

# On the design of the optimum numerical filter with a prefixed response

M. GALLI - P. RANDI (\*)

Ricevuto il 5 Ottobre 1967

**SUMMARY.** — The number of the coefficients of a numerical filter of an equispaced time series, corresponding to a given response is generally unlimited; this number should be limited for obvious computing reasons and not be greater than a certain value so that two data of the filtered series do not become dependent from each other, because of the filter.

In this work it is shown how to calculate a filter with a given number of coefficients whose response is the best approximation to a given response and meanwhile satisfies to certain desired properties.

**RIASSUNTO.** — In generale il numero dei coefficienti di un filtro numerico di una serie temporale a dati equispaziati, corrispondenti ad una risposta prefissata è illimitato; in pratica interessa che detto numero sia limitato per ragione di calcolo, e che non superi un certo valore per la necessità, spesso richiesta, che due dati della serie filtrata non siano diventati tra loro dipendenti a causa della operazione di filtro.

In questo lavoro si mostra come si possa determinare un filtro di cui sia dato il numero dei coefficienti e la cui risposta sia la più approssimata a una risposta prefissata, e che nel contempo soddisfi a certe proprietà desiderate.

## INTRODUCTION.

During the analysis of Cosmic Ray data from this laboratory, it was often found necessary to select or, as some people say, filter from a time series, waves with frequencies falling within a certain limited,

---

(\*) Istituto di Fisica Università di Bologna.

Gruppo Italiano di Fisica Cosmica; Consiglio Nazionale delle Ricerche.

or unlimited band. Similar problems arise also frequently in the analysis of time series of equispaced data of Geophysics.

An operation of numerical filtering on a series of equispaced data  $n_i$  ( $i = 1, 2, \dots$ ) consists of the calculation of the new series:

$$n'_i = \sum_{-r}^{+s} a_j n_{i+j} \quad (i = 1, 2, \dots) \quad [1]$$

where the  $(r+s+1)$  quantities  $a_j$  are called coefficients or weights of the numerical filter.

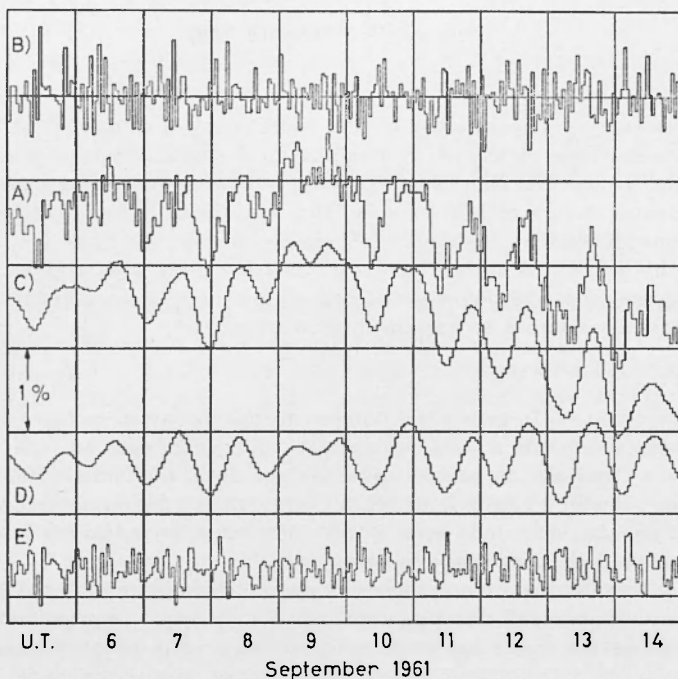


Fig. 1

- A) a sample of hourly barometric corrected data of the cosmic ray total ionizing component, from a scintillation counter working in Bologna;
- B) the prime difference series obtained from the time series represented by A);
- C) the series A) when filtered by the low-pass filter represented by curves C) and E) of Fig. 3;
- D) the series A) when filtered by the band-pass filter represented by curves C) and F) of Fig. 4;
- E) the series A) when filtered by the high-pass filter represented by curves C) and E) of Fig. 5.

