METEOROLOGICAL OFFICE
9 FEB 1979
LIBRARY

LONDON, METEOROLOGICAL OFFICE.

Met.O.16 Branch Memorandum No. 1.

Smoothing and filtering (of meteorological data). By LEE,A.C.L.

London,Met.Off.,Met.O.16 Branch Mem. No.1,1977,31cm.Pp.10,21 pls.

An unofficial document - restriction on first page to be observed.

Y42. J2
National Meteorological Library and Archive
Archive copy - reference only

Met O 16 Branch Memorandum No 1

METEOROLOGICAL OFFICE
9 FEB 1979
LIBRARY
128741

SMOOTHING AND FILTERING
(Of Meteorological Data)

by A C L LEE

This document represents the opinion of the author, and does not necessarily reflect the official view of the Meteorological Office. It should not be quoted except as a private communication with the written permission of the author.

Smoothing and Filtering
(Of Meteorological Data)

Introduction

In signal processing there is a very common situation where a relatively slowly moving "signal" is corrupted by "noise" which has a wide spectral content.

A meteorological example might be an attempt to monitor a synoptic variable such as wind direction or temperature from a synoptic station. A single sensor will measure the synoptic variable with measurement noise superimposed by effects such as atmospheric turbulence. There will also be effects due to topography and the friction layer, but this is a separate issue which for the present purpose will be ignored.

For the above example, we might ideally have a dense network of sensors and use spatial averaging over a "synoptic" scale length to smooth out noise. Clearly this is impracticable; instead some form of "time averaging" or filtering is often used, the assumption being made that the "noise" has a zero mean, and that if we can average for long enough the residual noise will be small.

Unfortunately we cannot "average" for an arbitarily long time; the synoptic situation also changes and we are interested in following certain rapidly changing situations such as occur during the passage of a front. One possible viewpoint is that we are interested in variations below a certain freqency (the "corner frequency") which will be essentially contributed to by the varying synoptic situation. We wish to suppress variations above this frequency, as these will consist essentially of noise. This requirement can be implemented using a suitable "filter". The choice of "corner frequency" will be related to the synoptic scale length of the desired spatial averaging.

The actual choice of corner frequency is not the subject of this paper, but for synoptic purposes we are probably interested in being able to follow variations which have a period of the order of 10 minutes. In general "noise" will have a continuous spectrum, and for any corner frequency there will always be a contribution to the final signal from the low frequency noise in addition to any high frequency noise which may not be adequately suppressed. At the same time we wish to be able to follow the more rapid synoptic variation without undue suppression or other distortion.

It must be realised that filtering is not the same as taking an arithmetic "average" over some period of time. If the latter is specifically required, it defines a specific type of filter. If, however, the problem is left more general, in the sense that the objective is to monitor synoptic variations while suppressing "noise" which has a zero average, this leaves more flexibility in the design of the filter. This extra flexibility can result in a filtered signal which more nearly approximates the "synoptic" signal than does the arithmetic average.

In the following sections the generally accepted "arithmetic average" will be compared with various types of polynomial filter. It is hoped to demonstrate that the polynomial filter can be superior to the arithmetic average, and that it admits of easy implementation in hardware or efficient computation in software.

Polynomial Filters

In recent years a class of mathematical filter known as "Polynomial Filters" have been well documented. They are so called because the frequency-domain relationship between the input and output signals can be expressed as a polynomial in the Laplacian operator s (s being a complex frequency). This is useful because most desirable signal transmitting media have a frequency dependence which can be closely approximated by a polynomial in s, and thus a mathematical polynomial filter has a close correspondence with its physical realisation.

A good deal of attention has been focussed on obtaining the optimum performance (in some carefully defined manner) from low order polynomial filters, which correspond to simple physical realisations.

Some examples of physical realisations of n-order filters are shown in Fig 1. These are intended to demonstrate the relatively simple hardware, or small number of multiplications per timestep in digital software, required for implementation. For further information see Met.0.16 Branch Memos 2, 3.

Figs 2-11 describe the characteristics of several low-pass types of polynomial filter. Each has been normalised in frequency such that the "corner frequency" (defined here as having 3dB attenuation) occurs at 1 radian/sec. The most widely known polynomial filter is the Butterworth, whose characteristics are shown in Figs 2-4. Fig 2 show the filter attenuation in dB against frequency on a log scale for several orders (n=1 to 10) of filter. Note the change in scale at $\omega=1$. The coefficients in the Butterworth polynomial Transfer Function have been arranged so that all of the derivatives of the Transfer Function with respect to frequency, except the "n"th, are zero at $\omega=0$. In this sense the Transfer Function is maximally flat at low frequencies, while at high frequencies the Transfer Function is dominated by the highest power of s in the Transfer Function.

The Butterworth is an excellent filter for separating sinusoidal signals at low and high frequencies. Unfortunately, the time by which the signal is delayed on passing through the filter (the Group Delay) varies with ω, and this applies to frequencies below the corner frequency. The distortion which this produces on non-harmonic signals can be clearly seen in Fig 4 which shows the filter response to a Dirac Impulse (of vanishing small width and unity area), and to a unit step. The filter "rings" for a considerable time.

The impulse response is of particular interest in filter theory. If we have a time series S(t) passing through a filter with impulse response I(t) it can be shown that the response of the filter R(t) to the time series is given by

$$R(t) = \int_{\tau=0}^{\tau=t} S(\tau)\cdot I(t-\tau)\cdot d\tau$$

or in other words the impulse response shows how the previous history of an input time-sequence is weighted to produce the output time-sequence.

Figs 5, 6 and 7 show the characteristics of a particular Chebychev filter. In this filter a considerable increase in the rate of change of attenuation near the corner frequency can be achieved at the expense of allowing a defined amount of ripple (here 0.5 dB) in the passband (below the corner frequency). As might be expected the group delay and impulse response are even worse, from our point of view, than the Butterworth.

The Butterworth and Chebychev filters are optimised in the frequency domain. We are primarily interested in following a time sequence, i.e. in the time domain. Figs 8, 9 and 10 show the characteristics of the Thompson (or Bessel) filter. This has coefficients arranged to give a maximally flat group delay, which it achieves at the expense of a much more gradual increase of attenuation with frequency. This filter is extensively used in radar where the shape of a time series must be undistorted, consistent with removing high frequency noise. The impulse and step responses underline the desirability of maximally flat delay.

Just to show that the story does not end here, Figs, 11, 12, and 13 show the characteristics of a linear phase with equiripple error filter. This is similar to the Bessel in that flat delay is aimed for, but in this case a defined ripple is allowed in the delay and the extra freedom is used to increase the rate of attenuation near the corner frequency. This means that high frequency noise is more effectively suppressed, and this may be more significant than the small amount of distortion due to variable delay in the passband. This concepts leads naturally to Optimal Filtering, aimed at neither frequency nor time domain characteristics, but rather at minimising the R.M.S. difference between the noise-free signals and the delayed filter output in some carefully defined sense. Contributions to R.M.S. error occur due to both variation in group delay and noise. Thus the filter design depends on a knowledge of the absolute noise spectrum, on the signal spectrum, and on a knowledge of the relative importance of different frequencies in the signal. Discussion of these filters is outside the scope of this paper, and the expected improvement over the first three filters discussed is likely to be small.

The Arithmetic Average

Where some sort of smoothing is required, it is common practice in meteorology to form an arithmetic average of the measured signal for some finite time, eg a ten-minute average wind direction or wind speed. Some measuring devices will integrate a signal, present the integrated values, and re-start a new integration. This form of sampled output is difficult to compare with a continuous filter as the results will in general be different, and there is no continuous trace to see how representative the samples are of the previous 10 minutes. Instead a filter will be considered which continuously presents the result of an arithmetic average over the previous T seconds.

Response to an harmonic input is:-

$$R(\omega) = \frac{1}{T}\int_t^{t+T} \sin(\omega\tau+\delta)\cdot d\tau = \frac{1}{T}\left[-\frac{1}{\omega}\cos(\omega\tau+\delta)\right]_t^{t+T}$$

$$= \frac{1}{\omega T}\left(\cos(\omega\tau+\delta) - \cos(\omega\tau+\delta+\omega T)\right)$$

$$= \frac{1}{\omega T}\left(2\sin\left(\omega\tau+\delta+\frac{\omega T}{2}\right)\cdot\sin\left(\frac{\omega T}{2}\right)\right)$$

The factor $\sin\left(\omega T+\delta+\frac{\omega T}{2}\right)$ represents the original harmonic input with a phase delay of $\frac{\omega T}{2}$, or a group delay of $\frac{T}{2}$ seconds at all frequencies. The remaining factor $\frac{2}{\omega T}\sin\left(\frac{\omega T}{2}\right)$ represents the amplitude response as a function of frequency. At low frequencies this tends to unity. For compatibility, choose $T \simeq 2.78$ seconds, so that we have 3dB attenuation at $\omega = 1$. Figs 14, 15 and 16 show the characteristics of this filter. The attenuation characteristic is of particular interest here. In the stop-band the attenuation peaks, which go to infinity, occur at frequencies whose period is an exact sub-multiple of the integration period. Apart from these peaks, the general level of attenuation is rather low, increasing at a rate of about 12dB per octave, or 20dB per decade. This means that high frequency noise, which usually contains most of the noise energy, is poorly suppressed compared to any of the previous filters. Fig 15 illustrate the highly desirable constant group delay. Fig 16's Impulse response indicates how the previous history had been uniformly weighted, and the step response gives an indication of how the filter integrates changes.

It is instructive to compare this response with the Bessel filter (Figs 8, 9, and 10). For a Bessel filter with order above 3-4 the passband frequency and group delay characteristics are remarkably similar to the Arithmetic Average Filter. This means that for signals completely within the passband - essentially the synoptic variations - both filters will give remarkably similar outputs, once the different group delays have been accounted for. In the stopband the Bessel is clearly better able to attenuate broad band noise, especially as the order number is increased. The Arithmetic Average appears to have an edge in the time domain as its group delay remains flat to the highest frequencies, but this is of little practical advantage as the higher frequencies are strongly attenuated.

Rather than comparing characteristics in the frequency domain, it is useful to compare them in the time domain. One approach is to compare the step responses. These are obviously not identical, but if we wish to "integrate" activity over some period - say 10 mins - we can match a straight-line approximation of the step response to the ideal straight-line step response of the Arithmetic Average. The final filter will, in fact, include contributions from outside the 10 minute period, but will have a slightly higher corner frequency than the matched passband filter described above (factor $\simeq$ 1.7). Thus it will have a smaller attenuation than the Arithmetic Average in the "passband", and yet retain higher attenuation over most of the "stopband". The important consideration is not necessarily the precision of the "10 minute period", but the ability to distinguish between signal and noise.

In many situations the different absolute group delays of the various Bessel filters and the Arithmetic Average will be unimportant as they are precisely known. However, in a forecasting situation it may be desireable to have the minimum delay between the actual situation and the presentation of the measurement at C.F.O. All forms of filtering involve a group delay (which may be much less than the communication delay). An example might be:

"10-minute" average of Met. Variable

Group Delay in minutes

Arithmetic Average	: 5
4-Pole Bessel, to have similar passband response to above	: 15
6-Pole Bessel, as above	: 19
4-Pole Bessel, to have "10 minute" integration	: 8.7
6-Pole Bessel, to "10 minute" integration	: 11

It is interesting that current meteorological practice is to measure "instantaneous" parameters such as pressure and filtered parameters such as wind direction at the synoptic time, although in fact the filtered signal is the best estimate available of the situation one group delay earlier.

The Exponential Average

One form of easily implemented filter that has been suggested for smoothing is the exponential average, so called because of its convolution integral:

$$V_{OUT}(t) = \frac{1}{\tau_0}\int_0^t V_{IN}(t) \cdot e^{-\frac{(t-\tau)}{\tau_0}} \cdot d\tau$$

The transfer function can be expressed in complex from

as $$R(s) \equiv \frac{V_{OUT}(s)}{V_{IN}(s)} = \frac{1}{(s\tau_0+1)}$$

and the filter corresponds to smoothing with a single real time-constant τ_0.

The characteristics of this filter correspond to a polynomial filter of order 1, and are plotted under the characteristics for Butterworth, Chebychev, Bessel etc filters described earlier. The plots are all identical as the single coefficient allowed is sufficient only to determine the corner frequency. It can be seen that this particular filter does not achieve even approximate constant group delay near the corner frequency, and its attenuation is worse than the Arithmetic Average in the stopband. The impulse and step responses are not shown, but are represented by exponentials decaying with time-constant τ_o to 0 and 1 respectively.

Some Effects of Sampling

One of the usual reasons for filtering data is that one wishes to record or transmit one value for the variable once every (defined) time period, where the intention is that the series of measurements thus obtained should be reasonably representative of the original data up to as high a frequency as is consistent with the limited time resolution. One important criterion is that a time series with samples spaced at uniform intervals of τ seconds cannot resolve any frequencies higher than the Nyquist frequency $N = 1/(2\tau)$ Hz. If a higher frequency f is sampled, the resulting time series will define a frequency f-N with the same amplitude as f, and the higher frequency f is said to have been "aliassed" into a lower frequency. This important result means that if we are to monitor some variable and report on its progress by means of one measurement every τ seconds, then the time series must first be filtered with a low-pass filter. Attenuation above the Nyquist frequency must be adequate for energy at these frequencies to be reduced in intensity to sufficiently below the signals in the passband, so that the alliassed contribution to the lower frequencies will be negligable. If one wishes to preserve the waveshape (as well as merely the energy content) a Bessel filter, or similar, will be used. This has a somewhat gradual increase in attenuation with frequency, and care must be taken in determining the ratio of Nyquist to corner frequency.

One method of implementing a filter is to use a digital processor. This involves sampling and digitalizing the input data stream, performing a mathematical manipulation corresponding to the filter, finally producing a filtered data stream. Because of the C.P.U. time involved in "filtering" the data stream, it is desirable not to sample the input stream too frequently. Provided that there is some limitation on the bandwidth of the input stream, the Nyquist frequency can be chosen as the lowest that will alias negligible energy into the input stream. If the sampling frequency (and hence the Nyquist frequency) is chosen too low, then we are no longer receiving a stream of data whose high frequency noise components can easily be filtered out with rejection of the order of 100 dB. Instead the high frequency components alias noise into a broad spectrum which includes low frequencies. This can easily lead to a situation where the aliassed noise can only be reduced by a factor proportional to the square root of the number of samples in the impulse response, and it becomes difficult to reduce this noise by much over 20-30dB.

