 \text{ V} - 2.75 \text{ V})^2 / F_p / 10^{2.783} = 0.0021 \text{ V}^2, \quad (\text{A.6})$$

where $P_{\text{threshold}}$ is the threshold used to make the decision whether or not a string was plucked. In terms of voltage, this threshold is 45.7 mV RMS.

Store Data: This software module is made active and inactive under control of an interrupt enable. When active, this module is invoked by an interrupt from the A/D block, which occurs on every sample. Its interrupt is enabled by the power-threshold detector after the decision has been made to start the data collection. It deactivates itself after it has collected a 1/2 second interval of signal (1000 samples) by disabling its interrupt. It then passes control to the DFT algorithm program with a jump instruction. The DFT algorithm program is located at address *DFT_algorithm*.

The data is stored in 1000 bytes of contiguous memory starting at address *data_buffer_address*. The data is stored in chronological order with the oldest located at the lowest address.

DFT Computation: The DFT block is implemented as a subprogram. It is activated by a jump instruction executed in the store-data block. After completion, it passes control to the valid-note-detector block with a jump instruction.

The DFT block computes a 1000-point DFT on the 1000 samples of signal collected by the store-data block. The DFT produces 1000 complex numbers that represent the amplitude and phase of the Fourier components. Because the input is a real signal, only the first 500 complex numbers have meaning. The k th complex number, say the number $A_k e^{j\theta_k}$, represents a sinusoidal function of time with period $2\pi k/T$, amplitude A_k , and phase θ_k , i.e. it represents the Fourier component

$$x(t) = A_k \cos\left(\frac{2\pi kt}{T} + \theta_k\right), \quad (\text{A.7})$$

where T (specified in the store-data block as 1/2 second) is the observation interval (time interval spanned by 1000 samples). Only the magnitudes, i.e. A_k s, are important, so only the first 500 magnitudes are stored.

The only parameters that have to be specified in the DFT block are the length of the word used to represent the magnitudes of the DFT and the address of the data structure where these magnitudes are stored.

The word length used to represent the amplitude of the sinusoids is discussed first. Since each component in the DFT is a weighted sum of the data, where the magnitude of the weighting is between 0 and 1, the maximum possible value that could be obtained for a byte-sized input is $256 \times 1000 = 256,000$. This would require an 18-bit word, which would require three bytes of memory. The maximum possible practical value for a DFT component from a guitar-string signal (this excludes the dc component due to the bias) is well below this. The result can be safely fit into 16 bits, which is a two-byte word. This provides far more resolution than is needed and the result could possibly be truncated to fit into eight bits. To be on the safe side, the word length for the magnitude of the DFT is specified to be 16 bits. This decision should be reviewed after the system analysis.

The magnitude of the first 500 of the 1000 frequency components will be stored in a buffer with address *DFT_results*. The first component, which is the

DC component of the DFT, is stored at address *DFT_results*. Each word is two bytes long with the most significant byte stored at the lower address. Therefore, the second component, which corresponds to frequency $1/T = 2 \text{ Hz}$, is stored starting at address *DFT_results* + 2.

Valid-Note Detector and Fundamental-Frequency Estimator:

The valid-note-detector block is implemented as a subprogram. It is activated by a jump instruction executed in the DFT block. After completion, it passes control to either the frequency-estimator block or the power-threshold-detector block. The way control is passed depends on where it is passed. If it is passed to the frequency-estimator block, it is passed by a jump instruction. If it is passed to the power-threshold-detector block, it is passed by enabling an interrupt.

When the frequency-estimator block gets control, it executes and then passes control to the LED-sequence-generator block with a jump instruction. The valid-note-detector block decides whether or not the signal collected is that of a plucked string. It does this by comparing the power of the strongest Fourier component to the total power (excluding the dc component) in the signal. If this ratio is greater than some threshold, the signal collected is declared to be that of a plucked guitar string. The graphs of the DFTs of the sounds from all six strings show that the strongest Fourier component has at least 1/20 of the total power in the signal. Therefore the power threshold for a valid plucking is set at 1/20.