As an illustrating example let us consider the series of hourly barometric corrected data of the cosmic ray total ionizing component, from a scintillation counter working in Bologna; represented by the curve (A) of Fig. 1. Now let us observe the position of diurnal oscillation minima around 2400 of September 6<sup>th</sup> and 7<sup>th</sup>, those around noon of September 10<sup>th</sup>, 11<sup>th</sup>, 12<sup>th</sup> and also the minima at earlier times of 13<sup>th</sup>, 14<sup>th</sup>. From the physical point of view, in that data series it is clear the existence of a diurnal wave of variable amplitude and phase, as shown by curve (C) of Fig. 1. It is also evident that if one extracts with conventional methods the diurnal wave with the bandwidth allowed by the length of the data series, one obtains only the mean value and phase over the whole interval of data.

An ideal method should allow the separation of that wave from fluctuations of shorter period and from drifts, that from the physical point of view have a completely different meaning.

A numerical filtering procedure as indicated by [1] that can at least partially satisfy those conditions, will obviously give a much better guarantee of independence of the amplitude and phase of a wave from the following, as the number of filter coefficients is small.

On the other hand one must keep in mind that the smaller the coefficient number, the more scattered is the frequency band filtered. Furthermore the larger the number of coefficients used, the larger is the computation quantity.

So one faces the following problem: given a certain frequency response function (\*), that out of physical reason might be considered the most desirable, and given also the maximum permissible coefficient number, find the coefficient set whose response is the most approximated to the one required.

This is essentially a problem of numerical calculus, which has only little to do with the statistical treatment of filters of which an abundant literature exists (<sup>1</sup>, <sup>2</sup>).

In the past, calculating procedures such as cymanalysis (<sup>3</sup>, <sup>4</sup>), period analysis, smoothing etc. (<sup>5</sup>, <sup>6</sup>), were devised for the separation in a time series, of a wave with a given period or eliminating short period fluctuations often of statistical origin.

---

(\*) By frequency response function we mean the fractional amplitude of a wave of a given frequency that has passed through after the filtering operation.

Here we will briefly revise the relationship between a given function and another one obtained from it by a filtering procedure; then we will solve the specific problem and give some examples of finding a filter for a series of equispaced data satisfying some given conditions.

#### FILTERING OF AN EQUISPACED TIME SERIES.

Given a function  $n(t)$  quadratically integrable defined in the interval  $(-\infty, +\infty)$  and a function  $g(t)$  also quadratically integrable defined in the interval  $(a, b)$ , the equation:

$$n'(t) = \int_a^b g(x) n(t+x) dx, \quad [2]$$

will define a function  $n'(t)$  obtained from  $n(t)$  through a filtering function  $g(t)$ .

If  $f$  is the frequency,  $\mu'(f)$ ,  $\mu(f)$  respectively the Fourier transform of  $n'(t)$  and  $n(t)$  and  $\psi^*(f)$  the conjugate complex of Fourier transform of  $g(t)$ , it will be (see Appendix I):

$$\frac{\mu'(f)}{\mu(f)} = \psi^*(f). \quad [3]$$

One can also show that:

$$\frac{v'(f)}{v(f)} = \frac{\varphi(f)}{2}, \quad [4]$$

$v'(f)$ ,  $v(f)$ ,  $\varphi(f)$  being the frequency spectra of  $n'(t)$ ,  $n(t)$ ,  $g(t)$  according to the definition [21] of Appendix I.

Equation [4] means that after filtering, the amplitude of a wave in the frequency interval  $(f, f + df)$  is reduced to the fraction  $\varphi(f)/2$ . Furthermore filtering, shifts a wave of frequency  $f$ , by the amount:

$$\eta = \text{arctg} \frac{\varphi_2}{\varphi_1}, \quad [5]$$

$\varphi_1$ , and  $\varphi_2$  being the real and the imaginary parts of  $\psi(f)$ .

