ICOT/IR/10

icot

A NOTE ON THE UNDERSTANDING AND DESIGNING OF NUMERICAL FILTERS

BY

G. S. QURAISHEE

1967

NATURAL ENVIRONMENT RESEARCH COUNCIL

institute of coastal oceanography and tides

A NOTE ON THE UNDERSTANDING AND DESIGNING OF NUMERICAL FILTERS

BY

G. S. QURAISHEE

1967

This Report was prepared before the Tidal Institute became the Institute of Coastal Oceanography and Tides.

Introduction

In tidal work numerical filters are used either to remove unwanted frequencies from the observations or to isolate a particular band of frequencies so that the filtered data can be analysed smoothly.

Principle of a Filter

If a certain sequence of operation, like summation or difference of observations n units apart or average of n successive observations, etc. is performed on a time series then a spectrum could be assigned to each such operation. Such operations will modify the spectrum of the data. A function which defines such a sequence of operations is known as a filter or smoothing operator. Thus filtering involves the carrying out of certain operations on the observations to create a new sequence of numbers so that the spectrum of the new sequence satisfies our requirement. Different filters suggest different operations on the data.

Type of filters

Basically there are two types of filters 'A_n' and 'S_n' type. A_n filters involve the addition of two observations $n\Delta t$ time units apart. S_n filters involve the subtraction of two observations $n\Delta t$ time units apart. α_n filters are extended form of A_n filters and involve the summation of n consecutive observations of a time series. Note that $\alpha_2 = A_1$. The formalism of response functions of these filters are :

$$A_n = 2 \cos \pi n f \Delta t$$

$$S_n = -2i \sin \pi n f \Delta t$$

$$\alpha_n = \frac{\sin \pi n f \Delta t}{\sin \pi f \Delta t}$$

Δt is the time interval of observation, f is the frequency in cycles per unit time, n as described above. Response functions for the above filters are shown in Fig. 1 to Fig. 3.

Fig. 1A and Fig. 2A show response functions of 'A' and 'S' filters respectively. The abscissa in these figures is in Δt units which can be converted to appropriate frequencies according to the time interval t. 'A' filters are low pass filters and 'S' filters are high pass. The number of frequencies filtered varies with the value of n. As the value of n increases the response functions start to oscillate, thus allowing to pass a portion of higher frequencies in the case of 'A' filters and lower frequencies in the case of 'S' filters. Application of a filter to filtered data is the equivalent of applying the product of two filters. Figs. 2A and 2B show the response functions of such multiple filters. A multiple filter of type $A_1\ A_1$ i.e. A_1^2 is a more efficient low pass filter than A_1. Similarly A_1^3 is better than A_1^2. The same effect is evident in S_1^2 and S_1^3 in suppressing the lower frequencies. Products of filters like $A_1\ A_2\ A_3$; $S_1\ S_2\ S_3$; $A_1\ A_1\ S_1$; $A_1\ S_2\ S_2$, etc. act like band pass filters (see Fig. 2C and Fig. 3A). Responses of α filters and their products are shown in Fig. 3B. α filters, particularly their products are very useful to the tidal and current analysis to cut off high frequencies. These filters are easy to apply to a time series as the sequence of operation involves only the summation of n consecutive numbers.

By understanding the responses of these basic filters several varieties of filters can be devised according to the nature of problems involved.

In tidal work mostly low pass filters are used. Most common low pass filters are of the following type:

$$\frac{\alpha_n \alpha_{n+1}^2}{n(n+1)^2} \quad \text{or} \quad \frac{\alpha_n^2 \alpha_{n+1}}{n^2 (n+1)}$$

let the observations be denoted by $\zeta_1, \zeta_2, \zeta_3, \ldots$ the application of above filters will require either 3n or 3n-1 terms respectively as evident from the following:

Application of α_{n+1} to ζ series means generating a new series, say x, where

$$x_k = \sum_{j=0}^{n} \zeta_{j+k} \text{ where } k = 1, 2, 3, \ldots 2n$$

Thus,

$$x_1 = \zeta_1 + \zeta_2 + \ldots \zeta_{n+1}$$
$$x_2 = \zeta_2 + \zeta_3 + \ldots \zeta_{n+2}$$
$$\vdots$$
$$x_{2n} = \zeta_{2n} + \zeta_{2n+1} + \ldots \zeta_{3n}$$

ζ series of 3n is reduced to x series of 2n terms.

