Digital hearing aids: A tutorial review*

HARRY LEVITT, Ph.D.
Center for Research in Speech and Hearing Sciences, Graduate School and University Center of the City University of New York, New York, NY 10036

Abstract—The basic concepts underlying digital signal processing are reviewed briefly, followed by a short historical account of the development of digital hearing aids. Key problem areas and opportunities are identified, and the various approaches used in attempting to develop a practical digital hearing aid are discussed.

ADVANTAGES OF DIGITAL HEARING AIDS

Digital hearing aids promise many advantages over conventional hearing aids. These include:

- Programmability;
- Much greater precision in adjusting electroacoustic parameters;
- Self-monitoring capabilities;
- Logical operations for self-testing and self-calibration;
- Control of acoustic feedback (a serious practical problem with high-gain hearing aids);
- The use of advanced signal-processing techniques for noise reduction;
- Automatic control of signal levels; and
- Self-adaptive adjustment to changing acoustic environments.

Only a few of these features are likely to be included in the first generation of wearable digital hearing aids because of the constraints on chip size and power consumption. But all of the above-mentioned features have already been demonstrated in a master digital hearing aid (12,14) and it is likely that (with the development of more advanced signal-processing chips) an increasing number of these attractive features will be incorporated in the wearable digital hearing aids of the future.

The many advantages offered by digital hearing aids can be subdivided into three broad groups:

1) Signal-processing capabilities that are analogous to, but superior to, those offered by conventional analog hearing aids.

2) Signal-processing capabilities that are unique to digital systems and which cannot be implemented in conventional analog hearing aids.

3) Methods of processing and controlling signals that change our way of thinking about how hearing aids should be designed, prescribed, and fitted.

The third type of advantage is the most subtle—and the most important. For example, a digital hearing aid can be programmed not only to amplify, but also to generate, audio signals. As such, the instrument can be programmed to serve as an audiometer in order to facilitate the measurement of audiological characteristics relevant to the prescriptive fitting of hearing aids (14,15). Using this approach, it is possible to circumvent the very difficult problem of correcting for the frequency-dependent differences in sound level between the traditional audiometer headphone and the patient's own hearing-aid receiver. The idea of using a hearing
aid as an audiometer was born out of the realization that the primary differences between the two devices are those of software (i.e., the controlling program) rather than fundamental differences in hardware.

Another important stimulus to our thinking has been the use of computer simulation. As noted elsewhere in this paper, the development of computer-simulation techniques to facilitate the design and development of vocoders and other telephone-oriented speech processing systems led to the realization that computers could also be used to simulate hearing aids, and that eventually digital processing of audio signals would be possible in a hearing aid. Further, the first working digital hearing aid was achieved using real-time computer simulation (12,14), thereby providing a glimpse of the many possible features that could be incorporated in the digital hearing aids of the future.

The results currently being obtained using computer simulation techniques are having a profound effect on our thinking with respect to which features should or should not be included in a modern hearing aid. Whereas, in the past, progress in the development of more effective hearing aids was limited primarily by what could be achieved technologically, the present situation, with the help of computer simulation and the realization of practical digital hearing aids, is one in which progress is limited primarily by our own lack of understanding of what is important in processing signals for hearing impairment.

HOW DOES DIGITAL SIGNAL PROCESSING WORK?

Analog-to Digital Conversion

The key element in a digital hearing aid is its use of sound signals that have been sampled discretely in time, so that they have become represented by a series of data points instead of by a continuously-varying value analogous to the waveform itself. Figure 1 illustrates the process whereby an analog signal is converted to digital form. The uppermost section of the diagram shows a continuous waveform; in this example, voltage as a function of time is shown. The signal is periodic with a period equal to 0.8 msec. It is sampled at regular intervals of 0.1 msec. The value of the waveform at each sampling instant is shown by a solid circle.

The middle section of the diagram shows the sampled data. Note that the signal is no longer continuous but consists of a series of discrete pulses. The height of each pulse is equal to the value of the continuous waveform at the sampling instant. Note that, at this stage of the conversion process, the pulse heights are still specified in analog form; i.e., the value of each sample can be recorded at its precise numerical level within the range of the signal's continuous variation. Only the sampling instant is defined discretely.