Let equation [1] be considered. Let  $i$  be the order number of the abscissa interval, and  $n_i$  the value of a possible function  $n(t)$  corresponding to the abscissa:

$$t = i \Delta t \quad (i \text{ integer}); \quad [6]$$

if one defines the step function:

$$n(t) = \begin{cases} (n_{i-1} + n_i)/2 & \text{for } t = (i - 1/2) \\ n_i & \text{for } (i - 1/2) < t < (i + 1/2) \\ (n_i + n_{i+1})/2 & \text{for } t = (i + 1/2), \end{cases} \quad [7]$$

and uses it in place of  $n_i$ , and if similarly one defines  $n'(t)$ ,  $g(t)$  and uses them in place of  $n'_i$ ,  $g_i$ ; if  $\Delta t = 1$ , then equation [1] can be written as follows:

$$n'(t) = \int_a^b g(x) n(t+x) dx. \quad [8]$$

As a consequence, one can easily see that allowing for the difference between a definite integral and its approximate calculation by a summation, [4] and [5] are with a very good approximation the results obtained using [1] even though the number of points describing a period of sinusoidal curve is as low as five or six.

#### DESIGN OF THE OPTIMUM NUMERICAL FILTER CORRESPONDING TO A GIVEN RESPONSE.

From equation [3], [4], and [5], it is evident there is an infinite number of filters having the same response. Therefore for reason of convenience only symmetrical filters will be considered. Now let  $\psi_o(f) = \phi(f)/2$  be the desired response of a symmetrical filter  $g(t)$ . Their coefficients will form the set:

$$g_k (k = -\infty, \dots, N, \dots, 0, \dots, N, \dots, +\infty) \quad (g_k = g_{-k}).$$

Since it is impossible to use a filter with an infinite number of coefficients, and since, as we have seen with the example at the beginning, it is desirable for the coefficients to be reduced to a limited number, it will follow that, to the subset of coefficients:

$$g_k \quad (k = -N, \dots, 0, \dots, N) \quad [9]$$

made of  $N+1$  different coefficients, will correspond a response  $\psi(f)$  which, using [8] will be given by:

$$\psi(f) = \sum_{-N}^N g_k [\sin \omega (k + 1/2) - \sin \omega (k - 1/2)]/\omega. \quad [10]$$

The best set of coefficients will be obtained by solving the following equation of obvious significance:

$$\int_0^{\infty} [\psi_0(f) - \psi(f)]^2 df = \text{minimum}, \quad [11]$$

together with the following optional condition:

$$\sum_{-N}^N g_k = C, \quad [12]$$

which might be useful for example, when we want to eliminate the slow variations, or when we want to keep the average value unchanged (\*). Substituting [10] into the first term of [11] and proceeding according to the method of Lagrange multipliers, we obtain the set:

$$\left\{ \begin{array}{l} g_k - 2(\gamma_{k+1/2} - \gamma_{k-1/2}) + \lambda = 0 \quad (k = -N, \dots, 0, \dots, N) \\ \sum_{-N}^N g_k - C = 0, \end{array} \right. \quad [13]$$

whose solution is:

$$g_k = 2[\gamma_{k+1/2} - \gamma_{k-1/2} - \gamma_{N+1/2}/(N + 1/2)] + C/(2N + 1), \quad [14]$$

where:

$$\gamma_k = \int_0^{\infty} \psi_0(f) [(\sin \omega k)/\omega] df. \quad [15]$$

In Appendix II examples of high-pass, low-pass and band-pass filters are given.

#### ACKNOWLEDGEMENTS.

We wish to express our gratitude to Dr. L. Abbondanza for her valuable help in calculations.

---

(\*) In fact from [10] we obtain:  $\psi(0) = \sum_{-N}^N g_k$ .

a) If  $\psi(0) = \sum_{-N}^N g_k = 1$ , a constant value and the very long waves are not affected by the filter, the same holds for the mean values over long enough intervals.

b) If  $\psi(0) = \sum_{-N}^N g_k = 0$ , a constant value of long enough waves are eliminated. Obviously one may impose on equation [11] not only one condition but as many as  $N$  independent conditions.

## APPENDIX I

If  $n(t)$  is any function quadratically integrable, when defining its Fourier transform as:

$$\mu(f) = \int_{-\infty}^{+\infty} n(t) e^{-i\omega t} dt, \quad (\omega = 2\pi f) \quad [16]$$

let:

$$\mu(f) = \nu(f) e^{i\alpha(f)}, \quad [17]$$

since the real part of  $\mu(f)$  is an even function, and its imaginary part is an odd function, one obtains:

$$n(t) = \int_0^{\infty} \nu(f) \cos [2\pi f t + \alpha(f)] df, \quad [18]$$

which shows that  $\mu(f)$  as defined by [17] represents the frequency spectrum of  $n(t)$ . Now, given a finite valued function  $g(t)$ , zero everywhere except in the interval  $(a, b)$ , the new function:

$$n'(t) = \int_a^b g(x) n(t+x) dx, \quad [19]$$

will be called the filtered function obtained from  $n(t)$ .