Application of α^2_{n+1} to ζ series is equivalent to the application of α_{n+1} to the x series to generate a new series, say Y, where

$$Y_1 = \sum_{k=0}^{n} x_{k+\ell} \text{ where } \ell = 1, 2, \ldots n$$

Thus,

$$Y_1 = x_1 + x_2 + \ldots x_{n+1}$$
$$Y_2 = x_2 + x_3 + \ldots x_{n+2}$$
$$\vdots$$
$$Y_n = x_n + x_2 + \ldots x_{2n}$$

x series of 2n terms is reduced to Y series of n terms.

Applying α to the Y series means obtaining the sum of n consecutive terms of Y series, to give a term Z_1.

Thus,

$$Z_1 = Y_1 + Y_2 + \ldots Y_n$$

on dividing Z_1 by the normalizing factor $n\ (n+1)^2$, will give the filtered value of the observations involving 3n terms. The above filters can be applied in the following two ways:-

(i) Apply the filter starting with the first term of the observations to give Z_1 and then apply again starting with the second term of the observations to give Z_2 and so on. The Z series will have the same time interval as the original series but several frequencies will have been removed (in the above case higher frequencies will be removed) according to the filter involved.

(ii) Apply the filter starting with the first term of the observations to give Z_1 and then apply again starting with the nth term of the observations to give Z_2 and so on. In this case the Z series will have the time interval $n\Delta t$. Thus using the above filters in this way we can change observational interval from a smaller unit to higher, at the same time smoothing out the unwanted frequencies according to the design of the filter.

Fig. 4 gives response functions of some of the typical filters. These curves suggest that for each filter, various frequencies are filtered in different ratios. Therefore percentage reduction of the various frequencies in each filter should be known so that it can be taken into account in the analysis. The abscissa in Fig. 4 is also given in terms of cycles per day so that the reduction in various harmonies can be computed from the graphs for any particular filter. Table 1 readily gives the percentage reduction in amplitude for different species with respect to each filter.

TABLE 1.

Δt Min	Smoothing Operator (Filter)	Number of Data Lost	Reduction in amplitude of the Species No. in Percentage			
			1	2	4	8
5	$\frac{1}{12^2\cdot 14}\alpha_{12}^2\,\alpha_{14}$	36	·93	3·69	14·31	48·28
10	$\frac{1}{6^2\cdot 7}\alpha_6^2\,\alpha_7$	17	·86	3·69	14·07	47·68
15	$\frac{\;}{4^2\cdot 5}\alpha_4^2\,\alpha_5$	11	·78	3·78	14·46	48·83

Filters can also be applied in reverse sense i.e. a low pass filter can be used as high pass by substracting the filtered value, after normalizing, from the appropriate value of the original data. Similarly a band pass filter can be used as band suppress filter.

<u>Time representation of a filter</u> (making a stencil of a filter)

It is sometimes convenient to represent a filtering operator in the form of a stencil for the use of computers. The method is that the filtering operators are represented by a set of polynomials* which have identical multiplicative and additive properties as filtering operators themselves. Expanding the polynomials and equating zero power as the central observation the required factors to be multiplied to successive observations can be calculated. Polynomials associated with basic filtering operators are given below:

$$A_{2n} \longleftrightarrow x^n + x^{-n}$$

$$S_{2n} \longleftrightarrow x^n - x^{-n}$$

$$2n+1 \longleftrightarrow \sum_{k=-n}^{n} x^k$$

$$**I \longleftrightarrow 1$$

* Known as Aragnol Polynomials.