Some Examples of Filtering

The above description illustrates the potential superiority of a Bessel Filter over a Moving Mean mainly in the Frequency Domain. However, data is generally used in the Time Domain. The result of filtering a time series depends, of course, on the time series itself; nevertheless a useful "feel" for the effects of a filter can be obtained by studying real data.

Fig 17 shows the results of filtering temperature data from a rapidly responding probe installed at Cardington. Temperature was measured every 0.1 seconds, and the top curve shows 1024 points (102.4 seconds) of raw data joined by straight lines. The overall frame height corresponds to 2.4°C. The middle curve shows the same data filtered by a Moving Mean filter having an integration period of 6 seconds (approximately 1/17 of the frame width). The bottom curve shows the raw data filtered by a 6-Pole Bessel filter having the same corner frequency as the 6-second Moving Mean. Curves (a) and (b) are shifted horizontally and vertically with respect to (c) to allow for Group Delay, and to avoid overlapping of the curves. Curve (c) commences at time zero, and the extent of the horizontal shift for (a) and (b) can be seen from their initial flat portions.

The effects of filter initialisation and Group Delay can be seen in the first 10-seconds or so of data. Because data is delayed on filtering, the first few seconds of filtered output do not correspond to real data, but rather depend on the initial values present in the filter stores. For cosmetic reasons these have been chosen to be near the expected raw data values, and this is the way to minimise the initial transient filter response. Where data is being measured continuously (e.g. surface temperature) the transient response occurs only on first installation and does not cause problems.

Because of its poor high-frequency attenuation, the Moving Mean filter exhibits structure on a time-scale significantly shorter than 6 seconds. As a consequence, values on the curve intended to give a measurement representative of the previous 6 seconds or so are sensitive to the exact time at which the value is chosen, i.e. the reported signal is noisy. If the intention is to send information on the shape of the filtered signal, the presence of the unwanted higher frequencies will make coding difficult.

A highly subjective way of comparing the filters is to decide which is the closest to the curve one would draw by eye through the data when "smoothing" over a 6-second interval.

Fig 18 shows curves derived from Cardington wind speed data. The overall width represents nearly 3 hours of data (4096 points at 2.5 seconds/point), and the overall height is 12 meters per second. The original data was recorded from a fast sensor at 0.1 second intervals. This would have shown excessive high-frequency variations on the diagram which would only have hindered interpretation by masking the lower frequencies. The Cardington data was thus filtered to the point where the time-series could be adequately represented by points spaced at 2.5 seconds (using a 6-Pole Bessel filter with corner frequency 0.36 radians/sec, or 1/f = 17.5 seconds). The filtered Cardington data was then used as "raw" data in Fig 18, and was again filtered by a 10-minute Moving Mean filter (approximately 1/17 of the frame width), and by a 6-Pole Bessel filter having the same corner frequency as the Moving Mean. The differences between the Moving Mean and Bessel Filters can again by seen.

Unfortunately the Cardington data only lasted for 1.78 hours. Fig 19 shows the same raw Cardington data, but with the time-scales expanded by a factor of 5, and all filters (including pre-filters) having corner frequencies increased by a factor of 5 giving 2-minute mean values.

The particular examples shown above are for the absolute value of a function. In meteorology one is often interested in the first derivative of functions such as the ones above. This is analogous to being interested in the shape of the contour on a chart rather than their absolute values.

Examples of a potential interest in time derivatives might be:

(a) Looking for variations in wind direction or speed, or in temperature, to denote the passage of a front.

(b) Use of Isallobars, being contours of derivatives of pressure. The shape of the contours are a function of the second derivative of pressure.

(c) Estimating wind speed and direction from the differential (or derivative) values of the radar range, azimuth and elevation of a balloon.

In all these examples the process of differentitation accentuates the higher frequencies and makes the filter rate of attenuation more important. In the time domain (as in Fig 17, and more obviously in Figs 18, 19) the Moving Mean curves are very much rougher than the corresponding Bessel curves, so that derivatives of temperature or wind speed (the tendancy of temperature or wind speed) will be much more noisy in the Moving Mean case than in the Bessel case. Similar comments have even greater force for higher derivatives. Thus the Bessel filter will help, for example, the bench forecaster to see the small chart-to-chart differences in wind direction with time indicating the passage of a weak front.

Meteorological Office

Ministry of Defence

February 1977

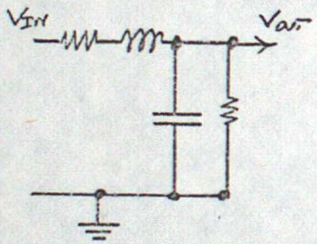
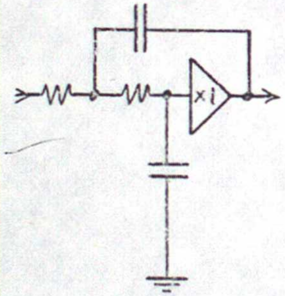
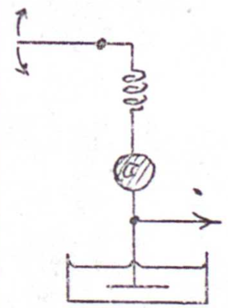
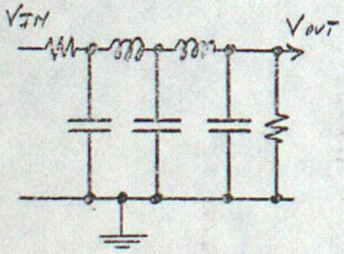
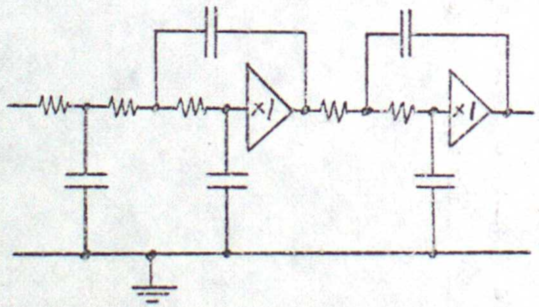
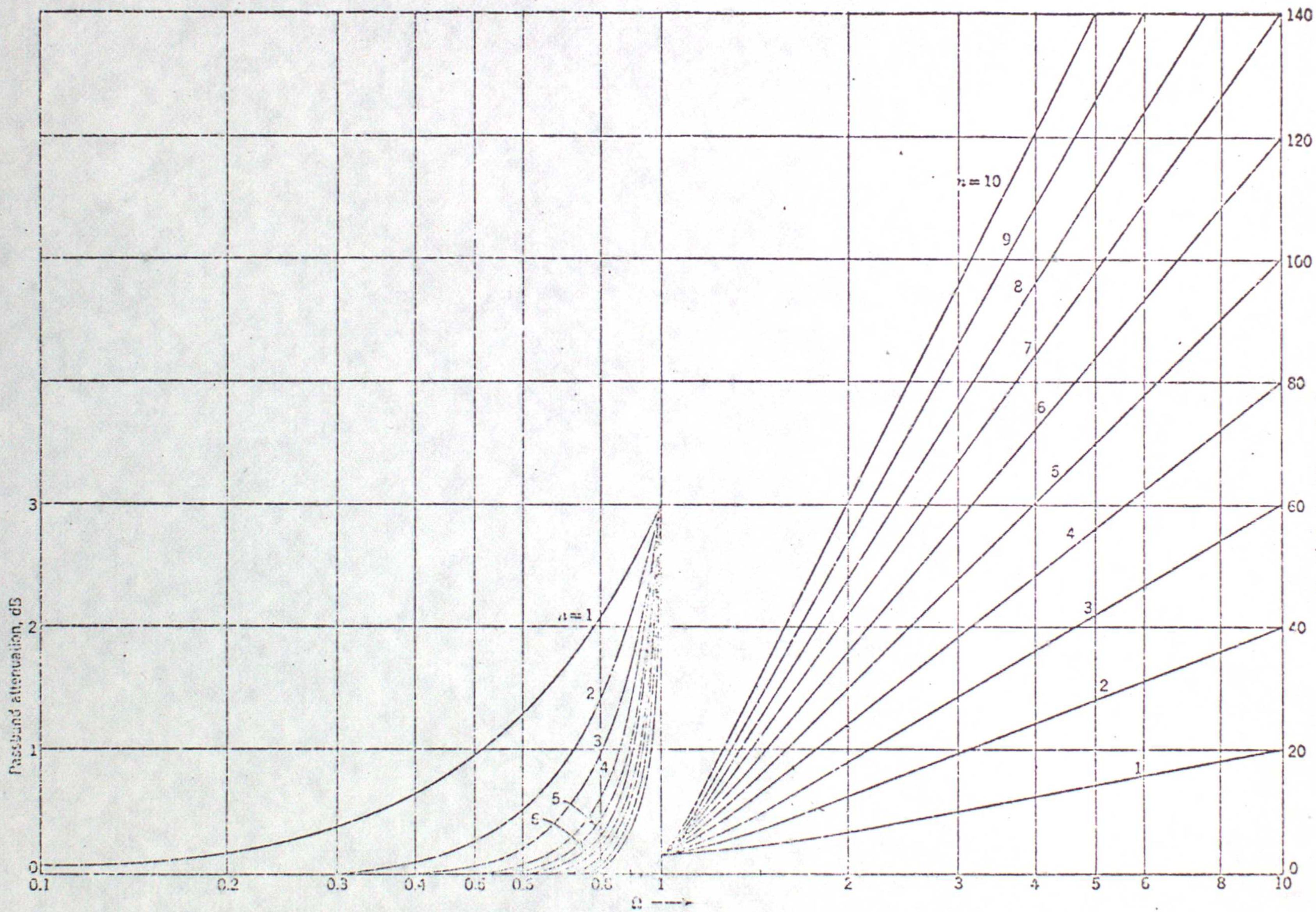
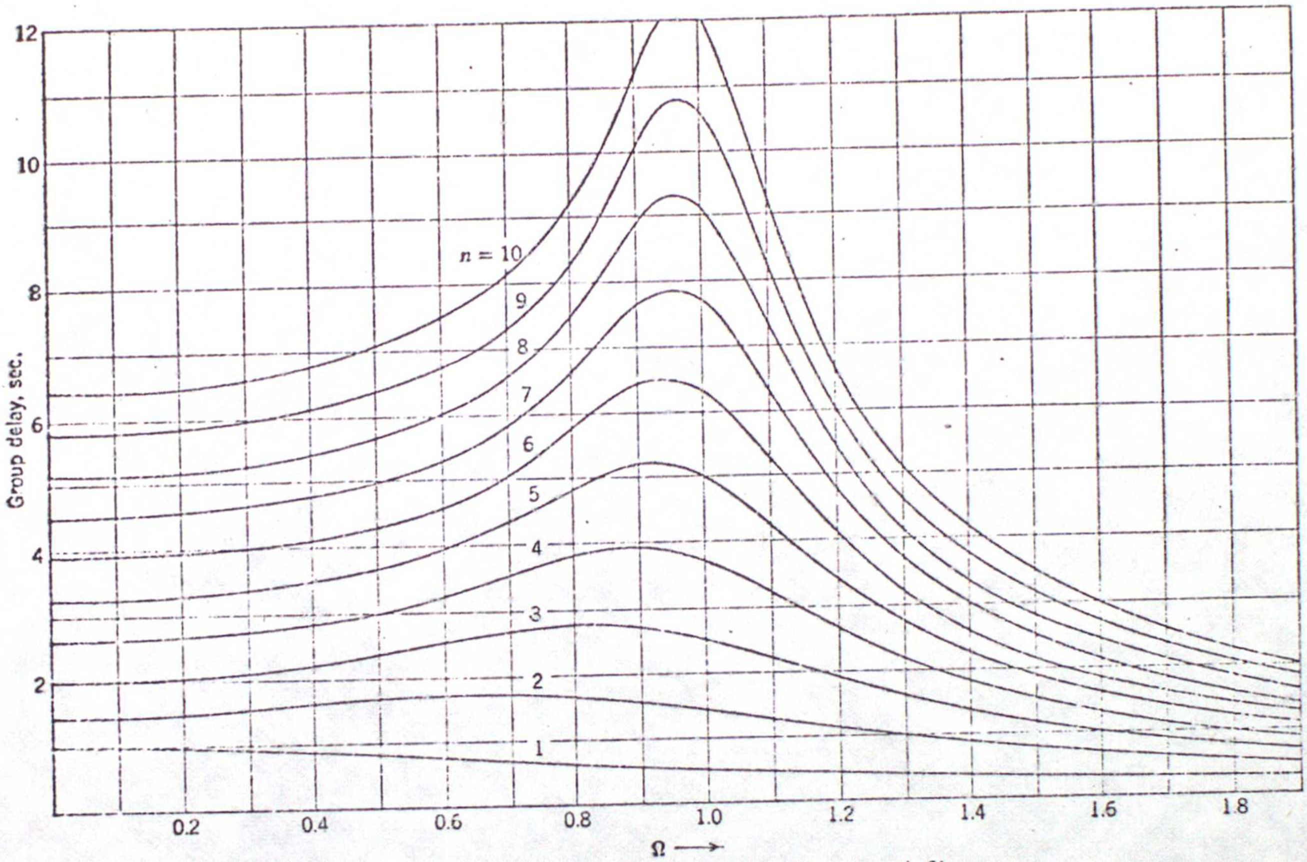
	Electrical Passive	Electrical Active (Sallen and Key Elements)	Mechanical	Transfer Function	Recursive Digital
2nd Order				$\dfrac{1}{(s+\alpha)^2+\omega^2}$	$V_{OUT}(n) =$ $\alpha_1 V_{OUT}(n-1)$ $+\alpha_2 V_{OUT}(n-2)$ $+\alpha_3 V_{IN}(n-1)$
5th Order			? ?	$\dfrac{1}{(1+\alpha_1 s+\alpha_2 s^2 + \alpha_3 s^3 + \alpha_4 s^4 + \alpha_5 s^5)}$	$V_{OUT}(n) =$ $\alpha_1 V_{OUT}(n-1)$ $+\alpha_2 V_{OUT}(n-2)$ $+\alpha_3 V_{OUT}(n-3)$ $+\alpha_4 V_{OUT}(n-4)$ $+\alpha_5 V_{OUT}(n-5)$ $+\alpha_6 V_{IN}(n-2)$

Fig 1- Some Filter Realizations

Fig 2.

