TECHNICAL REVIEW

No. 1 — 1978

Contents

Digital Filters and FFT Technique in Real-Time Analysis

by

R. B. Randall and R. Upton

ABSTRACT

The ever decreasing price of digital components has meant that all-digital methods of real-time analysis now predominate where previously analog or hybrid methods were used. This article discusses the two main methods of digital real-time analysis, namely FFT, (Fast Fourier Transform), and Digital Filtering. It shows that while FFT is ideal for producing an analysis with constant bandwidth on a linear frequency scale, Digital Filtering is the best method to use for constant percentage bandwidth measurements on a logarithmic frequency scale. Applications of the two methods are also discussed, particularly with respect to the choice of the correct analyzer for a given application, together with the whole concept of "real-time" and when real-time analysis is really necessary.

SOMMAIRE

Du fait de la diminution constante du prix des composants digitaux, les techniques entièrement numériques d'analyse en temps réel prédominent actuellement, alors que auparavant on utilisait les techniques analogiques ou hybrides. Cet article traite des deux principales techniques d'analyse numérique en temps réel, la FFT (transformée de Fourier rapide) et l'emploi de filtres numériques. Il montre que si la FFT est idéale pour l'obtention d'une analyse à bande constante sur une échelle de fréquence linéaire, l'emploi de filtres numériques constitue la meilleure méthode pour les analyses à pourcentage de bande constant sur une échelle de fréquence logarithmique. Cet article traite également des applications de ces deux méthodes, en particulier du point de vue du choix de l'analyseur convenant le mieux à une application donnée; il traite, en outre, du concept général de "temps réel" et indique dans quels cas l'analyse en temps réel est vraiment nécessaire.

ZUSAMMENFASSUNG

Der immer weiter sinkende Preis digitaler Komponenten hat bewirkt, daß volldigitale Methoden in der Echtzeitanalyse die früheren analogen oder hybriden Methoden immer weiter verdrängen. In diesem Artikel werden die beiden wichtigsten Methoden digitaler Echtzeitanalyse, nämlich FFT (Fast Fourier Transform — Schnelle Fourier Transformation) und digitale

Filter, diskutiert. Es wird gezeigt, daß FFT am geeignetsten für Analysen mit konstanter Absolutbandbreite auf linearer Frequenzskala ist, während sich digitale Filter als die beste Methode für Analysen mit konstanter Relativbandbreite auf logarithmischer Frequenzskala erweisen. Für beide Methoden werden Anwendungsbeispiele diskutiert, wobei speziell auf die Wahl des richtigen Analysators für einen gegebenen Fall eingegangen wird, zusammen mit dem gesamten Echtzeit-Problem und wenn Echtzeit-Analysen wirklich notwendig sind.

Introduction

In recent years there has been a rapid development in the field of frequency analysis instrumentation. In particular, the constantly falling price of digital components has meant that completely digital instruments are now very competitive with the older analogue analyzers previously used for frequency analysis, and have completely replaced them for certain applications. The name RTA (Real-time Analyzer) has become a commonplace term for a rapid analyzer producing a complete spectrum in parallel and displaying it on a continuously updated screen. This feature of rapidly producing continuously updated spectra is very valuable for trouble-shooting, but one of the purposes of this article is to discuss the importance of the concept of "real-time"; when it is and when it isn't really necessary.

The main purpose, however, is to discuss the pros and cons of the two major digital methods for obtaining a real-time analysis, viz. (recursive) digital filtering and FFT (Fast Fourier Transform). It is found that digital filtering is the best technique for constant percentage bandwidth analysis (on a logarithmic frequency scale) while FFT is the best technique to use for constant bandwidth analysis (on a linear frequency scale). Thus the choice of the correct analyzer for a given application depends greatly on which of these two forms of presentation of the results is preferred or required.

There may appear to be an inconsistency in the fact that Brüel & Kjær's 1/3-Octave Digital Frequency Analyzer Type 2131 (Fig.1) (based on digital filtering) is real-time up to 20 kHz whereas the Narrow Band Spectrum Analyzer Type 2031 (Fig.2) (based on FFT) is real-time only to 2 kHz. In fact there is no real inconsistency because a narrow-band analyzer would most often be used to analyze signals containing discrete tones (e.g. machine vibrations), and a discrete frequency must remain constant for an appreciable time before it can be recognized as such, whereas signals which change rapidly, and for which a real-time frequency up to 20 kHz is more likely to be necessary, will often be intrin-

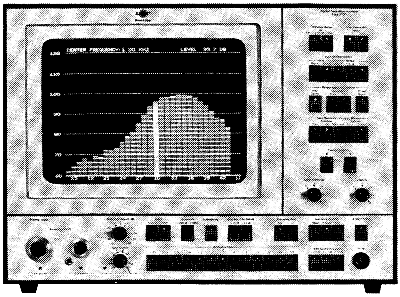


Fig. 1. Digital Frequency Analyzer Type 2131

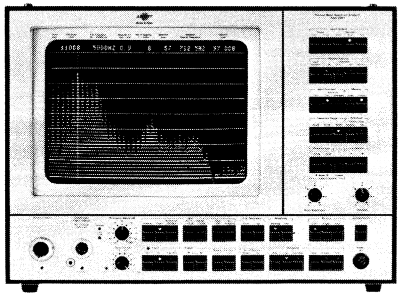


Fig. 2. Narrow Band Spectrum Analyzer Type 2031

sically broad-band and thus optimally analyzed in 1/3-octaves. This is a fairly broad statement, but will be enlarged upon later in the article when typical examples of applications of the two techniques will be given.

The prospective purchaser of an up-to-date "real-time" analyzer is faced with a difficult choice, because it is only rarely that he will be able to acquire both analyzer types. It is hoped, however, that the present discussion will aid the decision by giving an objective survey of the pros and cons of each type. In particular, there are situations where one or the other type is absolutely necessary, and a few situations where, in fact, both may be required. There are still a number of situations, however, in which the choice is not so clear-cut, but at least the information given here will allow the choice to be made objectively, with the appropriate weighting being given to those factors which play a significant role in a particular case.

Choice of Frequency Scale and Bandwidth

As mentioned, there is a basic choice to be made between constant absolute bandwidth and constant relative (percentage) bandwidth where the absolute bandwidth is a fixed percentage of the tuned centre frequency. Fig.3 compares these two alternatives on both linear and logarithmic frequency scales and illustrates one of the most fundamental differences between them.