If a valid note is detected, control is passed to the frequency-estimator block. This block estimates the frequency of the string plucked as follows. The buffer is searched to find the largest component. Then the amplitude of the Fourier component at half this frequency is compared to that of the largest component to see if a subharmonic exists. If the amplitude of the subharmonic is at least one third that of the largest component, the subharmonic is declared to be the fundamental frequency of the string plucked. Otherwise the largest component is taken to be the fundamental of the string plucked. The frequency of a Fourier component is determined from its position in the buffer. The component in the k th word has frequency $k/T = 2k$.

The result is stored as two's complement in a two-byte word at addresses *measured_fundamental_frequency*, and *measured_fundamental_frequency* + 1. It is stored as an integer in two's complement with the most significant byte at the lower address. The units are $1/T \text{ Hz}$, where T is the observation interval. In this case $T = 1/2$ seconds and the units are 2 Hz , i.e. an integer value of 128 indicates the fundamental is at 256 Hz. After the fundamental frequency is estimated and stored, control is passed to the LED-sequence generator with a jump instruction.

LED Sequence Generator and Display: This block compares the estimated fundamental frequency to the correct fundamental frequency for each of the six strings. On the basis of this comparison it sets the sharp and flat

LEDs for each of the six strings. If the frequency corresponds to “in-tune” for a particular string, then both the flat and sharp LEDs are activated to indicate “in-tune.” The outputs of this block are 12 logic lines called S1F (string 1 flat), S1S (string 1 sharp), S2F, S2S, . . . , S6S. A logic high indicates the LED is on.

The LEDs are activated for two seconds and then turned off. Immediately after turning off the LEDs, control is passed to the threshold detector. This is done by enabling the interrupt for the threshold detector and then executing a wait instruction.

Power-Regulator and Power-up-Reset Blocks: The power-regulator block must provide the electroacoustic block with a DC voltage of between 7 and 9 volts for a load current in the range 1 to 50 mA. The power supply must stabilize between 7 and 9 volts within 250 milliseconds of first reaching 7 volts. The total AC voltage, which includes frequencies from 1 Hz to 10 MHz, must not exceed 0.20 volts RMS.

The DSP chip is powered with 5 volts. The regulator must supply 5 V DC ± 0.25 V for a load current between 10 and 200 mA. The power supply must stabilize between 4.75 and 5.25 V within 250 ms after reaching 4.5 V. The total AC noise voltage must be limited to less than 0.25 V peak.

The power-up-reset line, which is normally less than 0.5 volts, must generate a pulse of 3.4 to 5 volts for 250 to 500 milliseconds with the leading edge occurring when the 5 volt power supply reaches 4.75 volts.

A.4.8 Job 6: Analyze the System

After specifying the blocks in the block diagram, Sarah analyzed the system. She has three areas of concern. The first is the effect of background noise on the performance of the threshold detector. What if the background noise was greater than the threshold? The second concern is the accuracy of the frequency estimation. The third concern is the way control was passed from block to block. She felt it may not work the way she specified.

To address her first concern, she calculated the power in the background noise at the input to the threshold-detector block. Her logic was this: The background acoustic noise is specified in the requirements specification to be at most 60 dB above the hearing threshold of 10^{-12} W/m². In going through the electroacoustic block, this noise would experience the same reference gain as the signal but could experience a different filter gain. An unlucky situation would have the background noise experiencing the maximum filter gain of +1.5 dB. This would make noise power at the output of the electroacoustic block 6 dB above the minimum signal power level.

At this point Sarah knows the power-threshold detector will not work as she specified it. She must either synthesize a new system based on a different concept, revise the current system, or see if the noise level specified in the requirements specification can be relaxed. She chose to explore the third