The bottom section of the diagram shows the sampled data signal in binary form. For the purposes of this example, eight arbitrary levels of quantization are shown. Each of the analog samples shown in the middle section of the diagram has been approximated by a sample with a height equal to one of these eight levels of quantization (The binary representation of these eight quantization levels appears on the ordinate.) The continuous waveform has also been reproduced here in order to show how well the binary samples approximate the original signal (at the sampling instants).

The process of converting an analog sample to binary form involves a sequence of binary “decisions.” The first decision is whether the signal is greater or less than zero. Consider the first sample by way of illustration: this sample has a value of 0.9 volts and therefore the answer to the first question is positive (i.e., sample value is greater than zero) and a 1 is assigned as the first digit of its binary representation. (The analog-to-digital conversion has been set to cover the range from -1 to +1 volts. This range has been chosen to exceed that of the signal being quantized.) After the first binary decision it is evident that the value of the waveform at the first sampling instant must lie between 0 and +1 volts.

The next binary decision is whether this sample value lies above or below the mid-value (0.5 volts) of this range. The answer is again positive (i.e., sample value greater than 0.5 volts) and a 1 is once again assigned as the next digit in the binary representation. It is now known that the waveform must have a value between 0.5 and 1.0 volts at this sampling instant. The mid-value of that range is 0.75 volts and so the third binary decision is whether the sample value is greater or less than 0.75 volts. The answer is again positive so once again a 1 is assigned as the next digit in the binary representation. It is
now known that the waveform must lie between 0.75 and 1.0 volts at this sampling instant.

The above process can be continued iteratively until the binary representation reaches the desired accuracy. (The accuracy of approximation doubles with each additional binary decision.) In the above example, three binary decisions were used leading to a binary representation consisting of three binary digits. The term binary digit is abbreviated as bit and the binary representation derived above is said to have “an accuracy of 3 bits.”

Referring once again to the lowest section of Figure 1, it can be seen that each binary sample lies either just above or just below the original waveform. The distance between quantization levels in this example is 0.25 volts (e.g., Sample #1 lies between 0.75 and 1 volt, Sample #2 lies between 0 and -0.25 volts, etc.). The largest error that can occur is half a quantization step, which in this case is 0.125 volts. As can be seen from the diagram, all of the binary samples are within 0.125 volts of the original waveform. In general, an N-bit analog-to-digital converter will have $2^N$ quantization levels and the error of quantization will not exceed $1/2$ of $1/2^N$, or $1/(2^{N+1})$, of the range of the analog-to-digital converter. In the above example, $N = 3$ and the range of quantization was 2 volts; the number of quantization levels was thus $8$ (= $2^3$) and the precision of quantization was $1/(2^4)$ or $1/16$ of the range, which is $2/16$ (= 0.125) volts.
The Problem of Aliasing

In the digitization of analog signals we encounter the risk of "aliasing." The problem and meaning of aliasing is best described in terms of an example. Section A of Figure 2 shows a sampled-data sequence representing a continuous waveform. Sections B, C and D show three waveforms, each of which would yield the same sequence of samples. Without further information there is no way of knowing which waveform is represented by the sampled-data sequence shown in Section A.

Only one of the waveforms shown in Sections B, C and D is the true waveform, i.e., represents the waveform that was originally sampled. The others are aliases of the true waveform. (It can be shown that there are an infinite number of possible aliases of the true waveform.) If sampling is to be used as a means of preparing continuous signals for digital signal processing, it is essential that there be no ambiguity as to which waveform is represented by a sampled-data sequence.
A Rule for Avoiding Aliasing

The waveform in Section B of the diagram differs from the other waveforms in one important respect, it is represented by more than two samples per period. The waveforms in Section C and D are sampled at a rate of less than two samples per period. For example, in Section C there are nine samples for five periods of the waveform; the rate of samples is thus less than two samples per period. This is true of any other alias corresponding to the sampled-data sequence shown in Section A. A simple rule is thus to assume that the true waveform is the one sampled at more than two samples per period. In order for this rule to hold, it is essential that the sampling rate always be greater than twice the highest frequency of the waveform being sampled. Stated another way, if it is known that the highest frequency in the signal does not exceed half the sampling rate (i.e., there are at least two samples per period of the highest frequency component in the signal) then there can be no aliasing errors.