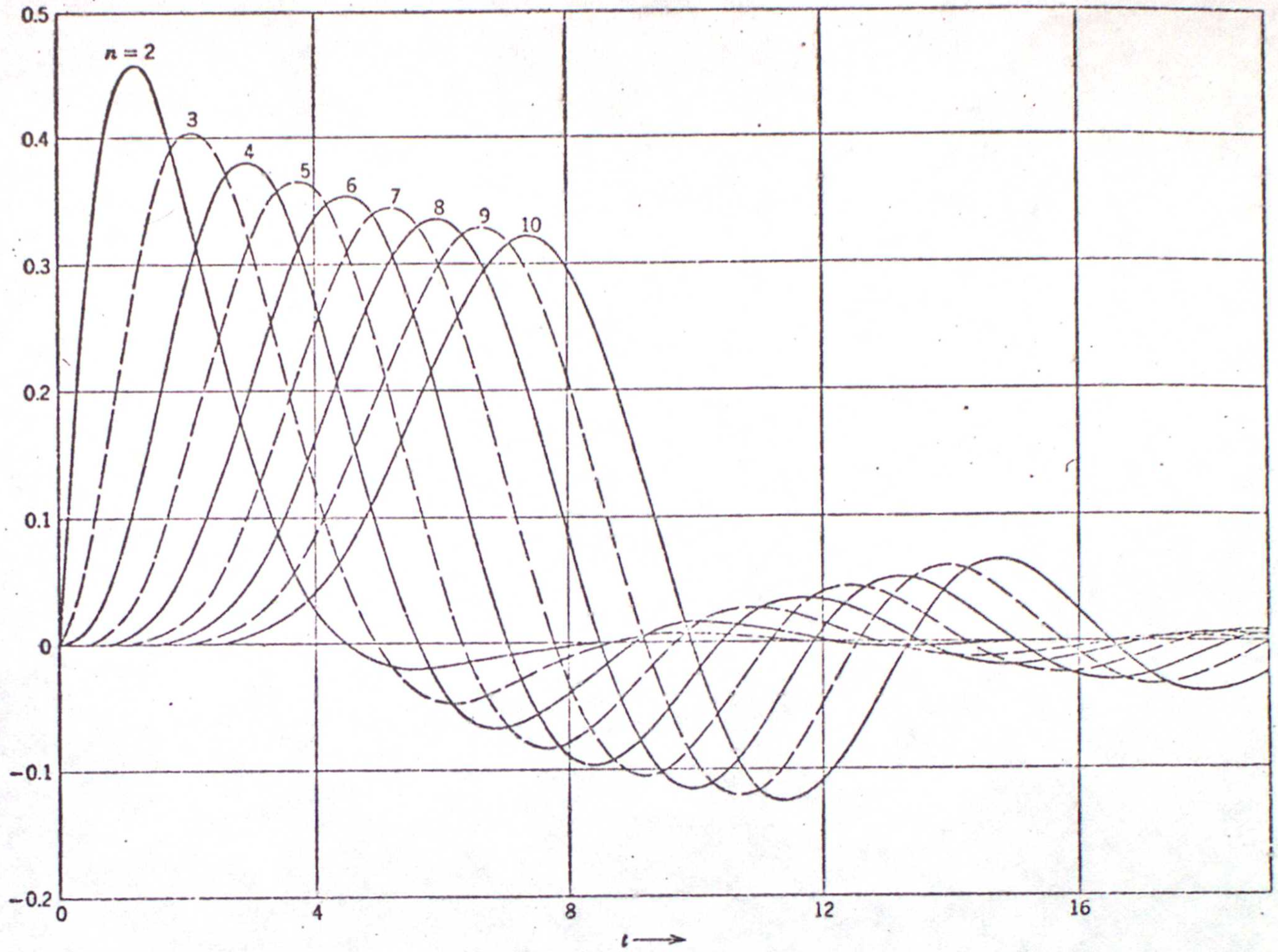


Attenuation characteristics for Butterworth filters.



Group-delay characteristics for Butterworth filters.

Fig 3



Impulse response for Butterworth filters.

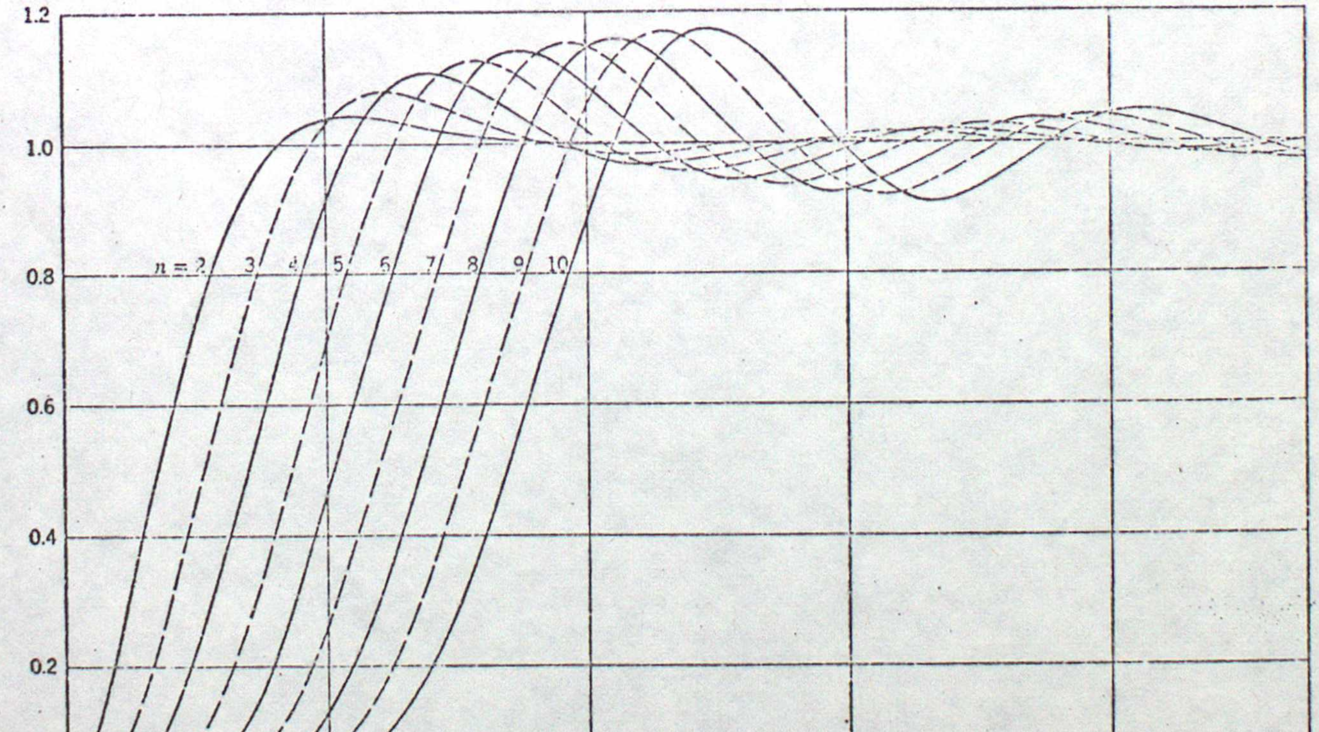
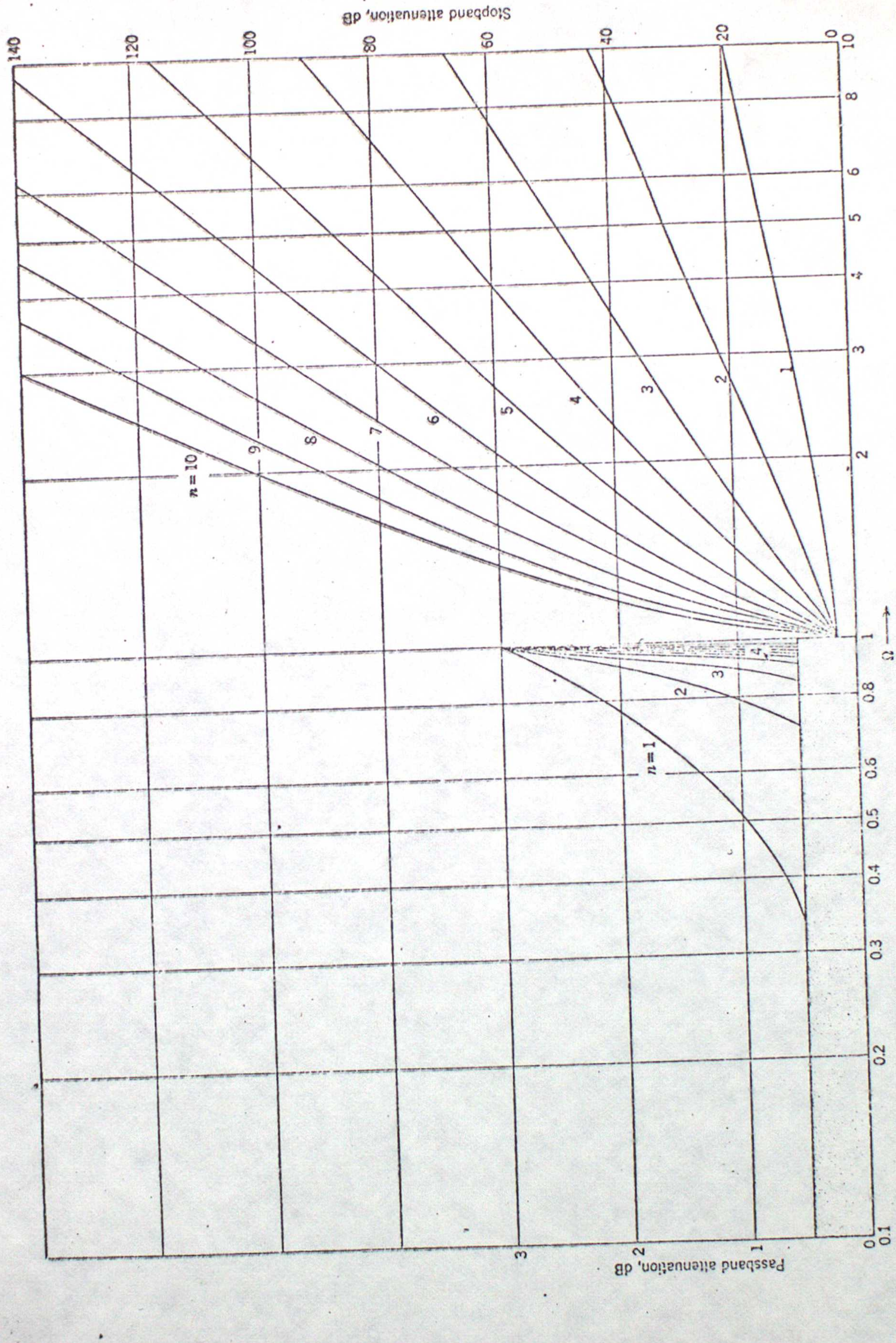


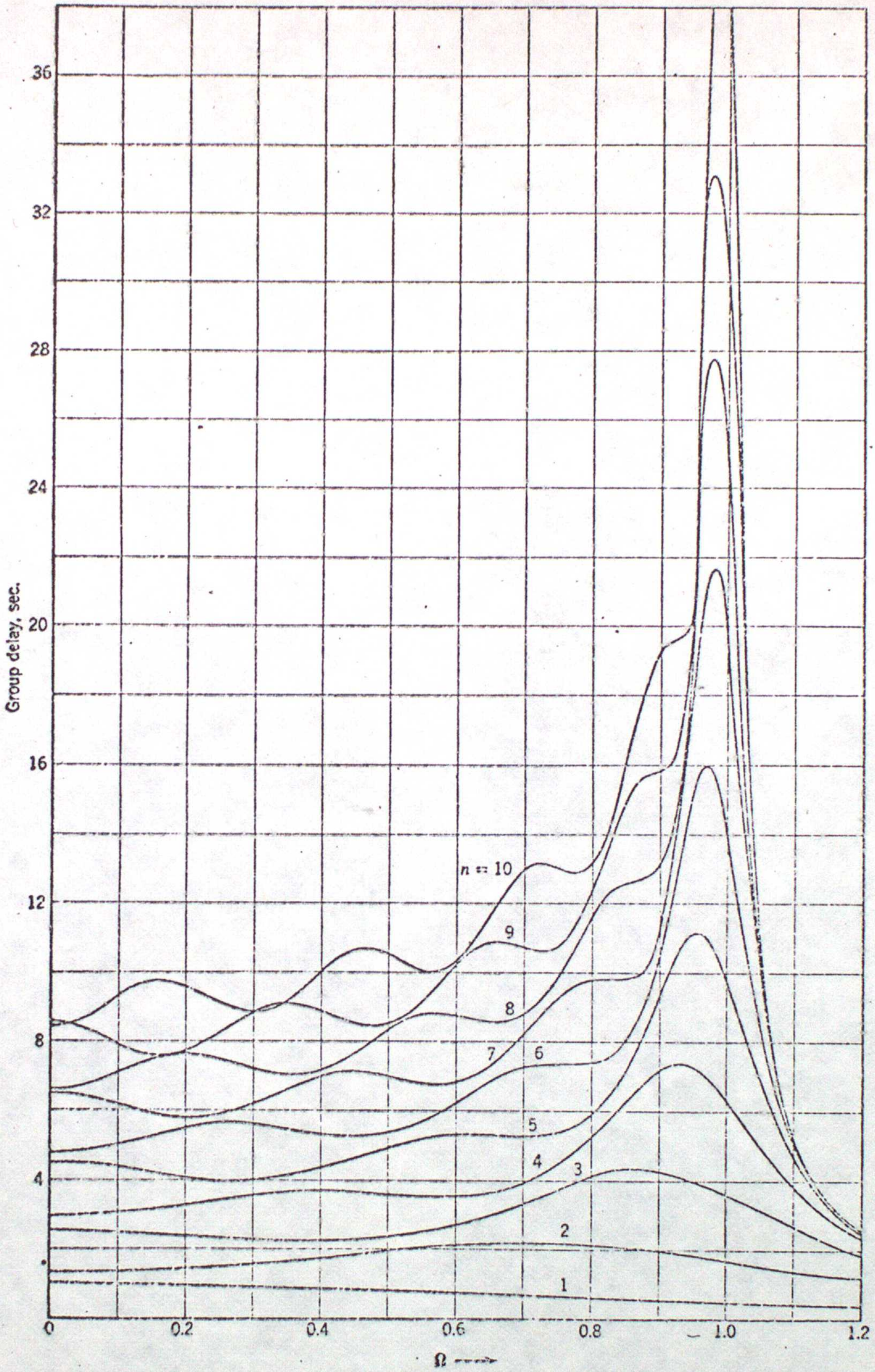
Fig 4.