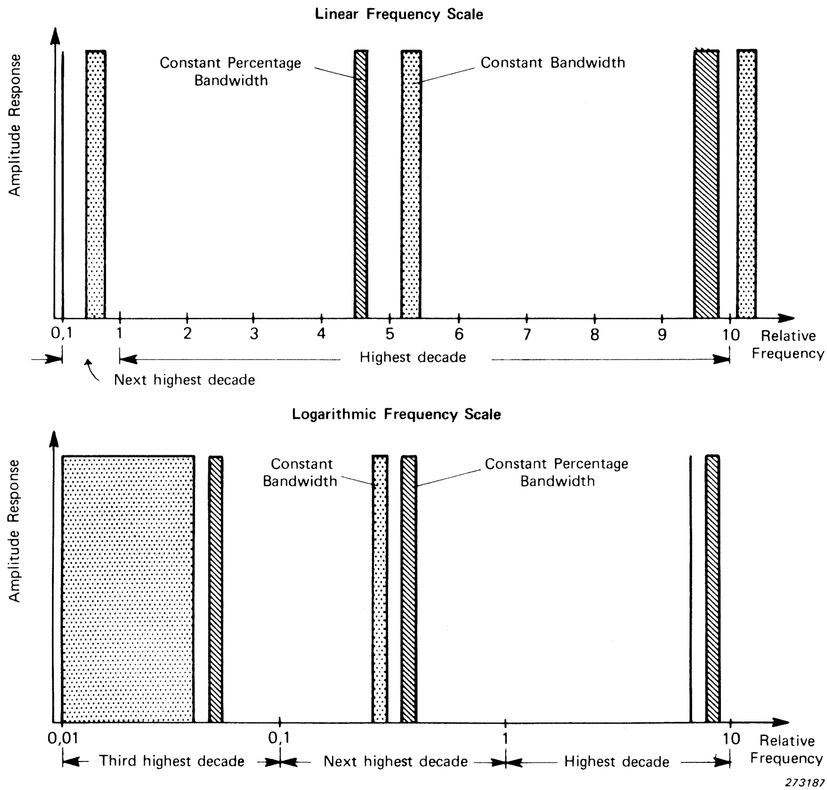


Fig.3. Difference between a constant bandwidth analyzer and a constant percentage bandwidth analyzer

Constant bandwidth gives uniform resolution on a linear frequency scale, and this gives equal resolution and separation of harmonically related components which facilitate detection of a harmonic pattern (Fig.4).

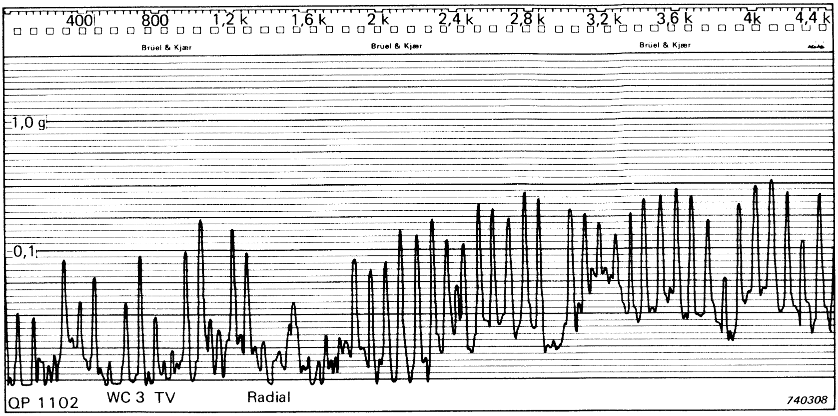


Fig. 4. Vibration spectrum having many harmonically related components

Other situations where this capability is desired are in the separation of equally-spaced sidebands (resulting from modulation) and in the detection of sub- and "inter"-harmonics which result from mechanical looseness (rattling). This sort of problem is usually associated with diagnosis. However, the linear frequency scale automatically gives a restriction of the useful frequency range to (at the most) two decades, as is evident from Fig. 3.

Constant percentage bandwidth, on the other hand, gives uniform resolution on a logarithmic frequency scale and thus can be used over a wide frequency range of 3 or more decades. (Fig. 3.)

A typical situation where this is important is in acoustical studies where it is desired to cover the entire audio range (20 Hz — 20 kHz). Another is in spectrum monitoring of machines where it is often necessary to cover the range from half running speed to say, three times a toothmeshing frequency; this can easily extend over 3 decades. Another feature of constant percentage bandwidth is that it corresponds to constant Q-factor (amplification ratio of resonant peaks) (Fig. 5). It is thus both natural and efficient to analyze spectra dominated by structural resonances on a logarithmic frequency scale with a constant percentage bandwidth somewhat narrower than the narrowest resonant peak.

Other grounds for using a logarithmic frequency scale (though not necessarily constant percentage bandwidth) are:

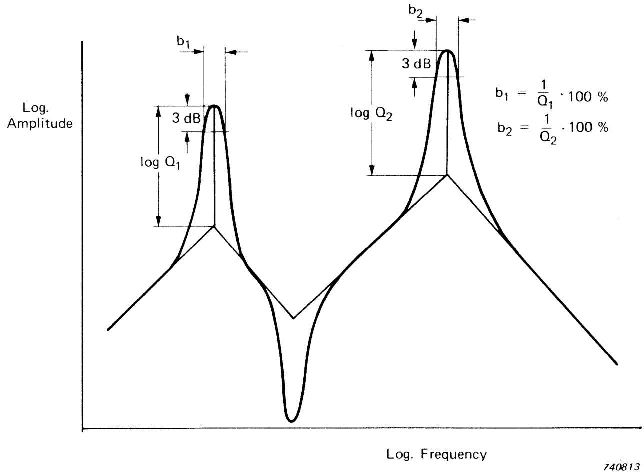


Fig.5. Relationship between the amplification factor Q and the relative bandwidth b

a) Small speed changes in, say, a machine only give a lateral displacement of the spectrum, thus simplifying direct comparison.

b) Certain relationships can most easily be seen on log-log scales such as for example, integration, which gives a change in slope of —20 dB/decade and thus means that constant velocities and displacements are represented by straight lines on an acceleration vs. frequency diagram.

Real-time analysis — when is it necessary?

Perhaps the most succinct definition of "Real-time" frequency analysis is "analysis of all the signal in all frequency bands all of the time". A necessary (though not sufficient) condition is that the calculation process must be at least as fast as the time taken to collect the data processed. Where the data is processed blockwise (as in the FFT procedure) then it is also necessary to store the new data arriving while a calculation process is going on, in order that none be missed. Where the data is processed on a sample-for-sample basis (as in digital filtering) then this intermediate storage is not required.

Before discussing when real-time analysis is necessary, it is perhaps a good idea to give some examples of situations where it isn't really necessary, even though a rapid analysis may be desirable.

1) For stationary signals — These are virtually defined as signals whose statistical properties do not change with time, and thus any portion of a given length is an equally valid sample of the signal. An average over a number of non-overlapping samples will give an equally valid result whether or not the samples are contiguous (the real-time case). Real-time analysis will, of course, give the result in the shortest possible time, but even if for example only 10% of the total signal is analyzed, the analysis time may still be short in comparison with the time taken to set up the measurement, and therefore quite acceptable. For example, an average of ten 400-line spectra with 20 kHz full-scale frequency would take 200 ms in real-time, but still only 2 s where only 10% of the total signal is processed. Both answers would be equally valid.

2) For transients (where the entire signal to be analyzed fits into the analyzer memory) — Only the recording in the memory has to be in real-time; the analysis can of course, be performed after the signal has been recorded.

3) For slowly changing non-stationary signals — Even though a signal is non-stationary, it does not necessarily have to be analyzed in real-time, this depending on how rapidly it is changing. For example, for machine run-ups and run-downs it is unlikely that the signal would change significantly in 200 ms (the calculation time for the 2031 Analyzer) even though the memory length in the 20 kHz frequency range is 20 ms. Thus, reading out a new spectrum every 200 ms (or even slower) would often adequately describe the changing signal.