In order to ensure that there are no aliasing errors, it is common practice to use an anti-aliasing filter prior to the sampling operation. An anti-aliasing filter is typically a lowpass filter with a very high rate of attenuation above the cutoff frequency, fc, where fc < 1/2 sampling rate. Similarly, in order to avoid the generation of spurious waveforms when a sampled-data sequence is converted back to analog form, it is common practice to use an anti-imaging filter after the digital-to-analog converter. It is usual for the anti-imaging filter to have the same lowpass characteristics as the anti-aliasing filter.

Figure 3 shows and describes two typical sampled-data systems. The first is an analog sampled-data system (often referred to simply as a sampled-data system). The second is a digital sampled-data system, referred to here as an all digital system. The latter is often referred to more simply as a digital system.

ELEMENTS OF A DIGITAL HEARING AID

The basic operations performed by a hearing aid are amplification, filtering, and output limiting. Amplification in a digital system is achieved by simply multiplying the samples representing the audio signal by a constant A. The magnitude of A determines the amount of amplification. Amplitude limiting is also achieved fairly simply, by setting a maximum allowable value for the samples contained in the digital representation. An alternative method of output limiting is to adjust the amplification constant, A, in inverse proportion to the short-term energy of the signal, thereby reducing the gain as signal level is increased. The first of these two methods is the digital equivalent of peak-clipping, (a technique commonly used in older conventional hearing aids); the second method is the digital equivalent of amplitude compression (a technique used increasingly in modern conventional hearing aids).

In contrast to the above operations, the process of filtering audio signals represents a particularly interesting and innovative application of digital techniques. Unfortunately, digital filters are difficult to describe without some use of mathematics. The Appendix provides an introduction to the key concepts underlying digital filtering. The non-mathematical reader may wish to skip the Appendix.

Digital filters, as illustrated in the Appendix, operate on sampled-data waveforms in much the same way (but not in exactly the same way) as an electronic filter operates on an electrical waveform. It may be noted that this filtering operation is also analogous to taking a running arithmetic average of the data sequence.

An important advantage of digital filtering over conventional electronic filtering is the potential for increased precision which may be several orders of magnitude greater than can be achieved in practice using conventional electronic (analog) components. For example, it is not difficult to create a digital filter with a very steep rate of attenuation (e.g., over 1000 dB/octave). It is extremely difficult to design and build such a filter from analog components.

A second important advantage is that the digital filter can be reprogrammed to have vastly different characteristics without any change in hardware; i.e., the same piece of equipment is used. This change can be made to take place in a fraction of a second.

A third fundamental advantage of the digital filter is that it can be programmed to include logical operations; e.g., switching itself off, or changing its characteristics, in response to preselected events in the signal.
Figure 3. This figure shows two typical sampled-data systems. The upper block diagram represents an analog sampled-data system (often referred to as a sampled-data system). It consists of an anti-aliasing filter, a sampling circuit, a signal processor for operating on sampled-data sequences, a circuit for waveform reconstruction, and an anti-imaging filter.

The second system is the corresponding all-digital system. It consists of an anti-aliasing filter, an analog-to-digital converter (this unit contains both a sampler and a circuit for converting the samples to binary form), a digital signal processor (this could be a general-purpose digital computer), a digital-to-analog converter, and an anti-imaging filter.
A BRIEF HISTORY OF DIGITAL HEARING AIDS

Early Applications of Digital Techniques

Sampled-data systems were first used in automatic control work as well as in the transmission of information (23,24). Theoretical analyses of these early systems led to the development of several important theorems, most notably the sampling theorem described earlier that determines the minimum sampling rate necessary to specify unambiguously a continuous signal of known bandwidth.