Step response for Butterworth filters.



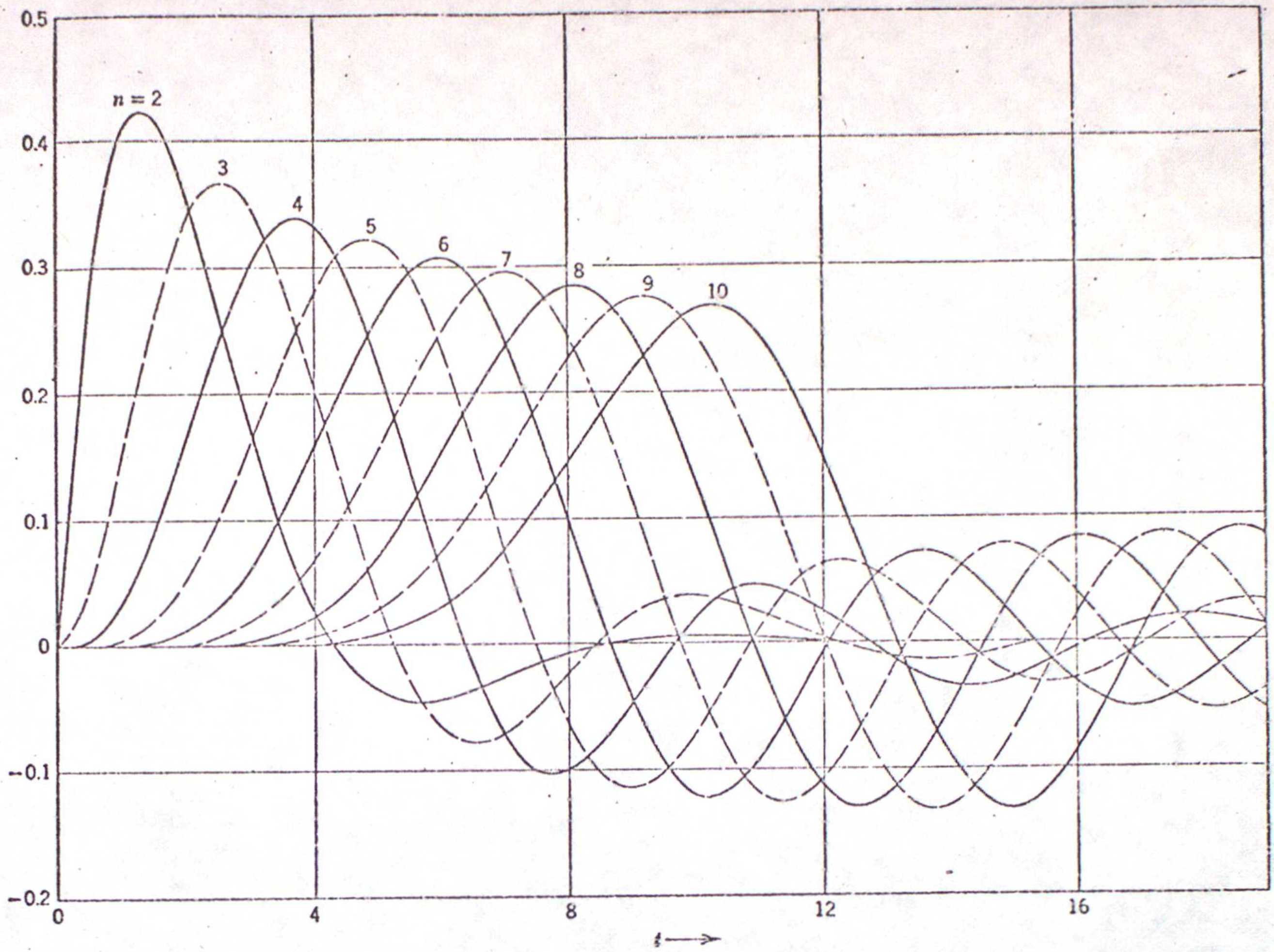
Attenuation characteristics for Chebyshev filter with 0.5 dB ripple.

Fig 5

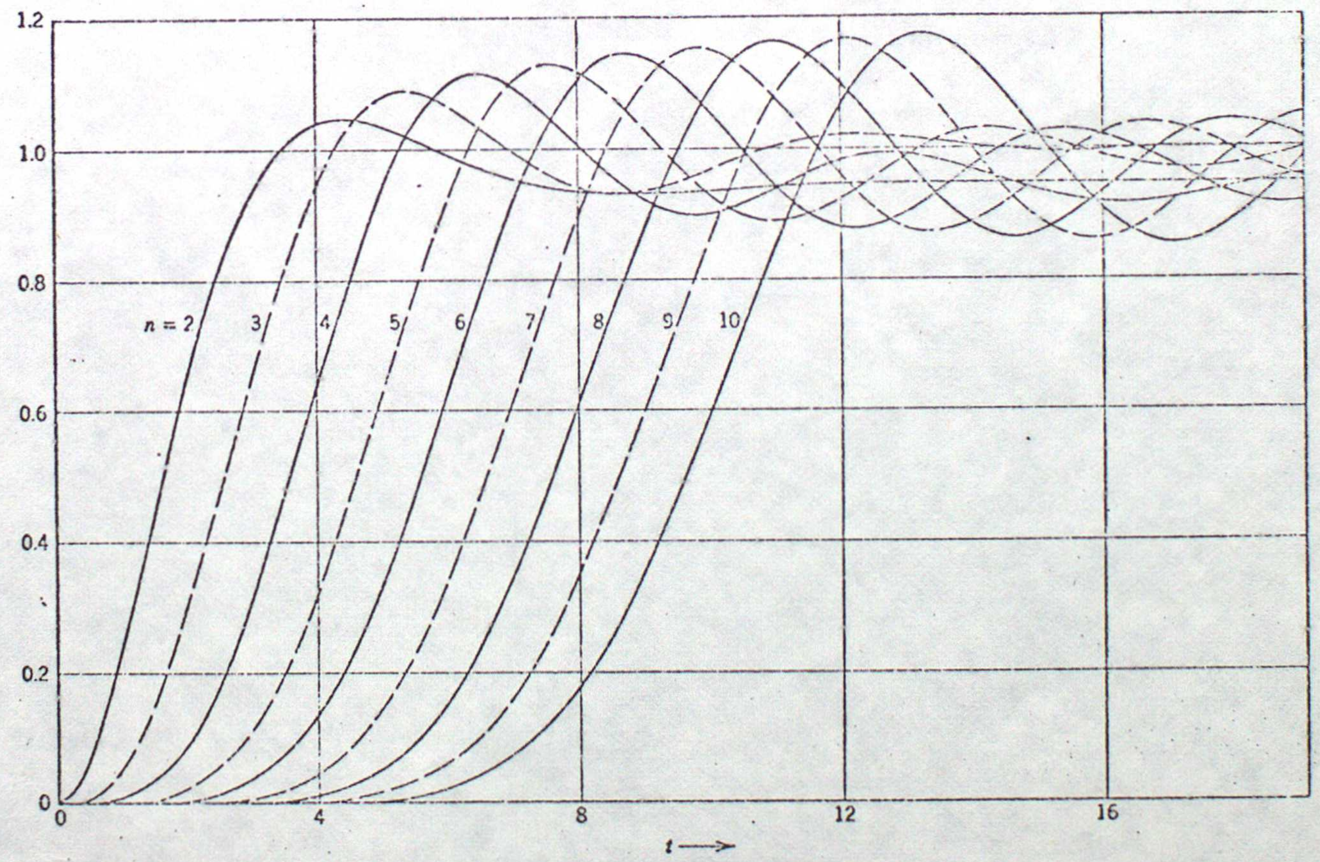


Group-delay characteristics for Chebyshev filter with 0.5 dB ripple.

Fig 6.

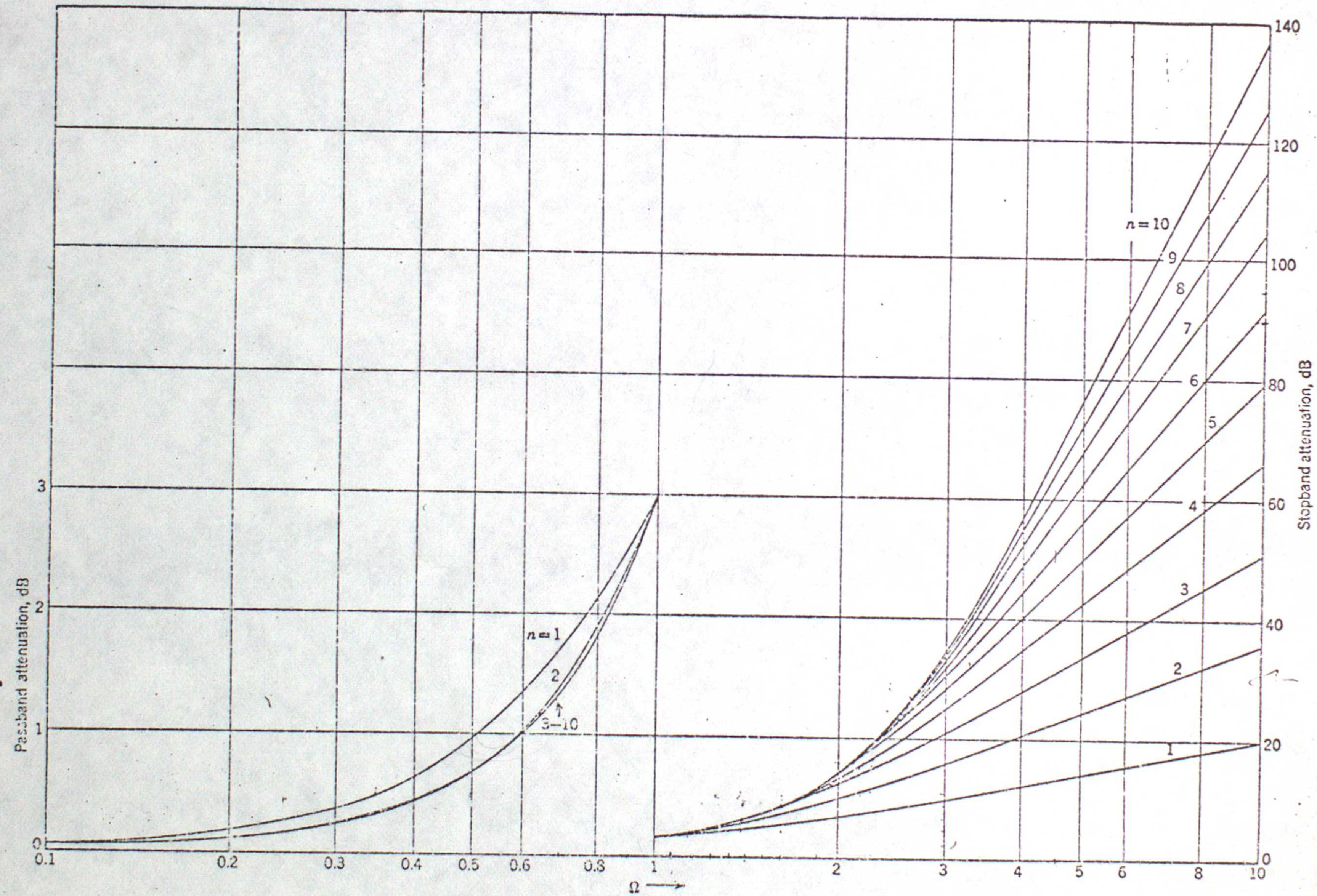


Impulse response for Chebyshev filters with 0.5 dB ripple.



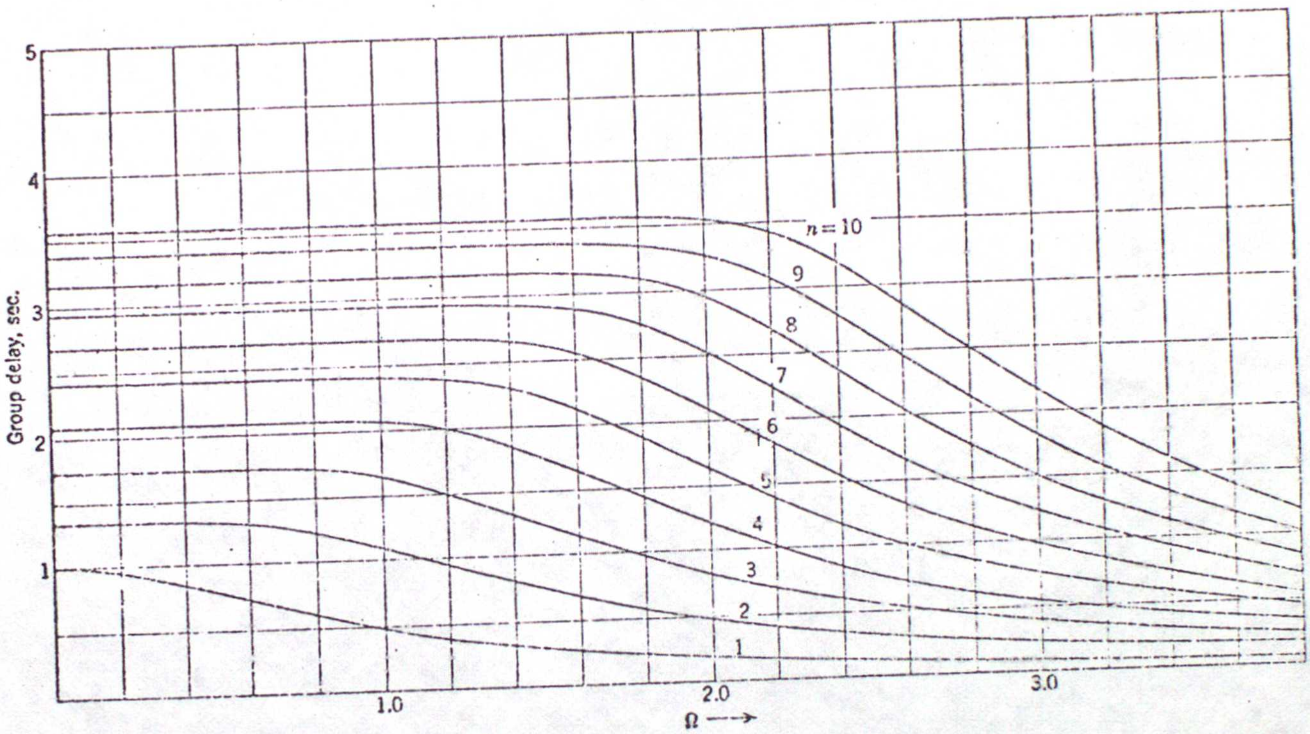
Step response for Chebyshev filters with 0.5 dB ripple.