Turning to the question of when real-time analysis is required, it will be seen to be only when the above considerations do not apply, viz., for rapidly changing non-stationary signals. Three typical examples should make this clear.

a) Speech analysis — Speech sounds last typically 40 ms or so, and there are rapid transitions between them, from vowels to plosives, for example. On the other hand, it must be kept in mind that to benefit from the real-time analysis of speech it is necessary to feed the results into a computer. Even appreciation of the results by eye may be difficult; the authors know of a case where a Digital Event Recorder Type 7502 was used to slow down a speech signal by a factor of

100:1 to permit the changes to be perceived visually on a real-time analyzer (with 100 Hz full-scale frequency thus corresponding to 10 kHz in the original signal). The analyzer used (Type 3348) was in fact capable of analyzing in real-time to 10 kHz. This also indicates how speech may be analyzed by an analyzer with lower real-time frequency. Intermediate recording on tape, and slowing down by 10:1 would permit analysis by an analyzer with 2 kHz real-time frequency (e.g. Type 2031).

b) Aircraft Flyover Noise — IEC and FAA regulations require that a large proportion (typically > 95%) of the total signal is analyzed and thus real-time analysis is unavoidable. Note that real-time operation of both filtering and detecting systems is implied.

c) Reverberation time measurements — Measurement of reverberation time in 1/3-octave bands requires that averaged 1/3-octave band sound pressure levels be sampled at short intervals (typically 1/10 of the reverberation time). This is discussed more fully in Section 6.

Digital Filtering

Only those aspects of digital filtering which are relevant to the discussion here will be touched upon; for a fuller discussion of digital filtering the reader is referred to Ref. 1.

A recursive digital filter is a digital processor which receives a sequence of digital values at its input, (usually samples of a continuous time signal), carries out a digital operation (involving a limited number of the previous output values) on each input value, and outputs a sample for each input sample. The continuous sequence of output values will in general, be filtered with respect to the input values, and the filtering operation can be designed to be equivalent to virtually any analogue filtration on the equivalent continuous analogue signal.

Fig.6 shows a typical 2-pole digital filter section as employed in the Analyzer Type 2131. Its filtering action can be made equivalent to virtually any analogue 2-pole filter by appropriate choice of the constants H_0, B_1 and B_2. It will be appreciated, for example, that for it to be equivalent to a simple (2-pole) damped resonator, the 3 coefficients give sufficient flexibility to vary the resonance frequency, damping factor, and overall amplification of the resonator.

Higher order filters can be obtained by cascading 2-pole sections (the z-

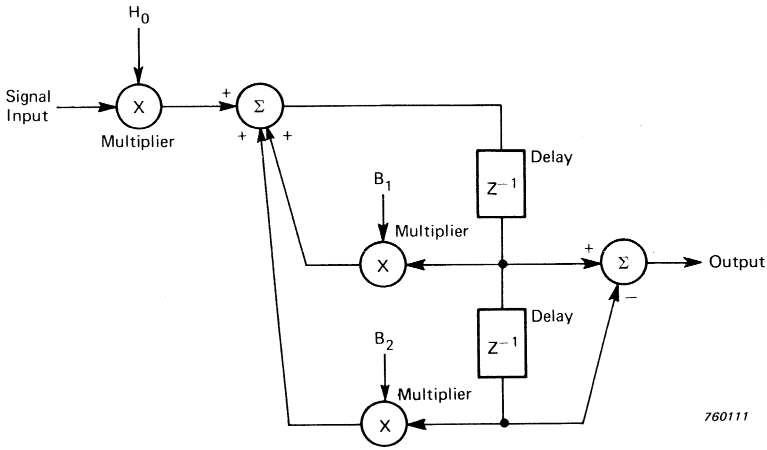


Fig.6. Block diagram of 2-pole digital filter used in the 2131

transform of the overall characteristic will thereby be the product of the individual z-transforms, see Ref.1). Ref.1 gives details of how equivalents of well-known analogue filter types can be designed. There are two very important points to remember about digital filters designed in this way:

1) The type (e.g. bandpass, lowpass, highpass) and class (e.g. Butterworth, Chebyshev) of filter is entirely determined by the filter coefficients (e.g. B_1, B_2, H_0 in Fig.6) and once decided results in permanently fixed filter characteristics.

2) The filter characteristics are determined only in relation to the sampling frequency; a digital filter knows nothing about actual physical frequencies. Thus the centre frequency and bandwidth (of, for example a bandpass filter) for the same set of coefficients, will be directly proportional to the sampling frequency; halving the sampling frequency will result in the same proportional bandwidth, one octave lower in frequency. This is one of the reasons why digital filters are most appropriate to constant percentage bandwidth analysis on a logarithmic frequency scale (i.e. one based on octaves).

Another is related to the efficiency with which the same hardware can be timeshared between a large number of different filters when the frequency scale is based on octaves. A discussion of the mode of operation of the Digital Frequency Analyzer Type 2131 should make this clear.

Fig.7 is a block diagram of the input and filter section of the 2131. It will be seen that the signal, shortly after entry, is converted into digital form and from then on all operations are digital. Before analogue-to-digital (A/D) conversion, the signal is first lowpass filtered with a 12-pole analogue lowpass filter having its cut-off at 27 kHz which is above the highest frequency of interest, viz. 22,4 kHz (the upper limiting frequency of the 1/3-octave centred on 20 kHz). This is done to avoid aliasing* (the sample rate of the A/D converter is 66,667 kHz).

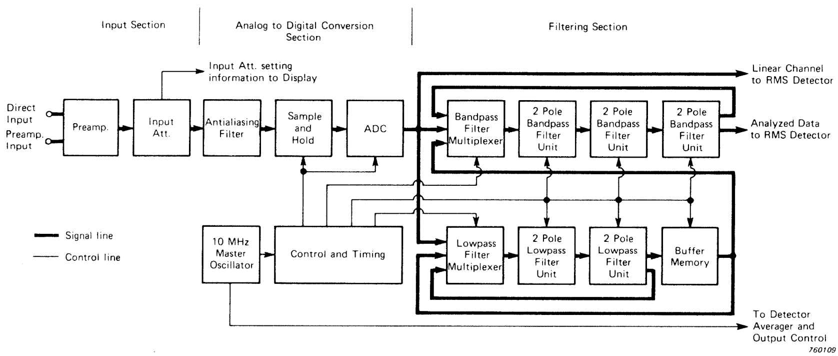


Fig.7. Block diagram of input and filter section of 2131

Each sample coming from the A/D converter is passed simultaneously through a 1/3-octave bandpass filtering section and a lowpass filtering section. In fact, each sample is passed through each section three times for the following reasons:

a) 1/3-octave filtering — The 1/3-octave filter section consists of three 2-pole filter units in series and for each pass, coefficients are used which give a 6-pole Chebyshev filter of 1/3-octave bandwidth. For each pass, the filter coefficients are changed so as to obtain successively the three 1/3-octave centre frequencies in each octave (e.g. 20 kHz, 16 kHz and 12,5 kHz in the highest octave). These three filter characteristics are illustrated in Fig.8.