Although the theoretical underpinnings of discrete digital signal analysis were well developed some time ago (17), it was not until the advent of the digital computer that these techniques started to take on a new, practical importance. A concomitant development of great consequence for the digital processing of audio signals was the development of analog-to-digital and digital-to-analog converters fast enough and with enough precision for the conversion of continuous audio signals to digital form, and vice versa. A typical system for telephone-quality speech would have a sampling rate of 10,000 Hz with 8-bit to 10-bit accuracy. A high fidelity system would require a sampling rate of at least 40,000 Hz with 16-bit accuracy or better.

Much of the early research on the digital processing of audio signals focused on speech analysis, speech synthesis, and vocoder design (4). During that early period in the development of digital audio, the time taken by a computer to process audio signals was extremely long. Typically, a fairly simple speech-processing algorithm (by today's standards) would take several hundred times real time. As a consequence, almost all of the research on the early development of digital signal processing for audio applications was done off-line. Even with this major limitation, it was far more efficient to develop experimental audio systems using off-line computer simulation rather than by building experimental prototypes. The latter approach is typically far more costly and time consuming than that of computer simulation. Further, experimental prototypes often do not meet all design requirements. The use of computer simulation has grown dramatically as a research and development tool and it is now widely used in almost all industrial research laboratories.

An important step in the development of computer-simulation techniques in digital audio was the introduction of high-level languages for facilitating the simulation of audio systems. One such language was BLODI, an acronym for Block Diagram Compiler, developed by Kelly, Lochbaum and Vyssotsky (11) in 1961. This language could be used to simulate any realizable audio system specified in block diagram form.

BLODI has been used for a wide range of applications including the computer simulation of a high-gain telephone with frequency shaping for hearing-impaired persons. That simulation, conducted by the author in 1967, was probably the first use of a digital computer in simulating a hearing aid. It was recognized at the time that a computer could also be used to adaptively adjust the frequency response and other characteristics of a hearing aid so as to best meet the needs of the user; however, because of severe practical limitations on the speed of simulation and the constraint that all processing be done off-line this potentially useful approach was not pursued until well over a decade later.

The advent of the laboratory computer brought the realization of a digital hearing aid one step closer. An important feature of the laboratory computer is the relative ease with which it can be used to control laboratory equipment. Computer-controlled audio systems for research in audition were developed soon after the introduction of the modern laboratory computer. The earliest of these systems were configured around a LINC-8, the predecessor of the highly successful DEC PDP-8 laboratory computer. An early system of this type was developed by Braida and his associates at the Massachusetts Institute of Technology (personal communication with author) and has been heavily used in research in psychoacoustics. A more modern version of the system has been used extensively in experiments on acoustic amplification and signal processing for hearing impairment. Although not designed specifically as such, this early system was in essence a computer-controlled master hearing aid used for research in acoustic amplification.

Quasi-Digital Hearing Aids

In a “quasi-digital” hearing aid, conventional analog amplifiers and filters are controlled by digital means. A simple, practical realization of this approach is to use the computer only for programming the hearing aid. Once programmed, the hearing aid is disconnected from the computer and is then used
A Quasi-Digital Hearing Aid.

The analog components of the hearing aid (amplifiers, filter and output limiter) are controlled by the digital controller. Note that not all the analog components need be under digital control.

in essentially the same way as a conventional hearing aid. This approach has not had much success commercially with hearing aids, although it has been used effectively with cochlear implant prostheses (1).

A closely related application of computers in this context is that in which a laboratory computer is used as an audiometer and prescriptive tool. Popelka and Engebretson (22) have developed a system of this type in which the computer is used to facilitate the measurement, display and interpretation of audiological data in order to prescribe a hearing aid more efficiently. The computer can also be used to search for the best hearing aid for a given patient from a data base containing detailed electroacoustic specifications of available hearing aids (10).

The next step in the development of a practical quasi-digital hearing aid is that in which both the analog components and the digital controller are combined in a single wearable unit. The basic architecture of such a hearing aid is shown in Figure 4. Note that the digital controller not only controls or programs the operation of the analog components (amplifiers, filter, and limiter), but is itself programmed by an external computer using a temporary connection. The external programming and control is performed in the clinic or hearing-aid dispensary when the unit is first prescribed and fitted.