From a preceding definition one obtains also:

$$n'(t) = \int_{-\infty}^{+\infty} \mu(f) e^{i\omega t} \left[ \int_a^b g(x) e^{i\omega x} dx \right] df$$

and:

$$n'(t) = \int_{-\infty}^{+\infty} \mu'(f) e^{i\omega t} df = \int_{-\infty}^{+\infty} \mu(f) \psi^*(f) e^{i\omega t} df, \quad [20]$$

$\psi^*(f)$  being the conjugate complex of the transform  $\psi(f)$  of  $g(t)$  and  $\mu'(f)$  the transform of  $n'(t)$ .

From equation [20] it can be seen that:

$$\mu'(f) = \mu(f) \psi^*(f), \quad [21]$$

introducing the frequency spectrum  $\varphi(f)$  of  $g(t)$  and the frequency spectrum  $\nu'(f)$  of  $n'(t)$  one obtains:

$$\nu'(f)/\nu(f) = \varphi(f)/2. \quad [22]$$

Since [21] may also be written in the form:

$$[\nu'(f)/2] e^{i\alpha'(f)} = [\nu(f)/2] e^{i\alpha(f)} [\varphi(f)/2 e^{-i\eta(f)}],$$

where  $\alpha'(f)$ ,  $\eta(f)$  are the arguments of  $\mu'(f)$ ,  $\psi(\bar{f})$ , one understands that:

$$\eta = \text{arctg } \varphi_2/\varphi_1,$$

is the phase shift introduced by the filter in the sinusoidal components of  $n(t)$ . The function

$$\varphi(f)/2 = \psi(f),$$

will be called the frequency response of the filter  $g(t)$ .

If  $g(t)$  is an even function, its response will be:

$$\psi(f) = \varphi(f)/2 = \int_{-\infty}^{+\infty} g(t) \cos \omega t \, dt.$$

#### EXAMPLE.

As an example of calculation of a filter response, one considers the filtered effect introduced by the prime differences of a series, which is a type of rather frequent calculation.

The series  $n'_i$  of the first differences obtained from a series  $n_i$ :

$$n'_i = n_{i+1} - n_i,$$

may be considered as obtained from  $n_i$  through a filter whose coefficients are:

$$g_0 = -1 \quad g_1 = 1.$$

Function  $g(t)$  as defined by [7] will be [see A] in Fig. 2]:

$$g(t) = \left. \begin{array}{lll} 0 & \text{for} & t < -1/2 \\ -1/2 & \text{''} & t = -1/2 \\ -1 & \text{''} & -1/2 < t < 1/2 \\ 0 & \text{''} & t = 1/2 \\ 1 & \text{''} & 1/2 < t < 3/2 \\ 1/2 & \text{''} & t = 3/2 \\ 0 & \text{''} & t > 3/2. \end{array} \right\}$$



Its Fourier transform will be:

$$\psi^*(f) = 2 [(\sin^2 \omega/2)/(\omega/2)]e^{i(\omega-3\pi)/2} .$$

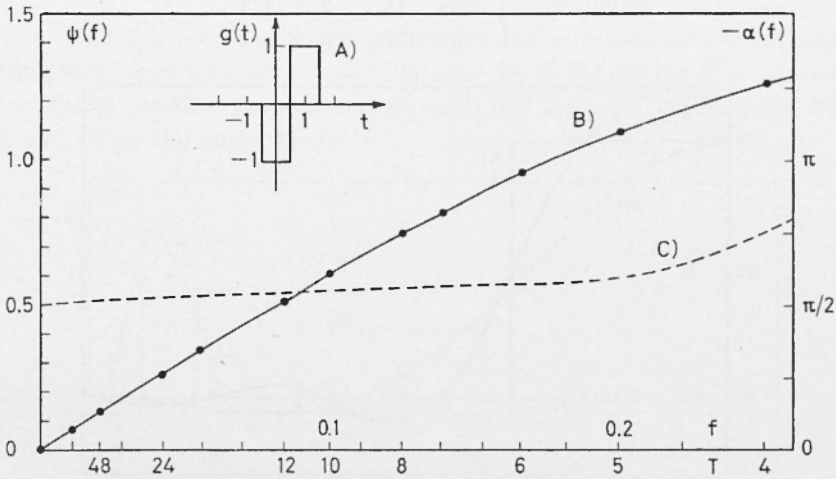


Fig. 2

- A) The filtering function corresponding to a prime difference filter i.e. the operation performed making differences between a term and its preceding;
- B) the frequency response of prime difference filter;
- C) the phase shift introduced by the prime difference filter.

Consequently the frequency response will be [see B) in Fig. 2]:

$$\psi^*(f) = \varphi(f)/2 = (2 T/\pi) \sin^2 \pi/T ,$$

and the phase shifts [see C) in Fig. 2]:

$$\alpha = -\pi/2 - \pi/T .$$

### APPENDIX II

Here below are shown some examples of calculation of filter coefficients of rather frequent use.

#### I. - DESIGN OF A LOW-PASS FILTER.

Let the following conditions be given:

- a) the maximum number of coefficients be  $2N + 1$  ;

b) the ideal frequency response be [see curve A) of Fig. 3]:

$$\psi_0(f) = \begin{cases} 1 & \text{for } 0 < f < 1/T \\ 1/2 & \text{» } f = 1/T \\ 0 & \text{» } f > 1/T \end{cases}$$

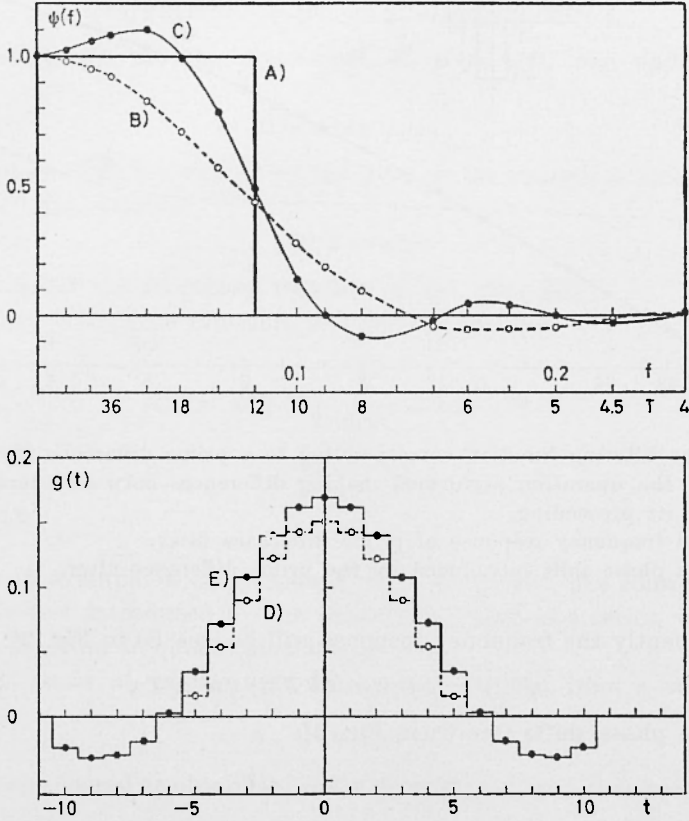


Fig. 3

- A) the ideal response of a low-pass filter;
- B) the frequency response of the filter coefficients represented by D) with  $T=12$  and  $N=5$ ;
- C) the frequency response of the filter coefficients represented by E) with  $T=12$  and  $N=10$ .

c) coefficients fulfil the equation:

$$\sum_{-N}^N g_k = 1 .$$

When doing the calculations one obtains:

$$g_k = \{ Si \pi (2k + 1)/T - Si \pi (2k - 1)/T - [2/(2N - 1)] Si \pi (2N + 1)/T \} / \pi + 1/(2N + 1) .$$

For  $t = 12$  and for  $N = 5$  the coefficients and correspondent frequency response are represented by curve D) and B) of Fig. 3; for  $T = 12$  and  $N=10$  one obtains the coefficients and the response represented by E) and C) of the same figure.