---

* Shannon's sampling theorem states that a sampled time signal must not contain components at frequencies over half the sampling frequency. "Aliasing" occurs when this condition is not met

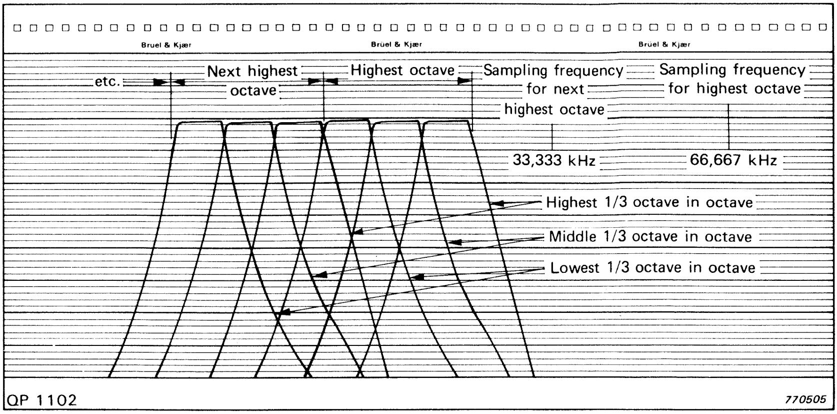


Fig.8. Filter characteristic vs. sampling frequency

b) Low-pass filtering — The low-pass filter section consists of two 2-pole filter units in series. Thus, during the three passes used to obtain the three 1/3-octave filtered values, it is possible to circulate the data value three times through the lowpass filter section, achieving 12-pole lowpass filtration (in this case, incidentally, a Butterworth filter was used). The cut-off frequency of the lowpass filter is one octave lower than the previous maximum frequency content.

The reason for the lowpass filtration is that it makes it possible to discard every second sample without losing any further information, i.e. once the highest octave in frequency is filtered away it is quite valid to use half the previous sample rate while still complying with Shannon's sampling theorem. These lowpass filtered samples with half the sampling frequency can now be fed back to the bandpass filter section, and since the filter characteristics are defined only in relation to the sampling frequency, the same filter coefficients will now give the three 1/3-octave filters one octave lower in frequency (see Fig.8).

In a similar manner, the same filtered samples can be fed back to the lowpass filter section and again filtered to one octave lower, once again allowing each second sample to be discarded, and so on.

It is possible to continue in this way for all lower octaves, thus obtaining the complete 1/3-octave spectrum. This explains the presence of the multiplexers at entry to both filter sections. These must keep track of where the next sample to be filtered is located and where the result is to be placed. It will be found that provided it is possible to process a sample from the A/D converter (i.e. in the highest frequency octave) plus one other sample in each sample period, it is possible to produce a parallel real-time spectrum for all octaves up to and including the highest. The limitation at the low frequency end is not in calculating capacity, only in being able to store the results, and so in the 2131 the frequency range is limited to a little over 4 decades. The reason why this is possible can be understood by reference to Fig.9. This is a table showing the order of processing, at least of the first few samples in each octave, both in the 1/3-octave section and in the lowpass filter section. The lowpass filtered values (with every second one discarded) are fed back to the next lower octave in both filter sections. Considering the number of samples to be processed in each octave, and calling the number in the 16 kHz octave M, the number in the 8 kHz octave is M/2, in the 4 kHz octave M/4, and so on. Thus the total number of samples to be processed in all octaves below the highest is M(1/2 + 1/4 + 1/8 + ...) = M, i.e. the same as the number of samples in the highest octave alone. Consequently, as shown in Fig.9, it is only necessary to process one data value from the highest octave plus one other in each sample period.

BANDPASS FILTER

SAMPLING PERIOD NUMBER

B.P. Filter Octave	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
16 kHz	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
8 kHz	1_1		3_1		5_1		7_1		9_1		11_1		13_1		15_1	
4 kHz		1_2				5_2				9_2				13_2		
2 kHz				1_3								9_3				
1 kHz								1_4								
500 Hz																1_5

LOWPASS FILTER

L.P. Filter Cut-Off Freq.																
12 kHz	1_1	2_1	3_1	4_1	5_1	6_1	7_1	8_1	9_1	10_1	11_1	12_1	13_1	14_1	15_1	16_1
6 kHz	1_2		3_2		5_2		7_2		9_2		11_2		13_2		15_2	
3 kHz		1_3				5_3				9_3				13_3		
1,5 kHz				1_4								9_4				
750 Hz								1_5								
375 Hz																1_6

Fig.9. Operation of 2131 digital filter unit

The operation of the analyzer can be converted to octave band filtering, once again basically by changing the filter coefficients. Since only one filter is to be calculated in each octave it is possible to recirculate the data values more than once through the bandpass filtering section. In the 2131, two passes are used (3 would have been possible) giving 12-pole filters of the Chebyshev type.

In a similar manner it is possible to perform for example, 1/12-octave analysis, simply by changing the filter coefficients. Since the 2131 can only calculate 3 filters per octave at a time it is necessary to make 4 passes, each time calculating a different set of three 1/12-octave filters in each octave (Ref.2). An alternative would have been to have limited the (real-time) frequency range to one quarter (5 kHz) to obtain the extra time necessary to make 4 times as many calculations in each octave. The major point, however, is that there is no limit to how far down in frequency one can go, from a given maximum real-time frequency, when the frequency axis is based on octaves.

As previously mentioned, the same filter coefficients will give exactly the same filter characteristics for each octave, simply by successively halving the sampling frequency. Within each octave, however, the characteristics are not in general identical. In the design, an attempt is made to obtain characteristics as close as possible to the equivalent analogue filters, which in general would have symmetrical characteristics on a logarithmic frequency scale. However, digital filters depart from the equivalent analogue filters in two ways:

1) Because of a periodicity introduced by sampling in time (Ref.2) the filter characteristic folds about the Nyquist frequency (half the sampling frequency) which has a non-uniform influence on the various filters within one octave (see Fig.10).

2) The folding is counteracted by the lowpass filter characteristic of the antialiasing filters, which thus have an influence on the overall characteristic. The lowpass filter is the same, however, for all filters in each octave and thus also has a different effect on their filter characteristics (see Fig.10).

It must be kept in mind that the deviations referred to above are very minor in nature, and in fact are only detectable from about —30 dB with respect to the reference level in the passband (Fig.8). The filters are well within the requirements of the most stringent standard specifications.

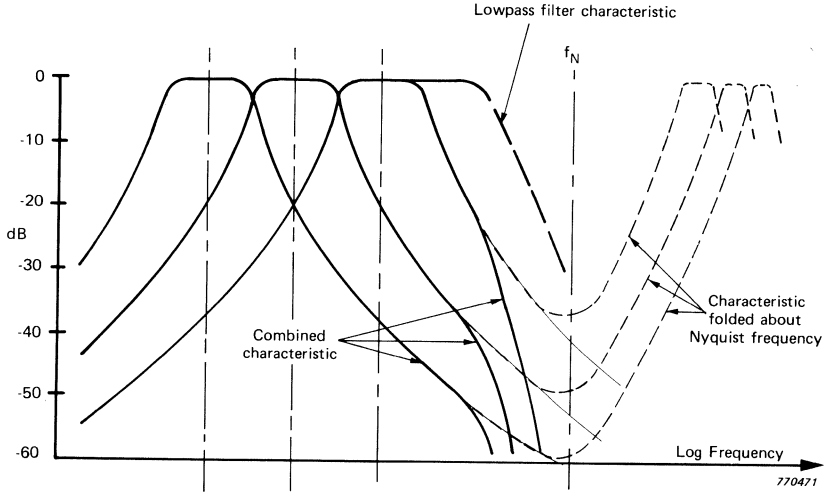


Fig. 10. Overall filter characteristics including lowpass filtration

While on the subject of digital filters, it should be pointed out that in theory it is possible to design filters having constant bandwidth on a linear frequency scale, but it is not then possible to time-share the hardware in the same efficient manner, and thus the FFT technique would normally be more economical. The reasons for this will now be discussed in detail.