The concept of a digitally-controlled analog hearing aid is very attractive from a practical perspective because of the low power consumption involved. The technology of low-power analog amplifiers and filters is well advanced, whereas the present generation of chips for digital signal processing draws relatively large amounts of power. The possibility of combining the low power consumption of analog components with the greater signal processing capability of digital components now appears to be a viable option. A hearing aid of this type has been developed by Etymonic Design, Inc., for Dahlberg Electronics (William A. Cole, personal communication). This also appears to be the approach taken by Mangold and Leijon (16) who have developed a programmable, multiband hearing aid. Similarly, Graupe, et al., (7) have developed an adaptive noise-reducing filter on a single chip using a quasi-digital approach. This chip is small enough and of sufficiently low power consumption to fit in a conventional behind-the-ear or in-the-ear hearing aid.
Another form of quasi-digital hearing aid (and which may have also been used in the applications cited above) is that of a sampled-data system in which the audio signal is sampled at discrete intervals in time, but the samples are kept in analog form during processing. Such a system can be implemented fairly easily in practice using switched-capacitance techniques. Consider, for example, a sequence of capacitors with each capacitor containing a charge representing a single sample in a sampled-data sequence. At each sampling instant, the charge on each capacitor is switched to the next capacitor in the sequence. The value of the charge on each capacitance is multiplied by a coefficient. The sum of these products is the output of the filter. The filter described here is essentially the same as that shown in Figure A3 of the Appendix.

The characteristics of this filter are determined by the choice of the coefficients. In a programmable version of the filter, the coefficients are adjusted by a digital controller. Note that although the incoming audio waveform is sampled discretely in time (hence the need for anti-aliasing and anti-imaging filters) the samples themselves remain in analog form throughout the processing. In an all-digital version of the above filter, the samples would be converted to binary form prior to processing.

The use of the switched-capacitance technique, or similar quasi-digital methods, has the important advantage that digital signal-processing techniques can be used without the need for analog-to-digital or digital-to-analog converters. The power consumed by an analog-to-digital converter increases rapidly with degree of quantization. The development of high resolution analog-to-digital converters of small size and low power consumption, suitable for use in a practical hearing aid, is still a difficult technical problem.

All-Digital Hearing Aids

In an all-digital hearing aid both the processing of the audio signals and the control of the processing are done by digital means. Further, all sampled waveforms are converted to binary form for ease of processing and then converted back to analog form after processing to drive the earphone. Graupe (6) appears to have been the first to implement such a system using an 8080 microprocessor. The approach used was conceptually similar to that of the digitally controlled analog system shown in Figure 4, except that the filter, limiter, and one or more of the amplifiers have been replaced by equivalent digital components. Further, because of the great flexibility of control afforded by the microprocessor, it was possible to program the system to be self-adaptive, thereby opening the door to the use of advanced signal-processing techniques for noise reduction and intelligibility enhancement. Although the 8080 microprocessor used by Graupe was both slow and relatively large in size, he clearly anticipated the day when microprocessors would be fast enough and small enough for use in a practical hearing aid.

At about this time (the late 1970s) the concept of an all-digital programmable hearing aid was also being developed in Germany (18) as well as at the Institute for Hearing Research in Nottingham, England (personal communication between M. Haggard and the author). The approach followed by the latter group was that of using a self-standing digital filter controlled by a standard microcomputer, a technique that was to be adopted later by several other research laboratories (15,25).

Although the concept of a digital hearing aid was anticipated at an early date, two major technical problems had to be resolved before anyone could develop a practical all-digital instrument. The first was the development of a digital signal processor fast enough to operate in real time. The second, more difficult problem (which has yet to be resolved satisfactorily) is that of developing digital circuitry that is small enough and sufficiently low in power consumption for practical use in a small, wearable unit.