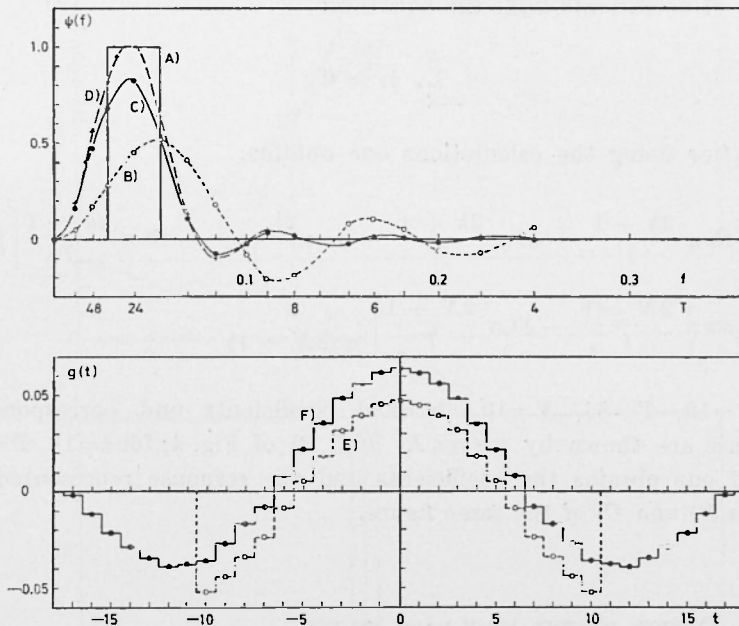


Fig. 4

- A) the ideal response of a band-pass filter;
- B) the frequency response of the filter coefficients represented by E) when  $t=18, T=36, N=10$ ;
- C) the frequency response of the filter coefficients represented by F) when  $t=18, T=36, N=17$ ;
- D) the frequency response of the filter coefficients represented by F) when multiplied by a factor for obtaining a 100% response for a period of 24 data intervals.

## II. - DESIGN OF A BAND-PASS FILTER.

Let the following conditions be given:

- a) the maximum number of coefficients be  $2N + 1$  ;

b) the ideal frequency response [see curve A) of Fig. 4] be:

$$\psi_0(f) = \begin{cases} 0 & \text{for } 0 < f < 1/T \\ 1/2 & \text{» } f = 1/T \\ 1 & \text{» } 1/T < f < 1/t \\ 1/2 & \text{» } f = 1/t \\ 0 & \text{» } f > 1/t \end{cases}$$

c) coefficients fulfil the equation:

$$\sum_{-N}^N g_k = 0 .$$

After doing the calculations one obtains:

$$g_k = \left[ \sin \pi \frac{2k+1}{t} - \sin \pi \frac{2k+1}{T} - \sin \pi \frac{2k-1}{t} + \sin \pi \frac{2k-1}{T} \right] / \pi - \\ - \left[ \sin \pi \frac{2N+1}{t} - \sin \pi \frac{2N+1}{T} \right] \frac{2}{\pi (2N+1)} .$$

For  $t=12$ ,  $T=36$ ,  $N=10$  obtained coefficients and corresponding response are shown by curves E) and B) of Fig. 4; for  $t=18$ ,  $T=36$ ,  $N=17$  one obtains the coefficients and the response represented by curves F) and C) of the same figure.