The Fast Fourier Transform (FFT)

The FFT procedure, which was rediscovered by Cooley & Tukey in 1965 (Ref.3) is a particularly efficient way of calculating the so-called "Discrete Fourier Transform" (DFT). The latter may be interpreted as a finite, discrete version of the Fourier Integral Transform.

The formulas for the DFT are as follows:

$$G(k) = \frac{1}{N} \sum_{n=0}^{N-1} g(n) e^{-j\frac{2\pi kn}{N}} \qquad \text{(Forward Transform)} \qquad (1)$$

$$g(n) = \sum_{k=0}^{N-1} G(k) e^{j\frac{2\pi kn}{N}} \qquad \text{(Inverse Transform)} \qquad (2)$$

where G(k) represents the (complex) frequency spectrum at N different frequencies from zero to (just less than) the sampling frequency f_s. The frequency increment (Δf) between spectrum samples = f_s/N and $k\Delta f$ is thus the actual frequency corresponding to G(k). Likewise, g(n) represents the time signal at N different points from zero to (just less than) T the total record length. The time increment between samples (Δt) = T/N and thus $n\Delta t$ represents the actual time corresponding to g(n).

Looking at equation (1) it will be appreciated that the N values of frequency ($k\Delta f$) are equally spaced on a linear frequency scale, and it can easily be shown (Ref.2) that the resolution bandwidth is a fixed proportion of the line spacing (i.e. constant bandwidth). In general, the time samples g(n) may be complex, but since in practice they are normally real, it is usual to transform N real values as though they were N/2 complex, and manipulate the result (Ref.4) so as to obtain N/2 frequency components up to the Nyquist frequency (half the sampling frequency). This does not however, change the constant bandwidth, linear frequency scale nature of the analysis.

It is possible to convert an FFT analysis to constant percentage bandwidth on a logarithmic frequency scale, but it is important to realise the restrictions that this imposes on the results. Taking the typical case of a 400-line analysis, as shown in Fig.11, at the maximum frequency (line 400) the relative bandwidth is 1/400 = 0,25%. One decade lower (line 40) the bandwidth is 2,5%, and thus for example conversion to 3% bandwidth is only possible one decade at a time. Two decades lower (line 4) the bandwidth is 25% and thus already greater than 1/3-octave (23,1%). The above figures assume a bandwidth equal to the line spacing, which applies for flat weighting of the data. In the normal practical case (with continuous signals) the data are weighted by a weighting function such as Hanning, to improve the filter characteristic, but this at the same time increases the bandwidth by 50%.

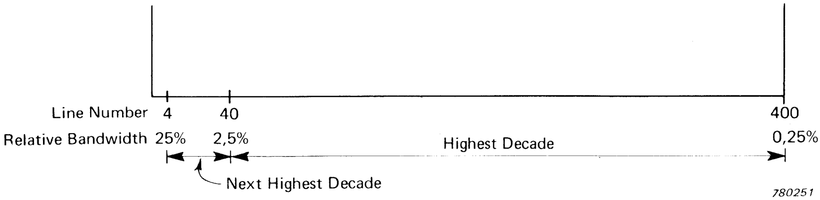


Fig.11. Relative bandwidth of a constant bandwidth spectrum

It is therefore common practice when converting to 1/3-octave bandwidth to do it $1\frac{1}{2}$ decades at a time. Let us look at what this means for the three decades 20 Hz — 20 kHz. First a spectrum to 20 kHz is generated, and converted for all frequencies above about 630 Hz. The time record corresponding to a single transform is 20 ms. It is then necessary to obtain a new spectrum to 630 Hz, which thus requires a thirty times lower sampling frequency and a thirty times longer time record. Even if averaging is performed on the linear spectra, before conversion, in order to improve statistical reliability, the same ratio of thirty must be maintained. Thus one only looks at the high frequency information 20 ms out of every 650 and the procedure is obviously not realtime. In practice, the time taken to make the conversion means that the total time to obtain a complete spectrum is even more than 650 ms, typically 2 s or so.

Converting only one decade at a time (such as is necessary with 3—4% bandwidth) means that the time ratio between highest and lowest decades is 100:1.

There is another detrimental effect of this conversion which has to do with filter characteristics. It will be seen from Fig. 12 that because the original filter characteristic is symmetrical on a linear frequency scale, the result of integrating a number of adjacent filters together will be also, and thus will be unsymmetrical on a logarithmic frequency scale.

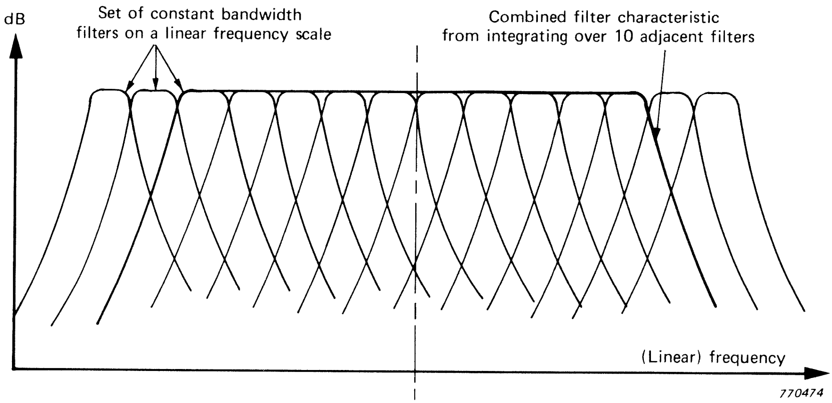


Fig. 12. Effect on filter characteristic of combining filters

Perhaps more importantly, over one decade the relative steepness of the flanks of filters obtained in this way will vary by a factor of 10:1 (and over $1^1/_2$ decades by 30:1). Thus, where two successive decades are fitted together there will be a sudden change of 10:1 in filter flank steepness (see Fig. 13) even though both may be within the tolerances specified for a particular filter class. It will be appreciated that this sort of variation in filter characteristic is at least an order of magnitude greater than that discussed in connection with digital filters in Section 4.