The first breakthrough came with the development of the array processor in which an array of numbers is processed simultaneously, instead of only one number at a time as in a conventional digital computer. The saving in processing time resulting from the use of this technique is sufficient to allow for real-time processing of audio signals. High-speed array processors were introduced toward the end of the nineteen seventies and shortly afterwards an all-digital hearing aid was developed configured around one of these units (12). Although still too large to be wearable, that instrument has been used effectively as a master hearing aid in a series of experimental investigations on the prescriptive fitting and evaluation of digital hearing aids (13,14,19).

Another important development was the introduction of a family of high-speed digital-signal-
processing (DSP) chips in 1982. Although not as fast as an array processor, these chips are fast enough for limited real-time processing of speech signals. Because of their small size, these chips can be packaged in a unit small enough to be wearable. Experimental body-worn digital hearing aids were developed soon after high-speed DSP chips became available (2,3,8,9,20). It is not clear which research group was first in the race to develop a body-worn digital hearing aid, since in addition to the research groups identified above, at least three other industrial research laboratories have developed instruments of this type but are extremely secretive about their progress.

The second major problem, that of reducing power consumption and physical size so that the digital chips are small enough to fit in a behind-the-ear or in-the-ear hearing aid, has yet to be resolved. Under normal operation, the various DSP chips currently on the market (as of March 1987) draw far too much current for a practical behind-the-ear unit. An important recent development is that of the application-specific DSP chip dedicated to a specific task and designed for very low power consumption. Application-specific analog chips for hearing aid use were developed well over a decade ago with highly successful results; the major research groups concerned with the development of a practical, wearable digital hearing aid are now actively engaged in developing their own low-power, application-specific DSP chips. It is likely that a practical solution will be found in the near future. At least one group, Project Phoenix in Madison, Wisconsin, has indicated that it will begin marketing a body-worn all-digital hearing aid before the end of 1987 and that a behind-the-ear unit will be available within a year. (Personal communication between K. Hecox, Project Director, and the author)

APPENDIX

Two Basic Classes of Digital Filters

Filtering may be viewed as a form of averaging or smoothing. A very simple form of averaging is to take the mean of successive pairs of samples. Consider, for example, the sampled-data sequence representing the waveform shown in Figure 1 of the text. The numerical values of this sampled-data sequence are reproduced in the third column of Table A1. The fourth column of the table shows the smoothing resulting from the operation:

\[ Y_i = 0.5 (X_i + X_{i-1}) = 0.5 X_i + 0.5 X_{i-1}, \]  

where \( Y_i \) is the smoothed sample value at sampling instant \( i \), and \( X_i \) is the value of the original sample-data sequence at sampling instant \( i \).