### III. - DESIGN OF THE HIGH-PASS FILTER.

Let the following conditions be given:

- a) the maximum number of coefficients be  $2N + 1$  ;  
 b) the ideal frequency response be [see curve A) of Fig. 5]:

$$\psi_0(f) = \begin{cases} 0 & \text{for } 0 < f < 1/T \\ 1/2 & \text{» } f = 1/T \\ 1 & \text{» } f > 1/T \end{cases}$$

c) coefficients fulfil the equation:

$$\sum_{-N}^N g_k = 0 .$$

After doing the calculations one obtains:

$$g_0 = 1 - (2/\pi) Si \pi/T + Si \pi (2N + 1)/T - 2/[(2N + 1) \pi] - 1/(2N + 1)$$

$$g_k = [Si \pi (2k - 1)/T - Si \pi (2k + 1)/T]/\pi + [Si \pi (2k + 1)/T] [2/(2N + 1) \pi] - 1/(2N + 1) .$$

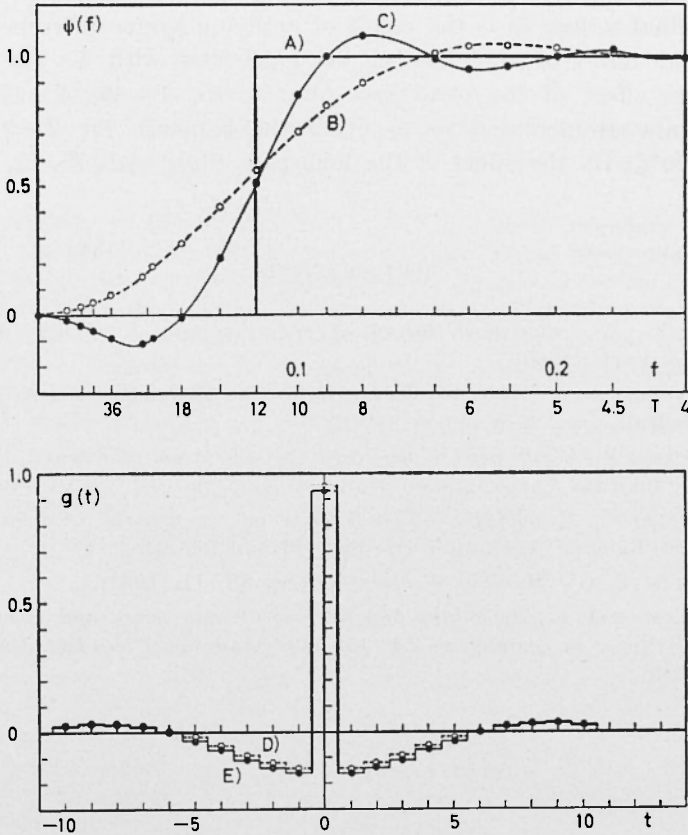


Fig. 5

- A) the ideal response of a high-pass filter;
- B) the frequency response of the filter coefficients represented by D) when  $T=12$ ,  $N=5$ ;
- C) the frequency response of the filter coefficients represented by E) when  $T=12$ ,  $N=10$ .

For  $T=12$  and  $N=5$  obtained coefficients and corresponding frequency response are represented by curves D) and B) in Fig. 5; for  $T=12$

and  $N=10$  one obtains coefficients and frequency response represented by curves E) and C) of the same figure.

#### IV. - AN EXAMPLE OF FILTERING OPERATION.

Figure 1 shows the effect of various filters on a time series: A) is the original series; B) is the result of applying prime difference filter; C) is the result of applying the low-pass filter with  $T=12$ ,  $N=10$ ; D) is the effect of the band-pass filter  $t=18$ ,  $T=36$ ,  $N=17$ , when coefficients are increased so to bring the response for  $T=24$  wave to be 100%; E) the effect of the high-pass filter with  $T=12$ ,  $N=10$ .

#### REFERENCES

- (1) LEE Y. W., *Statistical theory of communication*. J. Wiley & Sons, New York (1960).
- (2) WAINSTEIN L. A., ZUBAKOV V. D., *Extraction of signals from noise*. Prentice-Hall, Inc., New Jersey (1962).
- (3) VERCELLI F., *Guida per l'analisi delle periodicità nei diagrammi oscillanti*. R. Comitato Talassografico Italiano. Memoria CCLXXXV (1940).
- (4) VIANELLI S., *Metodologia statistica delle scienze agrarie*, «Edizioni Agricole» Bologna. 1, chapter III, pp. 737 and following (1954).
- (5) SHUMAN F. G., *Monthly Weather Review*, 85, 11 (1957).
- (6) HOLLOWAY J. L., *Smoothing and filtering of time series and space fields*. «Advances in Geophysics» 4, 351-389. Academic Press Inc. New York (1958).