Fig 7



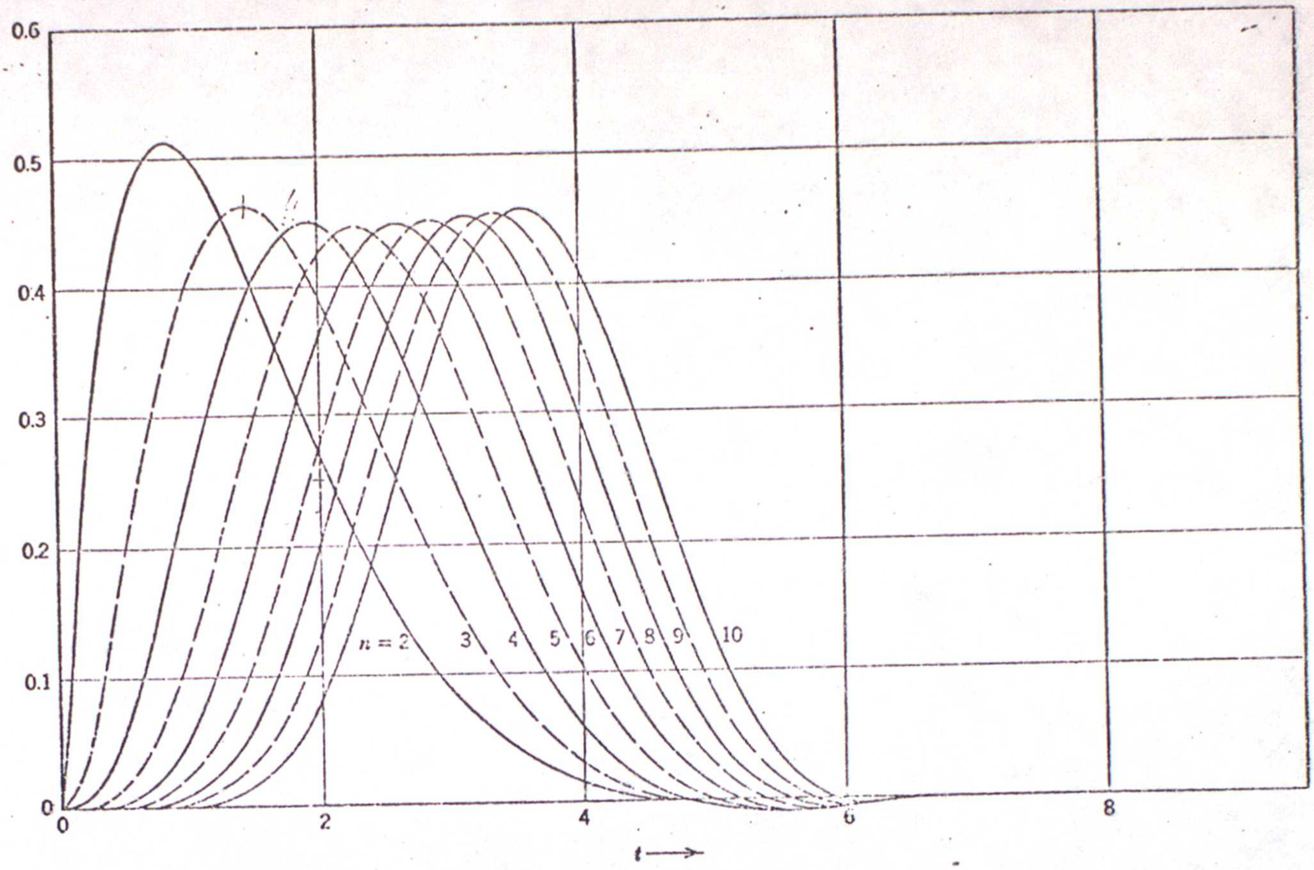
Attenuation characteristics for maximally flat delay (Bessel) filters.

Fig 8

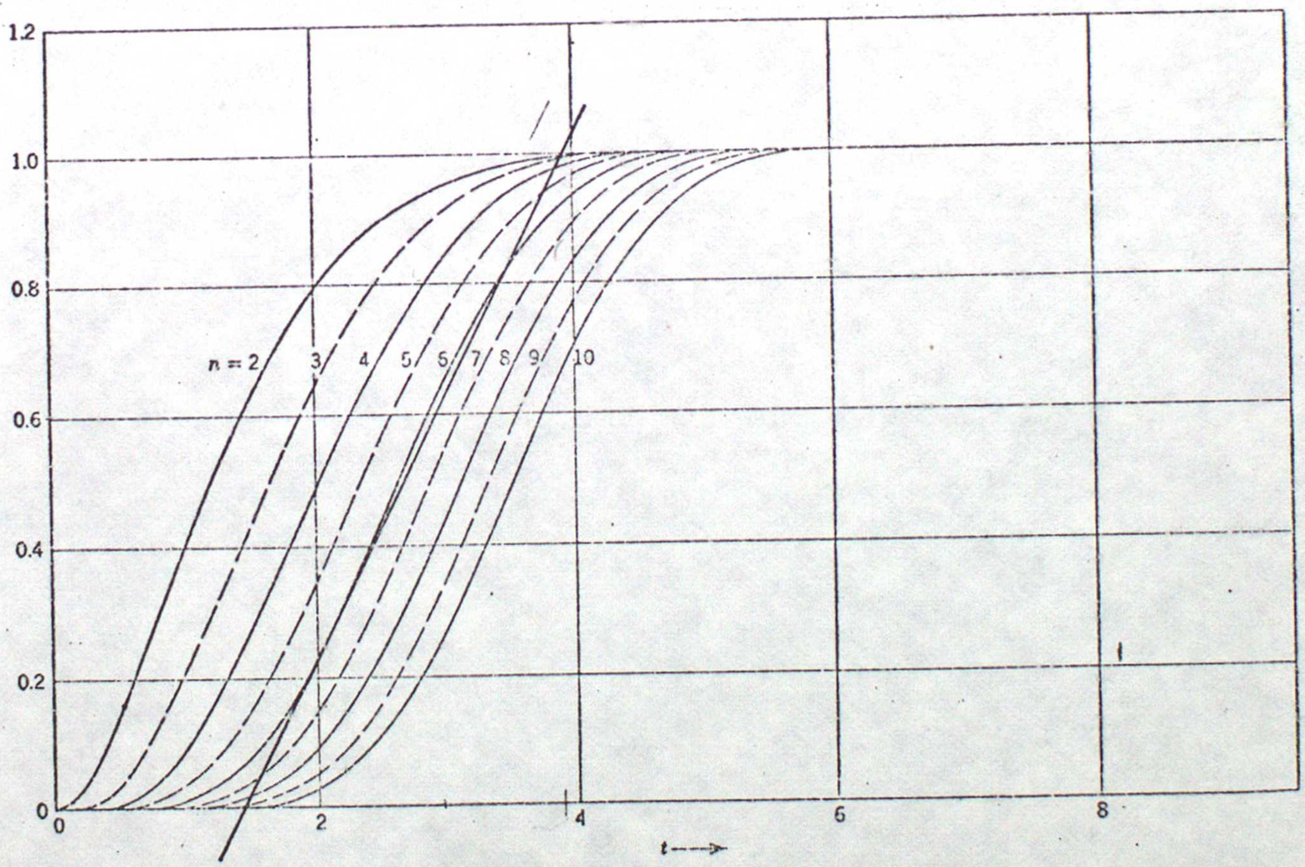


Group delay characteristics for maximally flat delay (Bessel) filters.

Fig 9

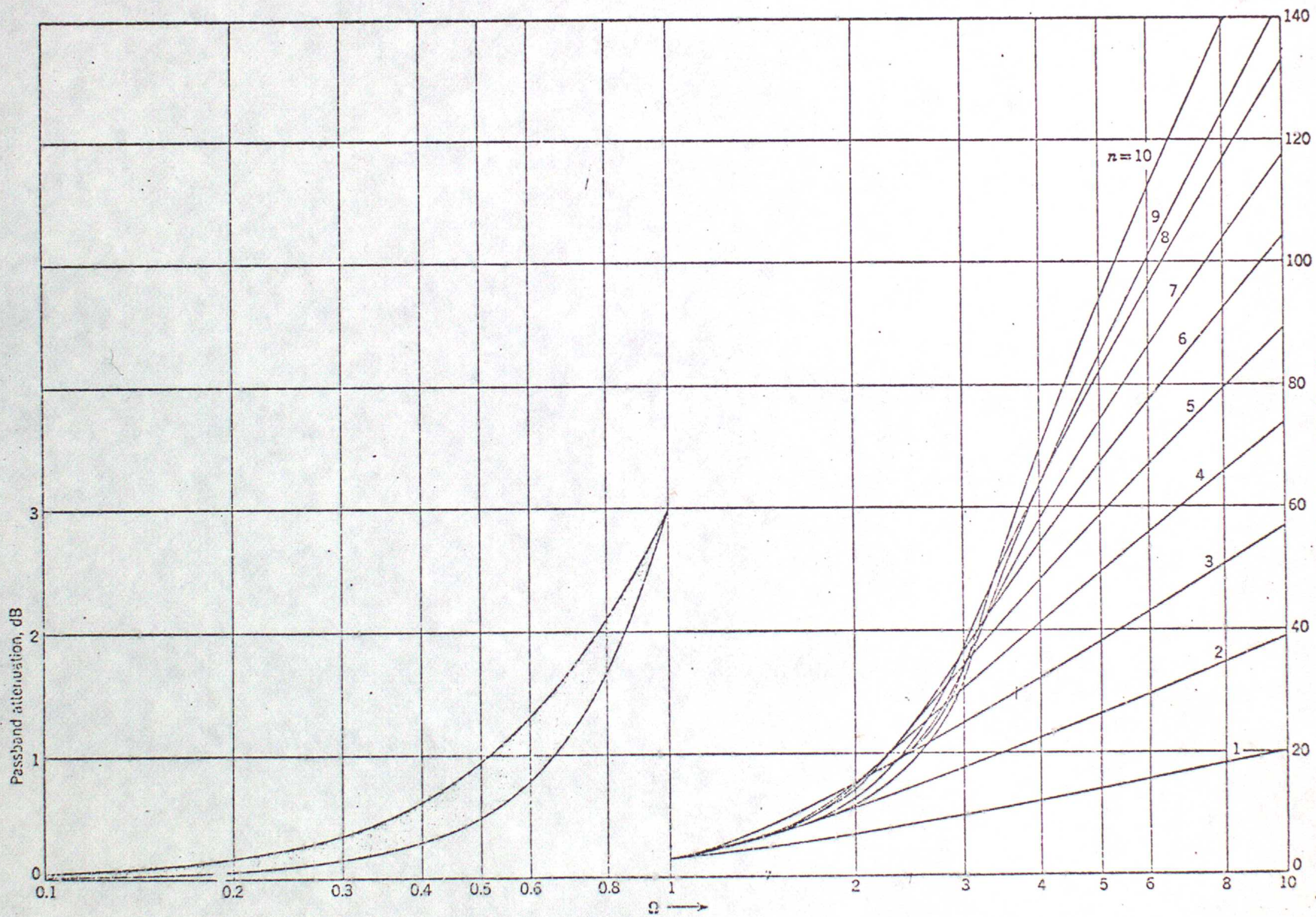


Impulse response for maximally flat delay (Bessel) filters.



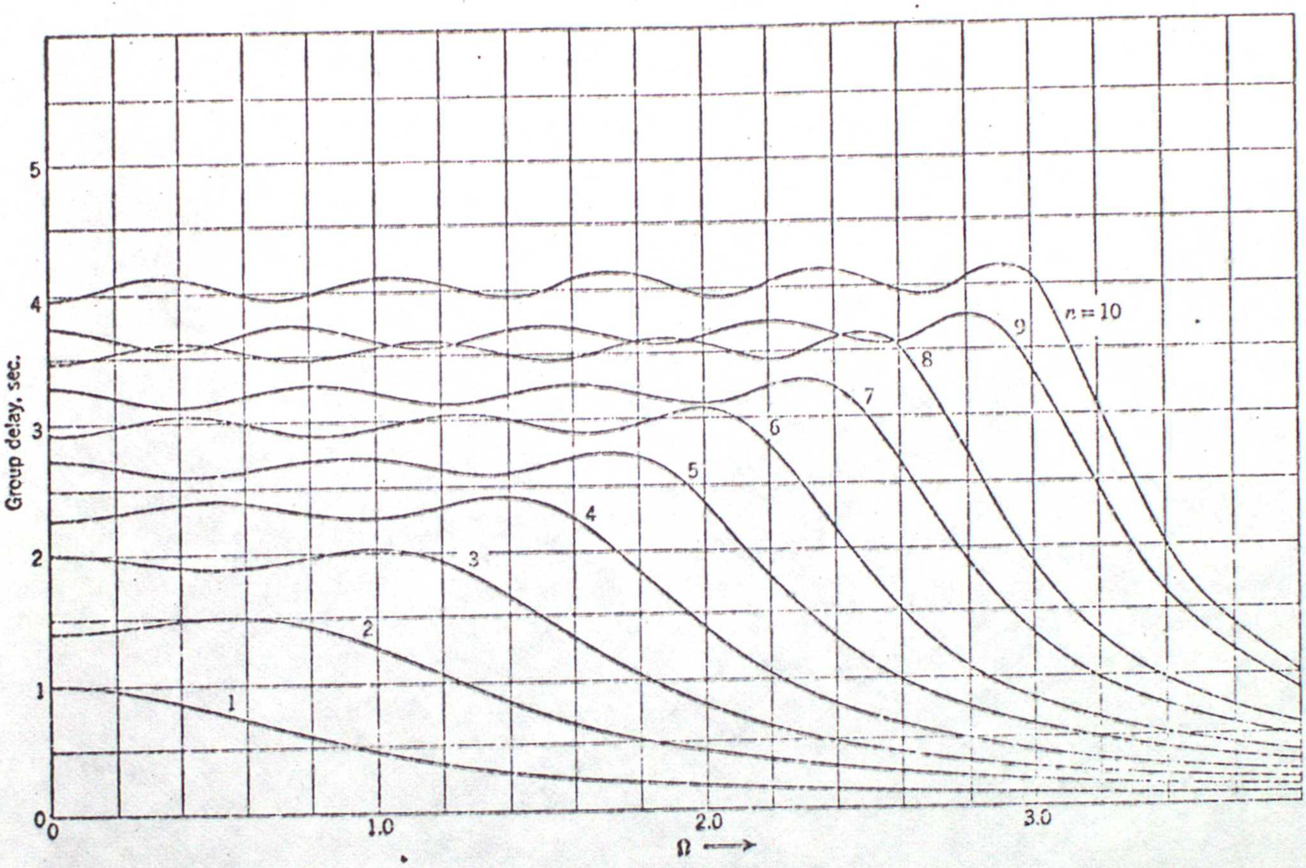
Step response for maximally flat delay (Bessel) filters.