<table>
<thead>
<tr>
<th>Sample Instant (ms)</th>
<th>Time (Eqn 1a)</th>
<th>X: (Eqn 1a)</th>
<th>Y: (Eqn 1b)</th>
<th>Yi (Eqn 1b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>.85</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>0.1</td>
<td>.9</td>
<td>4</td>
<td>.325</td>
</tr>
<tr>
<td>2</td>
<td>0.2</td>
<td>-1</td>
<td>4</td>
<td>.325</td>
</tr>
<tr>
<td>3</td>
<td>0.3</td>
<td>-0.43</td>
<td>-265</td>
<td>-.268</td>
</tr>
<tr>
<td>4</td>
<td>0.4</td>
<td>-1</td>
<td>-265</td>
<td>-.234</td>
</tr>
<tr>
<td>5</td>
<td>0.5</td>
<td>-0.43</td>
<td>-265</td>
<td>-.547</td>
</tr>
<tr>
<td>6</td>
<td>0.6</td>
<td>-1</td>
<td>-265</td>
<td>-.373</td>
</tr>
<tr>
<td>7</td>
<td>0.7</td>
<td>-0.43</td>
<td>-265</td>
<td>-.617</td>
</tr>
<tr>
<td>8</td>
<td>0.8</td>
<td>-1</td>
<td>-265</td>
<td>-.408</td>
</tr>
<tr>
<td>9</td>
<td>0.9</td>
<td>.9</td>
<td>4</td>
<td>.696</td>
</tr>
<tr>
<td>10</td>
<td>1.0</td>
<td>-1</td>
<td>4</td>
<td>.248</td>
</tr>
<tr>
<td>11</td>
<td>1.1</td>
<td>-0.43</td>
<td>-265</td>
<td>-.306</td>
</tr>
<tr>
<td>12</td>
<td>1.2</td>
<td>-1</td>
<td>-265</td>
<td>-.253</td>
</tr>
<tr>
<td>13</td>
<td>1.3</td>
<td>-0.43</td>
<td>-265</td>
<td>-.557</td>
</tr>
<tr>
<td>14</td>
<td>1.4</td>
<td>-1</td>
<td>-265</td>
<td>-.378</td>
</tr>
<tr>
<td>15</td>
<td>1.5</td>
<td>-0.43</td>
<td>-265</td>
<td>-.619</td>
</tr>
<tr>
<td>16</td>
<td>1.6</td>
<td>-1</td>
<td>-265</td>
<td>-.410</td>
</tr>
<tr>
<td>17</td>
<td>1.7</td>
<td>.9</td>
<td>4</td>
<td>.695</td>
</tr>
<tr>
<td>18</td>
<td>1.8</td>
<td>-1</td>
<td>4</td>
<td>.247</td>
</tr>
<tr>
<td>19</td>
<td>1.9</td>
<td>-0.43</td>
<td>-265</td>
<td>-.306</td>
</tr>
<tr>
<td>20</td>
<td>2.0</td>
<td>-1</td>
<td>-265</td>
<td>-.233</td>
</tr>
<tr>
<td>21</td>
<td>2.1</td>
<td>-0.43</td>
<td>-265</td>
<td>-.557</td>
</tr>
<tr>
<td>22</td>
<td>2.2</td>
<td>-1</td>
<td>-265</td>
<td>-.378</td>
</tr>
<tr>
<td>23</td>
<td>2.3</td>
<td>-0.43</td>
<td>-265</td>
<td>-.619</td>
</tr>
<tr>
<td>24</td>
<td>2.4</td>
<td>-1</td>
<td>-265</td>
<td>-.410</td>
</tr>
<tr>
<td>25</td>
<td>2.5</td>
<td>.9</td>
<td>4</td>
<td>.695</td>
</tr>
</tbody>
</table>

ACKNOWLEDGMENTS

I would like to thank Linda Ashour and Matt Bakke for their help in the preparation of this paper. I am also most grateful to Stanley Daly for his editorial input.
and $X_{i-1}$ is the sample value at the preceding sampling instant.

The effect of the smoothing operation given by Equation [1a] is to remove much of the ripple contained in the original sampled-data sequence. The sampled-data sequence represented by the $Y_i$ is, in essence, a filtered version of the original sampled-data sequence, $X_i$. The digital filter represented by Equation [1a] is shown schematically in the upper half of Figure A1. The symbol $D$ represents a delay of 1 sampling interval. The symbols $\times$ and $+$ represent multiplication and addition, respectively. Note that the coefficients $k_1$ and $k_2$ are both equal to 0.5 in this example. Note also that only four operations are required in order to implement this digital filter, (1 delay, 2 multiplications, and 1 addition).

Another very simple digital filter is shown in the lower half of Figure A1. In this case, the output of the digital filter, $Y_i$, is given by

$$Y_i = X_i + k_1 Y_{i-1}\ldots\hspace{1cm}[1b]$$

In this case, only three digital operations are required (1 addition, 1 delay, and 1 multiplication). The fourth column of Table A1 shows the output sampled-data sequence $Y_i$ for $k_1 = 0.5$ in the above filter. This filter is not quite as effective in removing the ripple from the original sampled data sequence, but it has other useful properties. Note that this filter also takes time to settle down. It is only toward the end of the second period of the original data sequence, $X_i$, that the output of the filter, $Y_i$, approximates a periodic function.

It is conceptually convenient to think of the smoothing operations given by Equations [1a] and [1b] in terms of the frequency responses of the associated digital filters. The filter corresponding to Equation [1b] is essentially a lowpass filter. In electrical engineering terms the frequency response of this digital filter is analogous (but not identical) to that of a single-pole lowpass network, as commonly realized by the simple RC network shown in Figure A2.