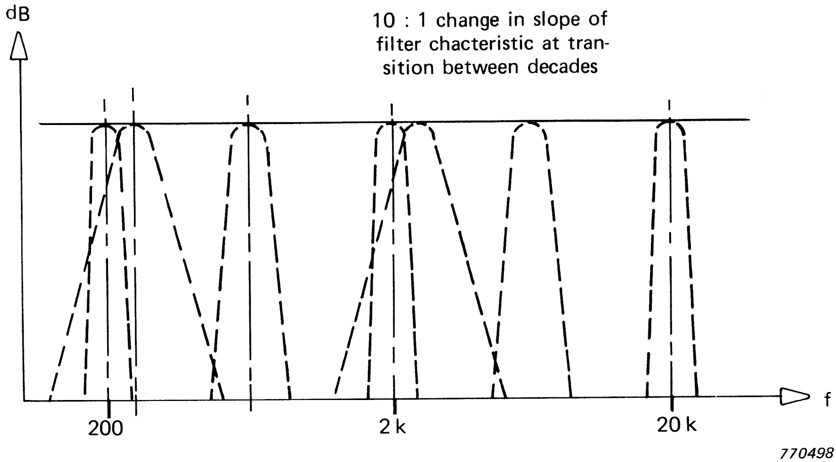


Fig. 13. Variations in filter characteristic with synthesized constant percentage bandwidth filters

There is one further point to watch when converting constant bandwidth spectra to constant percentage, in particular 1/3-octave. At the highest frequency in each linear spectrum it will be necessary to integrate approximately 100 lines together (actually 92,4 for 1/3-octave). Thus, the result obtained can be 20 dB higher than the original levels, and it is thus necessary to take this into account when judging linearity. The 2031 Analyzer, for example, has a linearity better than 0,1 dB or 0,01% of full-scale. The latter appears very small, and is in fact equivalent to —80 dB, but after integrating to 1/3-octaves the same accuracy would apply to —60 dB. This means that a 1/3-octave level measured at —60 dB could possibly be in error by as much as + 6 dB or —∞. This problem was in fact more prevalent with time compression analyzers, where the linearity specification was typically 0,1% instead of

0,01% and thus the whole problem is shifted up by 20 dB. It is also more likely with time compression than with FFT that the errors in linearity could occur over all the lines integrated into a 1/3-octave band.

In order to assist in assessing all the foregoing information some examples will be given of typical applications of the various techniques.

Examples where Digital Filtering Preferable
These include all the previously mentioned cases where real-time analysis of non-stationary signals in 1/3-octave bands is required, viz.

1. Aircraft flyover noise
2. Vehicle passby noise (1/3-octave)
3. Reverberation time measurements.

The latter will be discussed in some detail.

The reverberation time in a room/enclosure is the equivalent time taken for the smoothed sound pressure level to fall 60 dB (based on a linearized decay curve which may extend over less than 60 dB) on removal of the exciting signal. The same definition applies to individual 1/3-octave bands which are sometimes employed to give a better indication of the damping properties of the enclosure at different frequencies. Averaging is required to damp out random fluctuations (in particular when the exciting signal is random, such as pink noise), but the averaging time must be sufficiently short so as not to influence the results. Where exponential (i.e. RC) averaging is employed, as is normally the case, the detector level can itself not fall more than 8,7 dB per averaging time (Ref.2). Thus, so as not to limit the rate of fall of the measured signal (and allow some fluctuation above and below the mean line) the averaging time must not be greater than about 1/10 of the reverberation time.

For example, with the Analyzer Type 2131, it is possible to employ an averaging time of 31 ms and read out values at intervals of 44 ms. This permits the measurement of reverberation times down to about 0,3 s. It will be appreciated that the signal is then changing so rapidly that real-time operation is necessary. It was common in the past to make several measurements of the reverberation time and average them to obtain a more reliable result. Using a desk-top calculator in conjunction with the Digital Frequency Analyzer Type 2131 it is possible to employ another approach, where individual samples of the decay curve may be averaged to obtain a smooth result (Ref.5). This has the advantage that changes in slope can more easily be detected and measured (Fig.14).

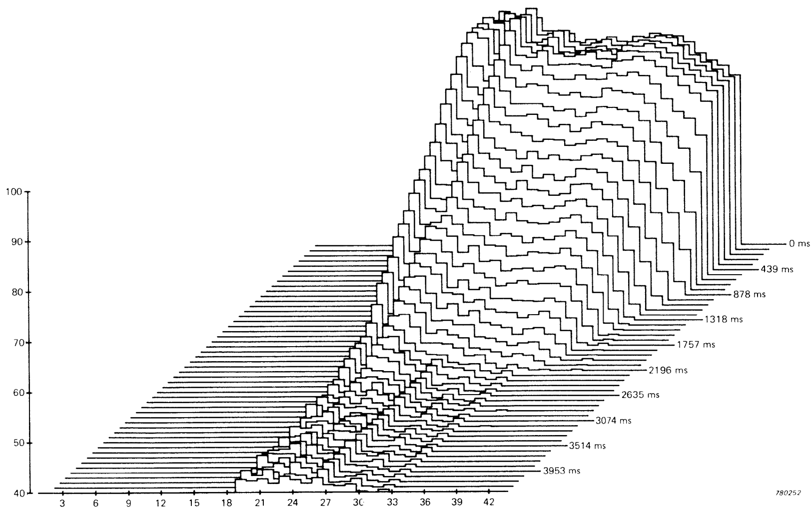


Fig.14. Typical time-frequency-amplitude landscape

Other cases where digital filtering is to be preferred, are 1/3-octave measurements of

1) continuous signals containing short impulses which might be missed by a non-real-time analysis method. Noise and vibration signals from slow speed reciprocating machines such as diesel engines and hydraulic pumps provide typical examples;
2) transients which are longer than the record length of an FFT analyzer which might otherwise be used. If an entire transient can be contained in the FFT analyzer memory (without losing important high frequency components) then the signal itself intrinsically contains less than 2 decades of frequency information and it will be valid to convert this single spectrum to 1/3-octaves over the highest two decades. When the transient cannot be contained in this way (which would often be the case with sonic booms, for example) it will be necessary to use a real-time 1/3-octave analysis method, e.g. digital filtering.

Examples where FFT preferable

These include all the previously mentioned cases where it is desired to detect equally spaced frequency components, e.g. harmonics, sidebands, and "interharmonics". The latter will be taken as a typical example:

Fig. 15 shows vibration spectra measured at the bearings of a rotating compressor before and after a maintenance shutdown. Before the shutdown only the low harmonics of the running speed (approx. 165 Hz) were evident, but after the shutdown not only a half-order subharmonic, but also the "interharmonics" of order $1\frac{1}{2}, 2\frac{1}{2}, 3\frac{1}{2}$ etc. are present. Sohre (Ref. 6) states that exact half-order components arise typically as a result of lack of tightness of mechanical assembly of (journal) bearing components. In the authors' experience such mechanical looseness ("rattling") tends to give rise to interharmonic components, also. Sohre points out the importance of distinguishing between this case and that of "oil whirl" where the frequency of the vibration is at just less than half the rotational speed (42 — 48%). Expressing the results on a linear scale with constant bandwidth immediately allows this distinction to be made purely from the pattern of the result.