Fig 10



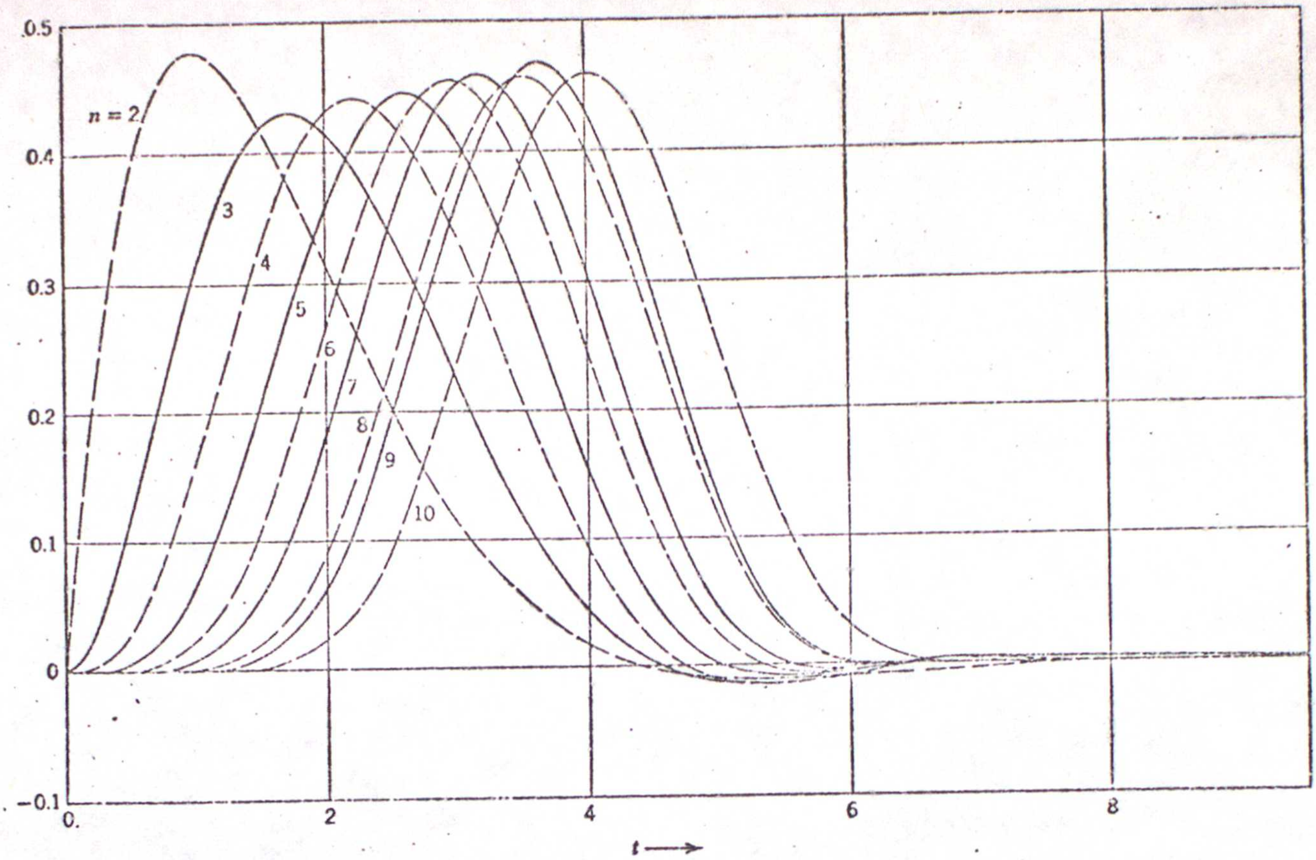
Attenuation characteristics for linear phase with equiripple error filter (phase error = 0.5°).

Fig 11

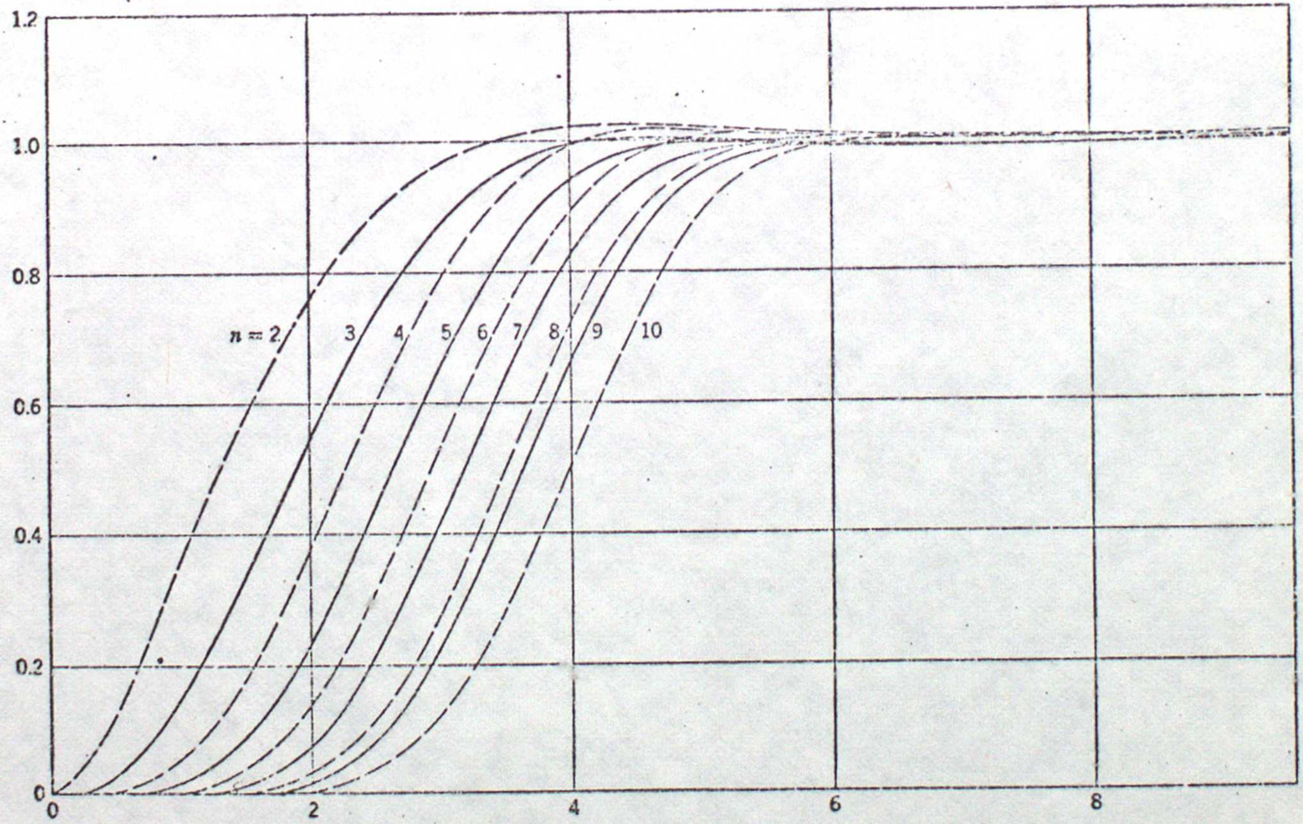


Group-delay characteristics for linear phase with equiripple error filter (phase error = 0.5°).

Fig 12



Impulse response for linear phase with equiripple error filters (phase error = 0.5°).



Step response for linear phase with equiripple error filters (phase error = 0.5°).

Fig 13

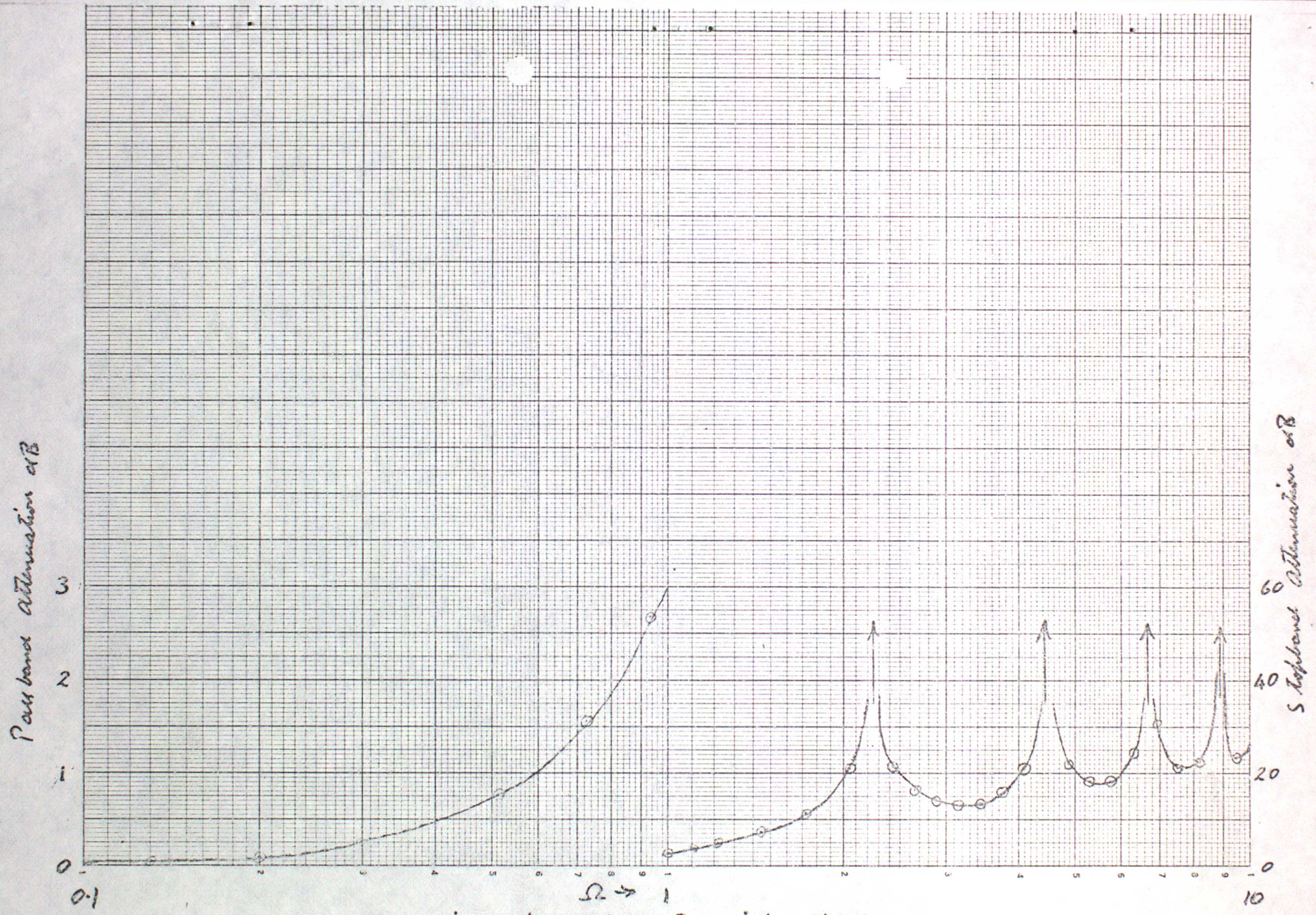


Fig 1k. Attenuation characteristics of Arithmetic Average Filter.