![Figure A1. Simple Digital Filters.](image)

The upper section of the diagram shows a simple finite-impulse-response (FIR) filter. The lower section of the diagram shows a simple recursive filter having an infinite impulse response (IIR).

Equations [1a] and [1b] can be generalized as follows:

$$Y_i = k_1 X_i + k_2 X_{i-1} + k_3 X_{i-2} + \cdots + k_n X_{i-n+1}\ldots\hspace{1cm}[2a]$$

$$Y_i = X_i + k_1 Y_{i-1} + k_2 Y_{i-2} + \cdots + k_n Y_{i-n}\ldots\hspace{1cm}[2b]$$

where $Y_i = \text{output sample at time instant } i$, $X_i = \text{input sample at time instant } i$, and the $k_j$ are constants ($j = 1,2,3,\ldots,n$).
The digital filter corresponding to Equation [2a] is referred to as a finite-impulse-response (FIR) filter since the output \( Y_i \) is determined by only the last \( n \) values of \( X_i \). In contrast, the digital filter corresponding to Equation [2b] is referred to as a recursive or infinite-impulse response (IIR) filter. In this case, the output \( Y_i \) is affected by all previous values of \( X_i \). In order to illustrate this point, consider the term \( Y_{i-1} \) in the simplest filter of this type, as given by Equation [1b]. The value of \( Y \) at sampling instant \( i-1 \) is determined from the previous value of \( Y_i \), i.e., \( Y_{i(-1)} = X_{i(-1)} + k_1 Y_{i(-2)} \). The value of \( Y_{i(-2)} \), in turn, is dependent on \( Y_{i(-3)} \), which in turn is dependent on \( Y_{i(-4)} \), and so on. If the constant \( k_1 \) is less than 1, then the earlier values of \( Y \) become progressively less important in deriving \( Y_i \).

The digital filter corresponding to Equation [2a] is shown in Figure A3. Electrical engineers will recognize this circuit as a transversal filter. The frequency response of the filter is determined by the coefficients \( k_1, k_2, \ldots, k_n \). Expressed mathematically, the frequency response of the filter is

\[
F(f) = \text{DFT} [k_1, k_2, \ldots, k_n]
\]

where \( F(f) \) is the frequency response of the filter specified at discrete frequencies \( f_i = (i-1, 2, \ldots, m/2) \) and \( \text{DFT} [k_1, k_2, \ldots, k_n] \) is the discrete Fourier transform of the set of coefficients.

Note that as the number of coefficients is increased, the precision with which the frequency response is specified is increased accordingly; i.e., the frequency interval between the discrete components \( f_i \) is proportional to \( 1/n \). Note also that each value of \( F(f_i) \) is a complex number that specifies both amplitude and phase of the frequency response at each discrete frequency. Since the coefficient array consists of \( n \) real numbers, the resulting discrete frequency response consists of \( n/2 \) complex numbers. For mathematical convenience, \( n \) should be an even number.

The digital filter corresponding to Equation [2b] is shown in Figure A-4. The frequency response of this filter is obtained by finding the roots of an \( n \)th order polynomial in \( f \) with coefficients \( k_1, k_2, k_3, \ldots, k_n \). This is not always easy to do when \( n \) exceeds 3. A case of particular interest occurs when \( n = 2 \); i.e., when the digital filter shown in Figure A-4 consists of only two delay loops. The frequency response of this filter is relatively easy to derive mathematically and is equal to that of a simple resonant circuit. If \( f_0 \) is the resonant frequency of the circuit and \( \sigma_0 \) its bandwidth (i.e., the resonance
is defined mathematically as consisting of the complex pole-pair $\sigma_o \pm j2\pi f_o$, then the coefficient $k_1$ is given by $2e^{-\sigma_o D}\cos(2\pi f_o D)$ and $k_2$ is given by $-e^{2\pi j D}$.

The examples cited above represent two basic classes of digital filter. There are, of course, many variations of these basic filter types, the analysis of which can be quite complex. For further information on this topic, the reader is referred to the classic texts of Gold and Rader (5) and Oppenheim and Shafer (21), among others.
REFERENCES