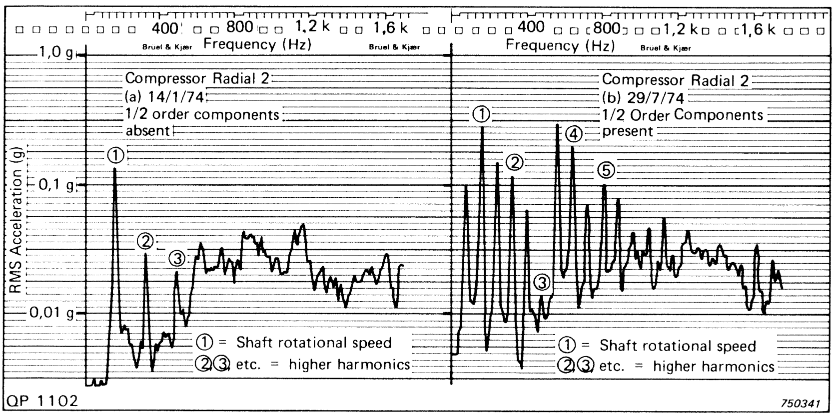


Fig. 15. Half order components presumably arising as a result of mechanical looseness

In general it can be said that diagnosis, in particular that based on discrete tones, is aided by having the spectrum expressed on a linear frequency scale with constant bandwidth.

A case in point is the evaluation of vehicle pass-by noise on a narrow band basis (e.g. Ref.7). This is not standardized for evaluation purposes, but is useful for diagnostic purposes in order to locate dominant noise sources.

Another case where constant bandwidth analysis on a linear scale is absolutely necessary (though not necessarily in real-time) is in cepstrum analysis. Cepstrum analysis is useful (Ref.8, 2) for echo detection and removal, voice pitch detection and speech transmission, and in the detection and separation of mixtures of families of harmonics and sidebands.

Examples where choice is difficult

1) Where it is desired to use a real-time analyzer for "machine health monitoring", the choice is not clear-cut. We have seen that for detection of faults in machines it was desirable to have constant percentage bandwidth on a log frequency scale, while for diagnosis of faults it was preferable to have constant bandwidth on a linear scale. However, the diagnosis would probably normally take precedence, particularly in view of the fact that the problems mentioned with respect to spectrum comparison for fault detection can be eased by:

   a) Using a tacho-signal to control the clock of an FFT analyzer to obtain an order analysis and thus reduce the influence of small speed fluctuations.

   b) Analyze the same signal in several different frequency ranges to cover the total range of possible faults.

   c) The above two measures are somewhat time-consuming, so as an alternative where many spectrum comparisons are to be made, it would probably be justified to obtain a desk-top calculator in order to make conversions to constant percentage bandwidth. The calculator could also be used then for making the comparisons automatically and thus reduce a lot of the drudgery associated with visual comparison. Note that this is one situation where synthesis of constant percentage bandwidth is justified, as the signals in question are stationary, and the importance of filter characteristic is not so great since the results do not have to be used for acoustical purposes; they only need to be repeatable.

2) Where the basic purpose of the analyzer is to aid in reducing the

noise produced by machines or other products there is also a problem. In order to evaluate and state the results correctly, it will usually be necessary to make 1/3-octave measurements satisfying the relevant acoustical standards on filter characteristics etc. On the other hand narrow-band analysis will be desirable for diagnosis of the source of troublesome noise components. The choice here depends on which is most important, the acoustical results or the diagnosis. For simpler machines, a 1/12-octave or 1/24-octave bandwidth analysis might be adequate for diagnosis in which case digital filtering would be the preferred method.

3) For continuous monitoring of the frequency spectra of a number of machines digital filtering would again be preferable because the real-time operation would permit the shortest possible analysis time for each monitoring point and thus maximize the number of channels surveyed by one instrument. It goes without saying that this monitoring is best carried out using constant percentage bandwidth over a wide frequency range, though once a fault is discovered it is best to have constant bandwidth analysis for diagnostic purposes. In large monitoring schemes it would be justified to have a separate FFT analyzer for such trouble-shooting, since the digital filter would be 100% occupied with the permanent monitoring.

4) Quality control by spectrum monitoring is a similar problem, but is complicated by the fact that there may only be a limited number of suspected faults, and thus it may be sufficient to monitor a constant bandwidth spectrum over a limited frequency range in which case the FFT analyzer may be the best choice. Once again, synthesized constant %-bandwidth would also be valid here since the signals are normally stationary, and it is only necessary to ensure repeatability. On the other hand digital filtering may be more economical because it produces the results more rapidly.

It is hoped that these examples will give an idea of the factors which typically play a role in the decision of which analyzer type is best for a given purpose.

References

1. RABINER, L. R. and GOLD, B.: "Theory and Application of Digital Signal Processing". Prentice-Hall, N.J., 1975.

2. RANDALL, R. B.: "Application of B & K Equipment to Frequency Analysis". B & K Publication, September, 1977.

3. COOLEY, J. W. and TUKEY, J. W.: "An Algorithm for the Machine Calculation of Complex Fourier Series". Math. of Comp., Vol.19, No.90, pp.297—301, 1965.

4. COOLEY, J. W., LEWIS, P.A.W. and WELCH, P.D.: "The Fast Fourier Transform Algorithm: Programming Considerations in the Calculation of Sine, Cosine and Laplace Transforms". J. Sound Vib. (1970), Vol.12, No.3, pp.315—337.

5. UPTON, R.: "Automated Measurements of Reverberation Time using the Digital Frequency Analyzer Type 2131". B & K Technical Review, No.2, 1977.

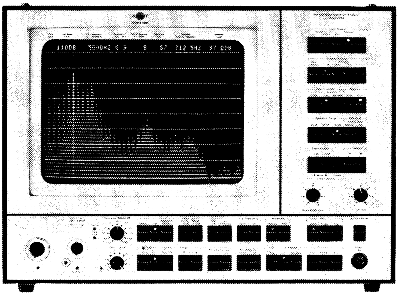
6. SOHRE, J. S.: "Operating Problems with High Speed Turbomachinery. Causes and Correction". Presented at ASME Petroleum Mech. Eng. Conf. Sept.23, 1968, Dallas, Texas.

7. CHUNG, J. Y. and WILKEN, I. D.: "A Technique for Measuring the Sound of a Moving Tyre". J. Sound Vib. (1977), Vol.55, No.1, pp.9—18.

8. CHILDERS, D. G., SKINNER, D. P. and KEMERAIT, R. C.: "The Cepstrum: A Guide to Processing". Proc. IEEE, Vol.65, No.10, Oct.1977, pp.1428—1443.

News from the Factory

Narrow Band Spectrum Analyzer Type 2031



The Narrow Band Spectrum Analyzer Type 2031 calculates a 400 channel constant bandwidth RMS power spectrum of continuous and transient vibration and acoustic signals. The analyzer uses the FFT algorithm to transform records of 1024 samples of the input signal into the frequency domain. The time taken to generate a single spectrum is less than 200 ms, giving real-time operation up to greater than 2 kHz. The frequency range of the instrument is pushkey selectable in a 1-2-5 sequence from 0 — 10 Hz to 0 — 20 kHz. Alternatively, by using an external sampling source, the frequency range can be made to float, allowing the 2031 to track phenomena of variable frequency, as in, e.g., order analysis.

The recording takes place in a transient recorder having two parallel memories. While the free running mode of the recorder is used for analyzing continuous signals, the triggered mode is used when analyzing intermittent or transient data. The triggering source may be internal with a level variable in 200 steps across the input voltage range or external from a TTL pulse. In the triggered mode, the after trigger recording setting can be adjusted between 0,0 to 9,9 record lengths in steps of 0,1 record lengths — a very useful feature for analysis of transient data. Also in the analysis of cyclic processes, it allows part of the signal to be gated and analyzed independently from the rest, e.g. the opening or closing of a valve in a machine cycle.