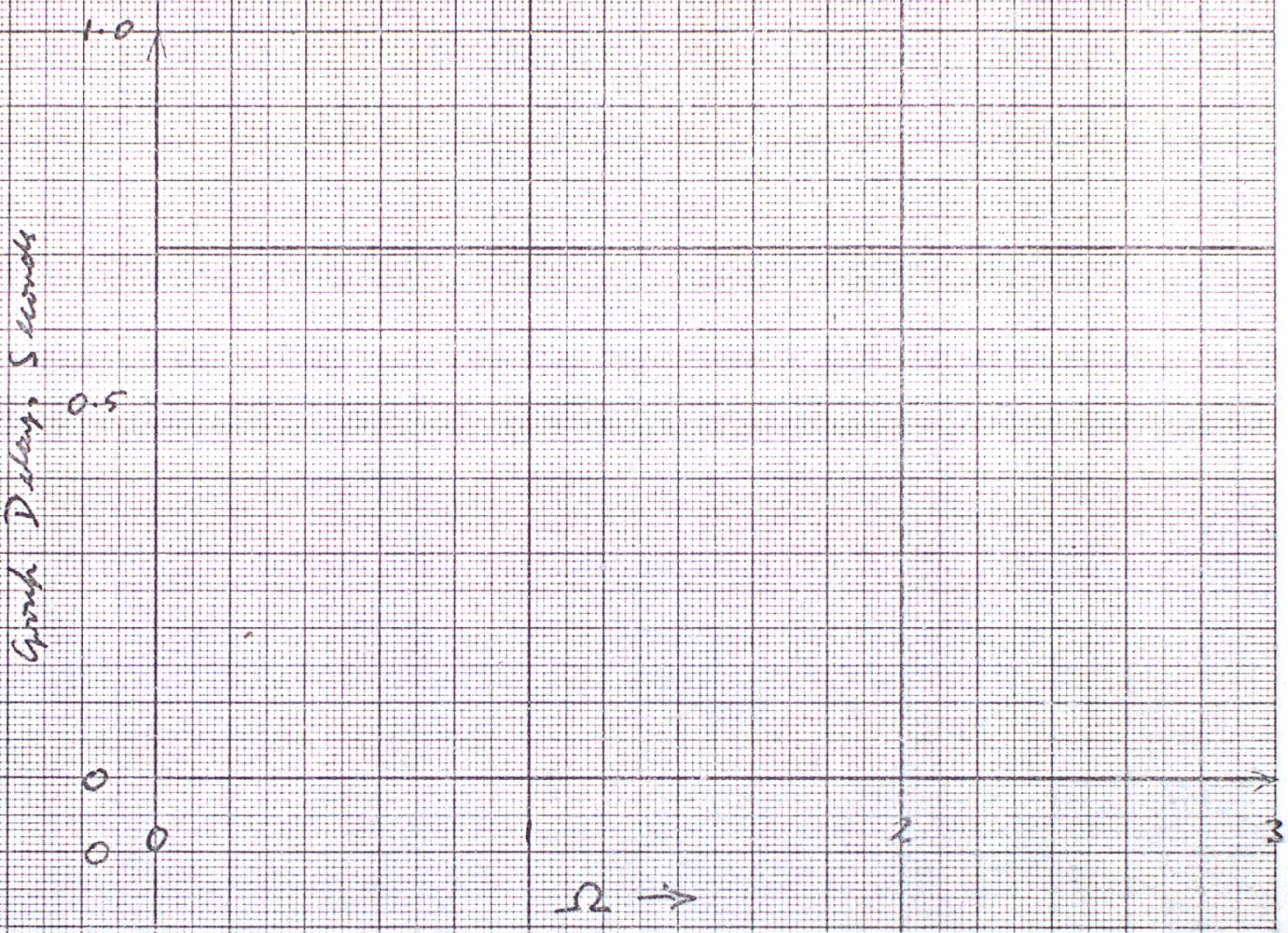
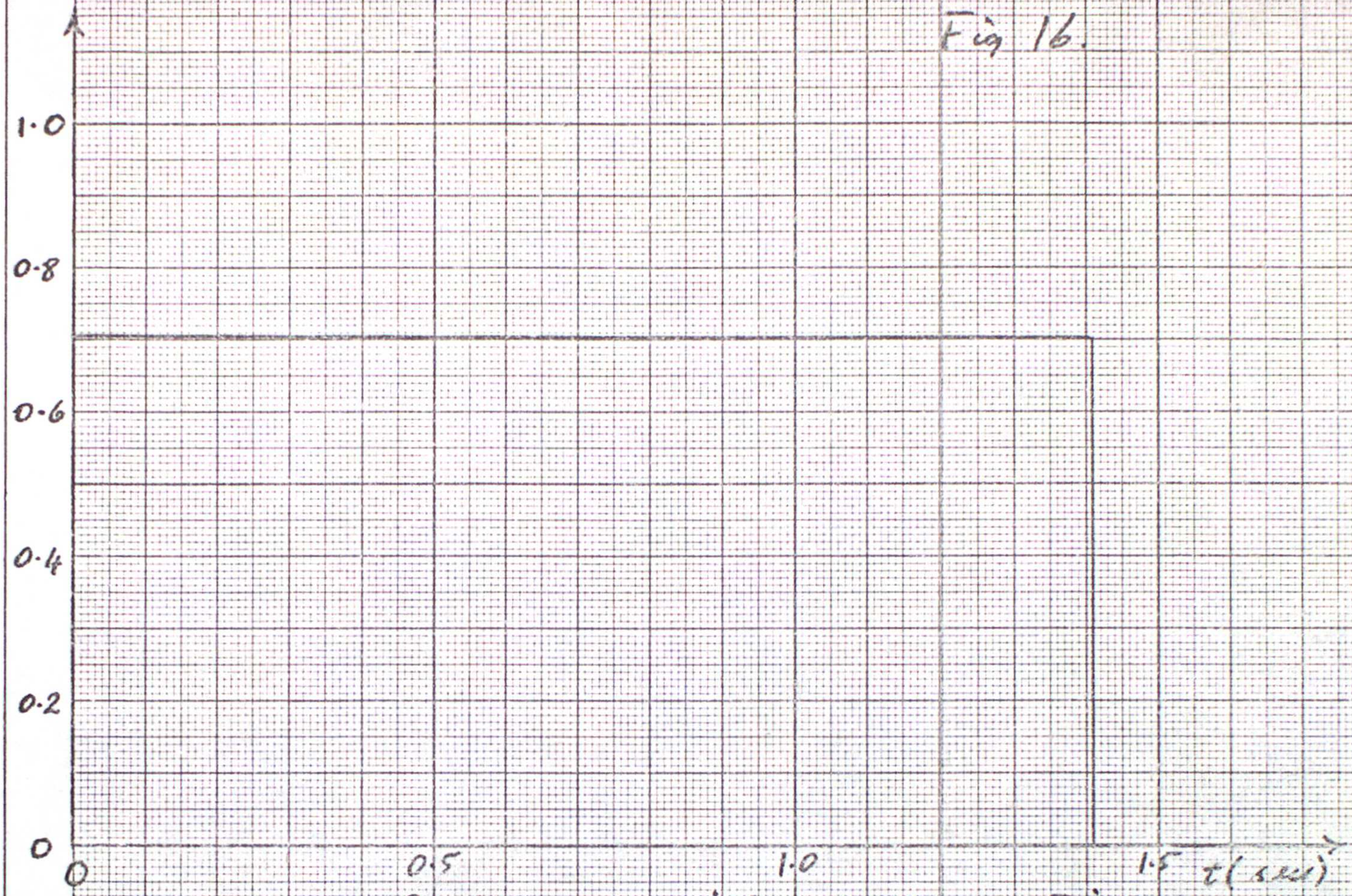
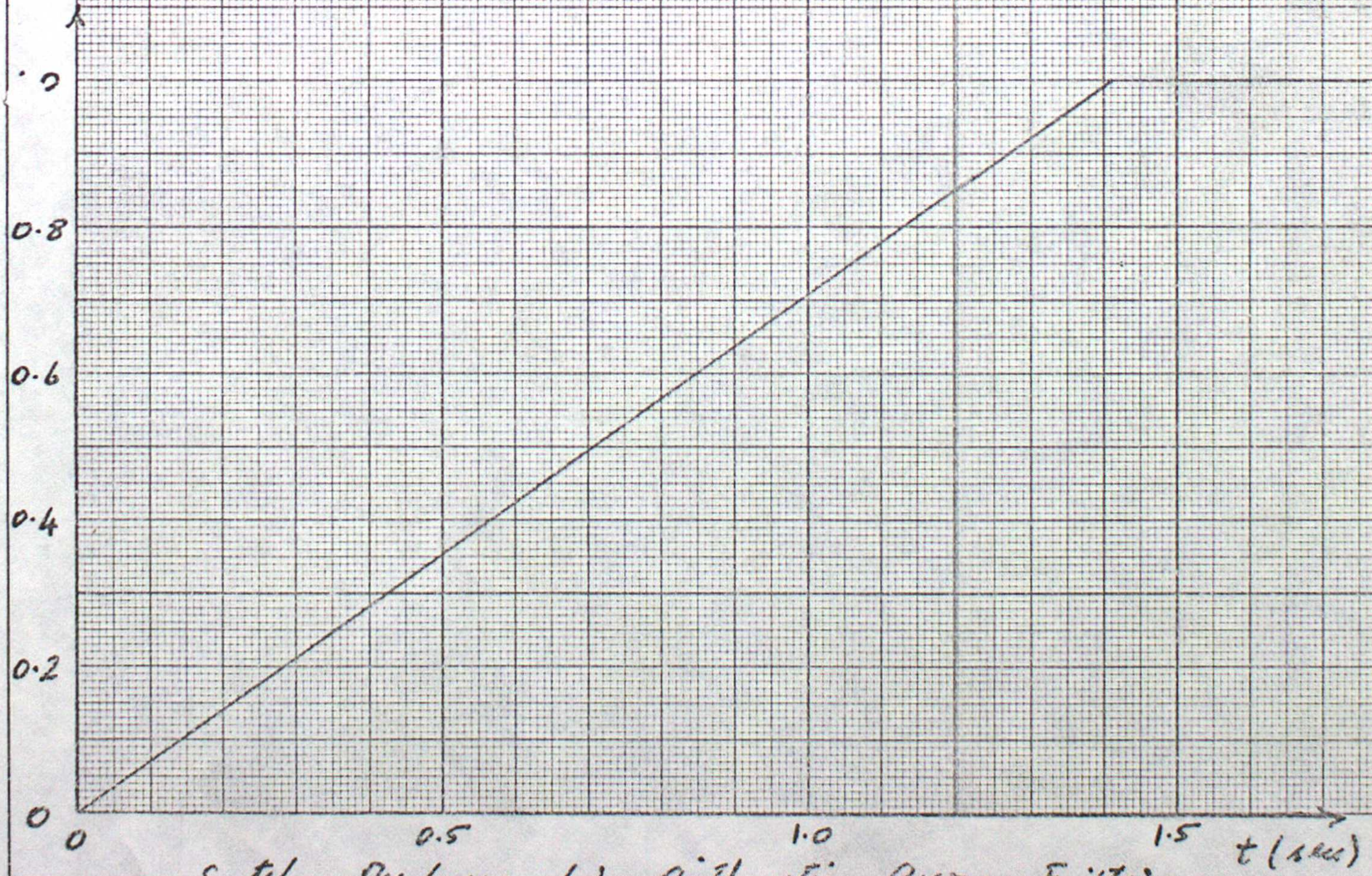


Fig. 15 Group Delay Characteristic for Arithmetic Average Filter.

Fig 16.



Impulse Response for Arithmetic Average Filter



Step Response for Arithmetic Average Filter

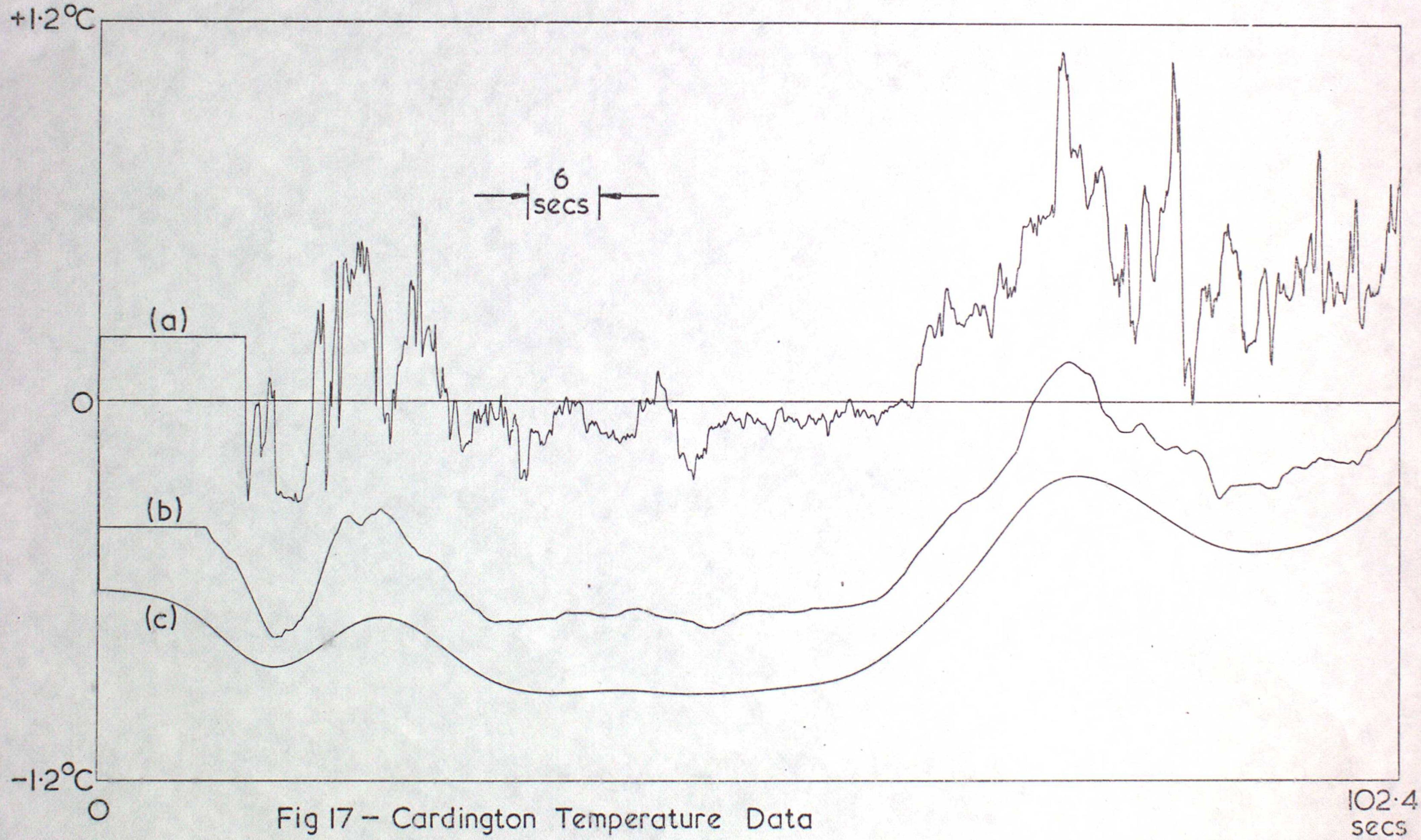


Fig 17 – Cardington Temperature Data

(a) Raw Data (b) 6 - Second Moving Mean (c) Equivalent 6 - Pole Bessel Filter

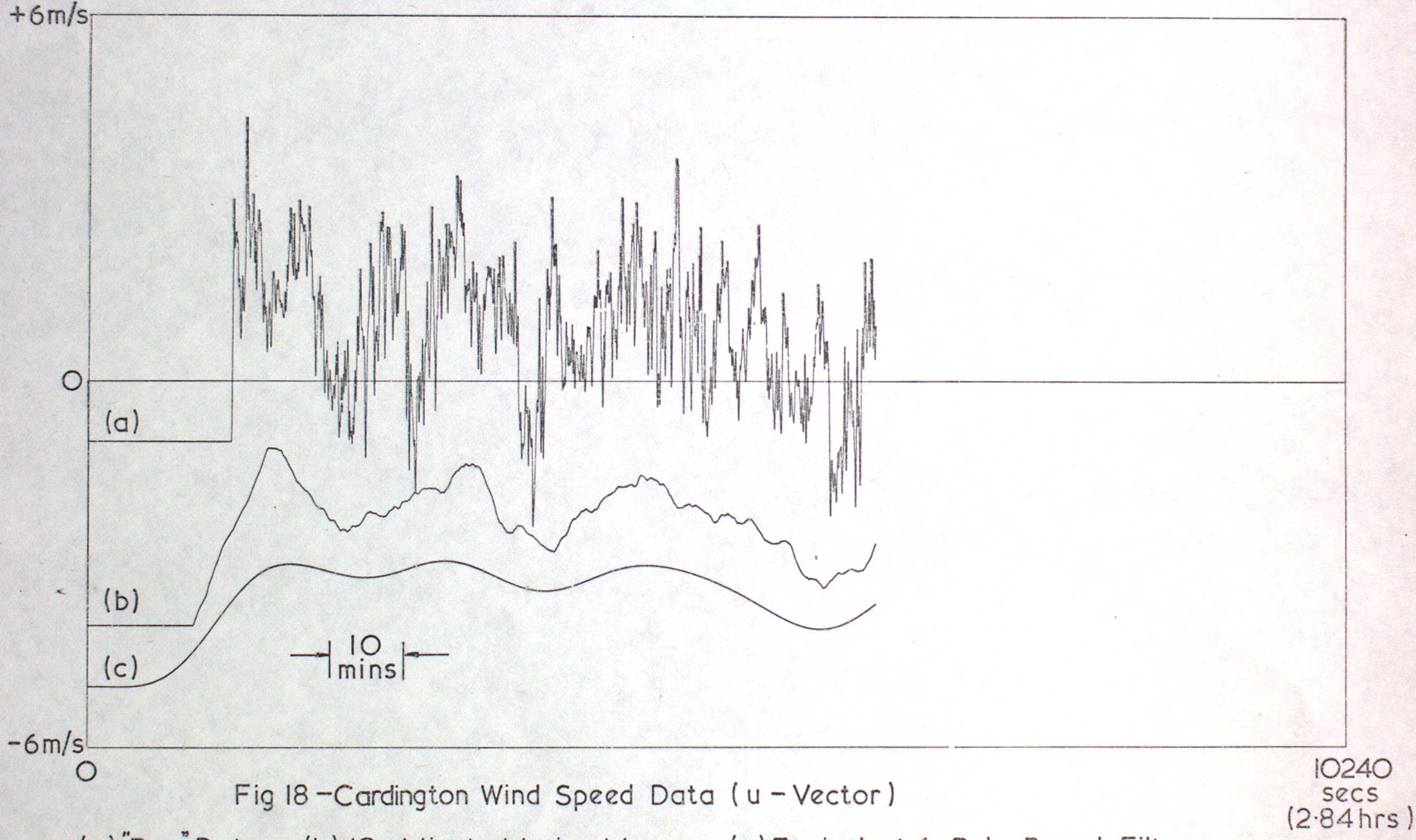


Fig 18 – Cardington Wind Speed Data (u – Vector)

(a) "Raw" Data (b) 10 - Minute Moving Mean (c) Equivalent 6-Pole Bessel Filter

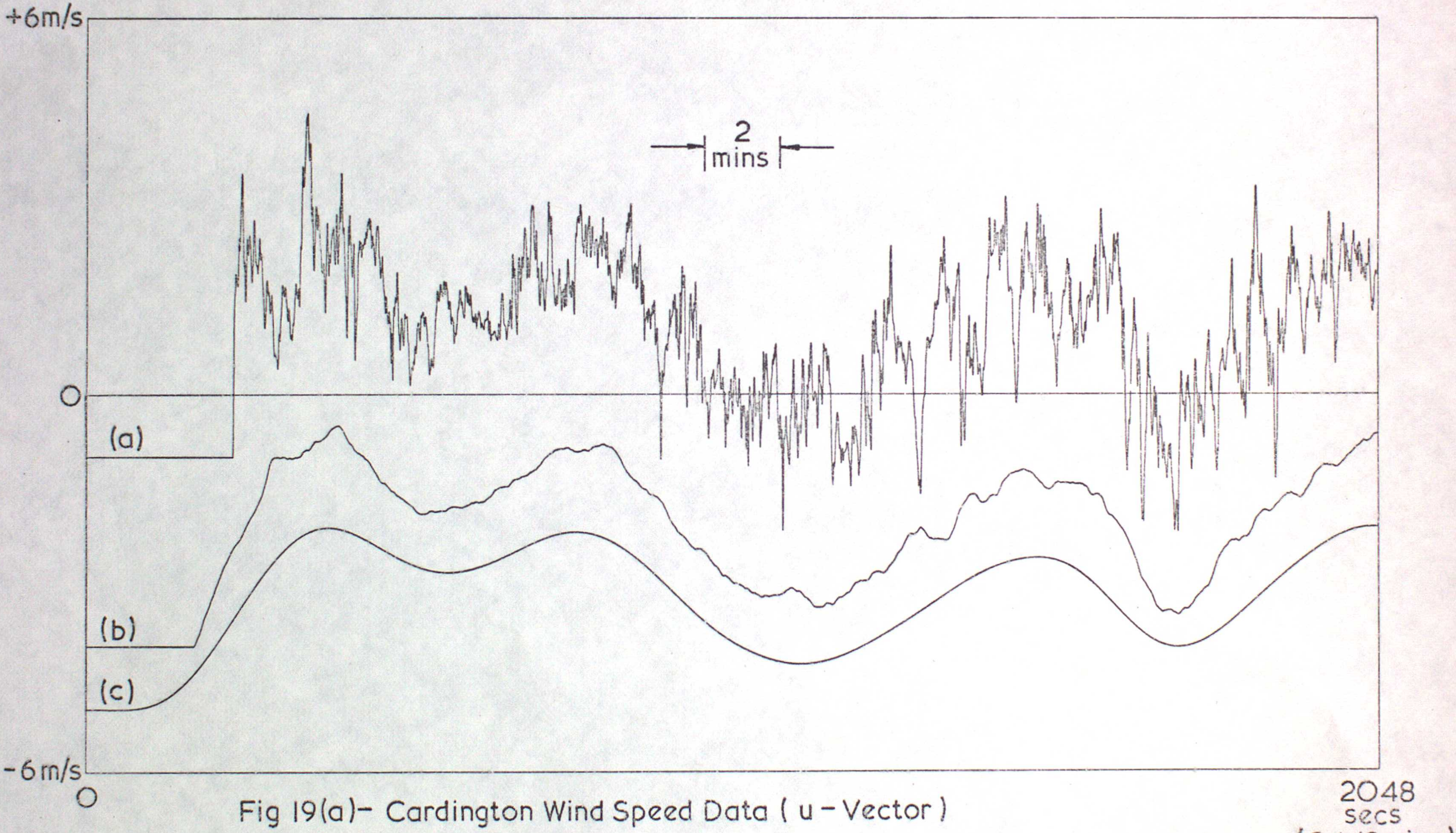


Fig 19(a)– Cardington Wind Speed Data (u – Vector)

(a) "Raw" Data (b) 2-Minute Moving Mean (c) Equivalent 6-Pole Bessel Filter

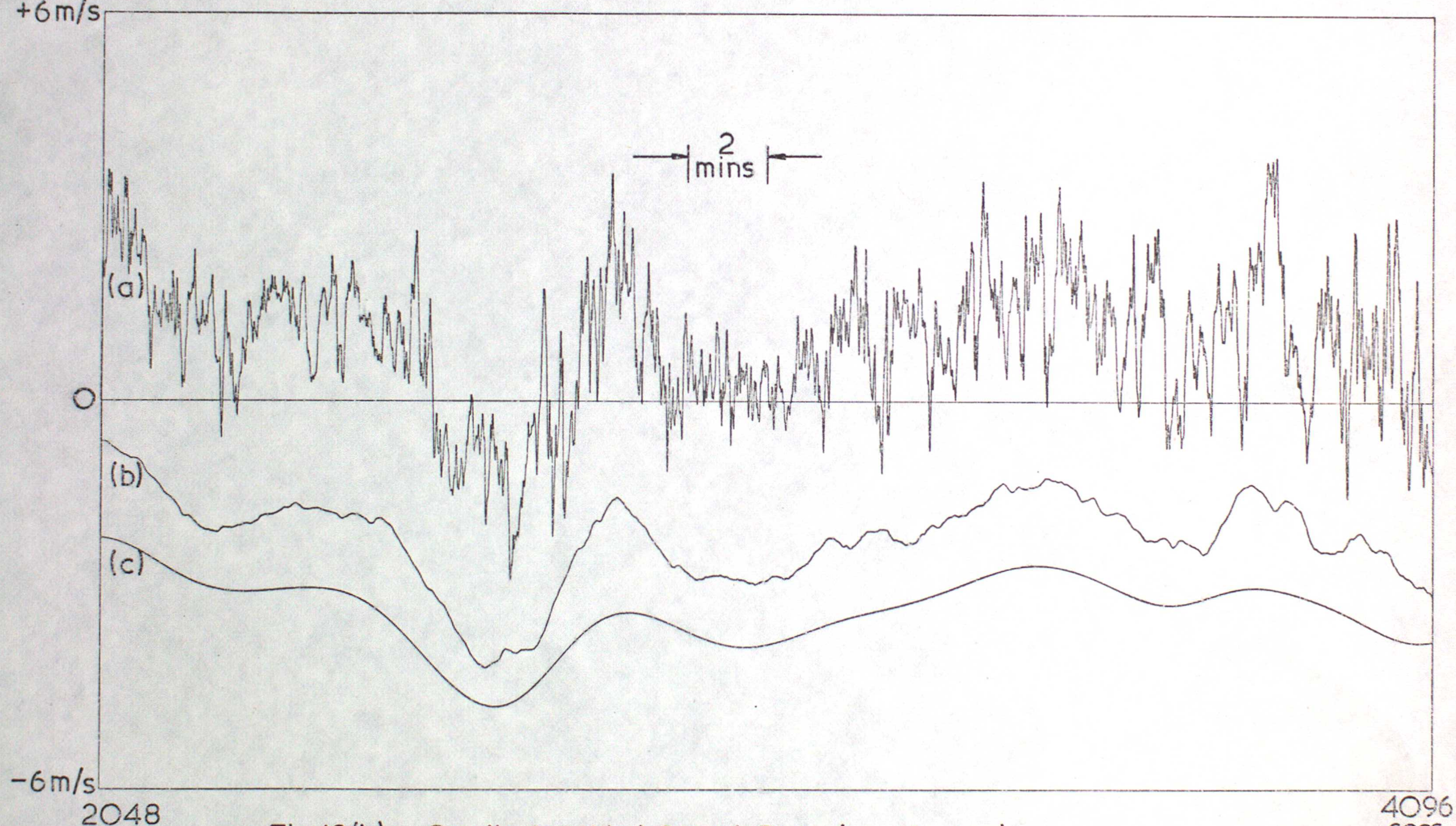


Fig 19(b) – Cardington Wind Speed Data (u – Vector)

(a) "Raw" Data (b) 2-Minute Moving Mean (c) Equivalent 6-Pole Bessel Filter

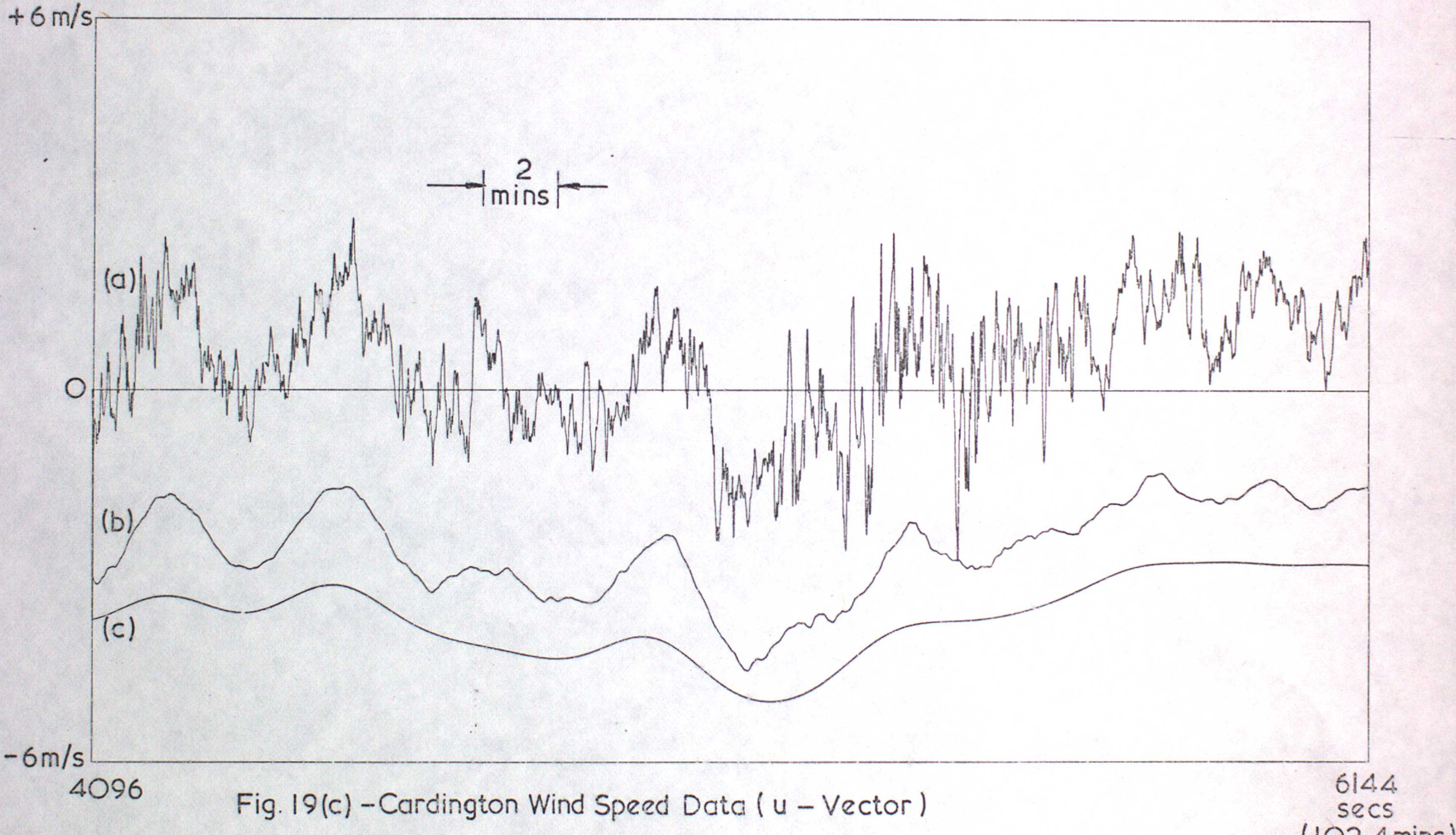


Fig. 19(c) – Cardington Wind Speed Data (u – Vector)

(a) "Raw" Data (b) 2-Minute Moving Mean (c) Equivalent 6-Pole Bessel Filter