After Fourier transformation the data can be exponentially or linearly averaged over 1 to 2048 spectra. A hold max. mode is also available, allowing the maximum value occurring in each channel to be held.

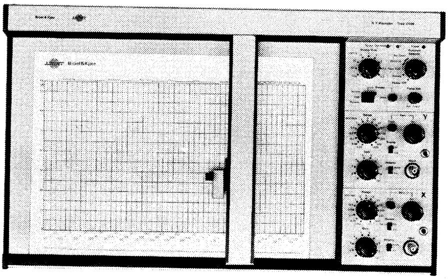
Both the instantaneous and the averaged spectrum can be displayed on an 11" calibrated screen which may also be used to show the time function. Either of these spectra can be stored in a reference spectrum memory and recalled later for comparison with new, incoming data. A further display mode gives the difference between the two spectra, whereby changes in spectra can be identified and transfer function magnitudes can be measured where sufficient stationarity exists. The frequency and RMS amplitude of any channel in the spectrum, or the relative position of any sample in the time function can be read from alphanumeric displays on the screen by using the line selector.

The amplitude display range of the 2031 can be set to 80 dB, 40 dB, or 20 dB. As the full scale level can be varied over 80 dB in 10 dB steps, a 40 dB or a 20 dB amplitude window from the displayed spectrum can be expanded to fill the entire screen. Similarly when a time function is displayed, the time axis can be expanded by a factor of 3.

Read-outs from the display can be plotted using a Level Recorder Type 2307 or (optional), an X—Y Recorder via the analog output, or digitally transferred to an IEC compatible peripheral via the IEC interface. The IEC interface also allows complete remote programming of the 2031 from an IEC compatible desk top calculator and input and output of all displayed data along with the complex spectrum and original (rather than displayed) time function.

Prior to recording and transformation, the input signal is passed through an antialiasing filter which is automatically selected with the frequency range. Before transformation the time signal may be weighted using the Hanning weighting to improve selectivity in the analysis of continuous signals.

X—Y Recorder Type 2308



X—Y Recorder Type 2308 is designed for fast, accurate, linear DC recording of slow and rapidly changing signals. Its fast slewing speed of 1000 mm/s (overshoot less than 1% of full scale) and maximum acceleration of 100 m/s ^{2} make it ideal for most laboratory work where hard copies of signal waveforms, Lissajous plots, frequency responses and analyses are required.

The writing system of the recorder accepts push-fit fibre-tipped pens and has a large writing area (185 × 270 mm) with electrostatic paper hold that firmly grips most types of paper up to A4 (DIN) size. The pen and the pen carriage are driven by separate low intertia, servo motors which are fully protected against excessive drive current and overloads.

Both the X and Y channels have high impedance floating inputs with normal and inverted input modes. 15 calibrated sweep sensitivities between 0,02 and 500 mV/mm may be selected which may be continuously adjusted between settings by separate potentiometer controls. The overall record linearity and accuracy obtained is better than 0,1% and 0,2% of full scale respectively.

The built-in sweep generator may be used to control the X or Y sweep of the recorder. It has 9 sweep rate settings between 0,2 and 100 mm/s plus "Forward", "Hold", "Reverse" and "Reset" modes. The ramp voltage output of the generator may be used to remotely tune voltage controlled types of frequency analyzers and signal generators for automatic recording of frequency responses and frequency analyses. Separate zero set and ramp set controls permit synchronous tuning with the X or Y sweep of the recorder and enable zero and full scale pen deflections to be set anywhere within the writing area.

For simple and straight-forward operation "Power Off", "Stand-by", "Paper Hold" and "Pen Drive" functions are on one multifunction control switch. The recorder may also be rack or bench mounted in either vertical or horizontal planes.

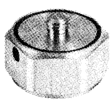
Calibrator for B & K Hydrophones Type 4223

This calibrator is a high level precision sound source which provides a rapid and easy method for the calibration in air of sound measuring systems which use B & K hydrophones as transducers. The principle of operation of the calibrator is that of a pistonphone in which a sound pressure is produced in the coupler cavity of the 4223 by four pistons which oscillate back and forth in phase. The frequency of the calibration tone is maintained within ± 2% of 250 Hz by a transistor circuit. The calibration accuracy obtained is within ± 0,3 dB.

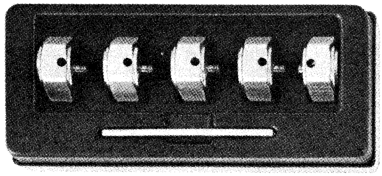
For connecting the different size B & K Hydrophones to the calibrator, three different couplers are included. The sound pressure levels produced in the coupler volumes by the 4223 are 159 dB, 162 dB, and 166 dB re. 1 μPa when used with B & K Hydrophones Types 8101, 8100 (and 8104) and the 8103 respectively. These high sound pressure levels enable accurate calibrations even in noisy surroundings. The sound pressure level in the coupler volume can be monitored with a 1/2" microphone which enables the calibration to be traceable to NBS. This facility also permits calibration of sound measuring systems which terminate with a 1/2" microphone when the dummy hydrophone supplied with the 4223 is placed in the coupler.

The calibrator is powered by six alkaline batteries, the condition of which may be checked by setting the control switch to the "Batt." position. This should result in a higher frequency tone than emitted in the "On" position (approx. 320 Hz with new batteries).

Mechanical Filter for Accelerometers UA 0559



UA 0559



UA 0553

When carrying out vibration measurements it may be desired to prevent the accelerometer from detecting high frequencies for one of the following reasons:

1. When low frequency vibration is measured (especially low levels) high level, high frequency acceleration may mask the low frequency components because of preamplifier overload, distortion, lack of electronic filters, etc.

2. When the resonant frequency of the accelerometer is excited by high frequency vibration (which may well be of insignificant level) wide band measuring errors and overload may occur. By using the mechanical filter UA 0559 the useful dynamic range of the measuring instruments can be increased by more than 20 dB.

3. To protect the accelerometer from damage if it is likely to be subjected to transient shock levels beyond its maximum capability.

By interposing the mechanical filter between the accelerometer and the test object, the transverse and main axis resonances (the Q factors of which are typically 30 dB in amplitude) are substituted by a highly damped resonance response of only 3 to 4 dB amplitude. The accelerometer's own axial resonance is suppressed by 25 to 30 dB and shifted up in frequency due to decoupling of the mass above the filter.

When a tunable filter is not available, the frequency range of interest can be extended by choosing an accelerometer with a lower mass, and reduced by adding mass loads on top of the mechanical filter respectively. Hereby a specific upper cut-off frequency can be achieved — e.g. when measuring vibration on human hand, arm, or body where the frequency range is normally limited to below 1000 Hz.

The stiffness and damping of the filter medium, butyl rubber, are dependent on ambient temperature, optimum damping being obtained between 20° and 50°C. A tommy bar included with the set can be inserted through the filter to relieve the rubber core of excessive torque when the accelerometer is tightened down.