** I is the unit filtering operator having 1 as multiplier to k^{th} observation and 0 for the rest, where k^{th} observation is the Central observation of a sequence. e.g. If $(I + A_4)$ is applied to a sequence $\zeta_1, \zeta_2, \zeta_3, \ldots$

Applying A_4 will give $\zeta_1 + \zeta_5$
Applying I^4 will give ζ_3 (Note that ζ_3 is the Central observation of this filter). Therefore applying $(I + A_4)$ will give $Z_1 = \zeta_1 + \zeta_3 + \zeta_5$.

In the above n may be an integer or a multiple of $\frac{1}{2}$, indicating spreading out of date points on each side of central observation. Suppose we want to make a stencil of a filter S_2^2. This means application of S_2 filter twice i.e. $S_2\ S_2$. S_2 is represented in polynomials as $(x^1 - x^{-1})$. In this the coefficients of $x^0 = 0$, $x^1 = 1$, $x^{-1} = -1$. In stencil form:

	Multipliers	terms
S_2	0	k^{th}
(In stencil form where	1	$(k+1)^{th}$
k^{th} term taken as	-1	$(k-1)^{th}$
central observation)	0	Remaining

where k is the k^{th} term of the data series.

In terms of data say, $\zeta_1, \zeta_2, \zeta_3, \ldots$ this means generating a new sequence Y so that,

$$Y_1 = \zeta_3 - \zeta_1$$
$$Y_2 = \zeta_4 - \zeta_2$$
$$Y_3 = \zeta_5 - \zeta_3$$
$$\vdots$$
$$Y_{n-2} = \zeta_n - \zeta_{n-2}$$

Note that this filter is spread out to one data point on each side of the central observation. Similarly applying S_2^2 will give in terms of polynomials $(x^1 - x^{-1})^2 = x^2 - 2x^0 + x^{-2}$ which in stencil form:

	Multipliers	terms
S_2^2	-2	k^{th}
(In stencil form where	0	$(k\pm1)^{th}$
k^{th} term taken as	1	$(k\pm2)^{th}$
central observation)	0	Remaining

In terms of data say $\zeta_1, \zeta_2, \zeta_3, \ldots$ the application of S_2^2 means generating a new sequence of Y so that,

$$Y_1 = \zeta_1 - 2\xi_3 + \xi_5$$
$$Y_2 = \zeta_2 - 2\xi_4 = \xi_6$$
$$\vdots$$
$$Y_{n-4} = \zeta_{n-4} - 2\xi_{n-2} + \xi_n$$

Note that this filter is spread out to two data points on each side of the central observation. Similarly it can be shown that the cube of this filter will spread out to 3 data points on each side of the central observation and so on. Thus spreading out of a filter can be calculated from its polynomial notation by equating zero value of n as the central observations and then highest value of $\pm n$ will indicate spreading.

Conversion of a filter to a stencil can be understood more clearly by the following examples :

Example 1

To make a stencil of Doodson Diurnal Filter (x_1) for hourly observations where filtering operator is $A_2\ A_4\ A_6\ S_{12}^2$.

Representing the filtering operator by polynomials

$$A_2\ A_4\ A_6\ S_{12}^2 \longleftrightarrow (x^1+x^{-1})\ (x^2+x^{-2})\ (x^3+x^{-3})\ (x^6-x^{-6})^2$$

Simplifying the right hand side will give

$$= (x^3+x^1+x^{-1}+x^{-3})(x^3+x^{-3})(x^6-x^{-6})^2$$
$$= (x^6+x^4+x^2+1+x^{-2}+x^{-4}+x^{-6})(x^6-x^{-6})^2$$
$$= (x^6+x^4+x^2+1+x^{-2}+x^{-4}+x^{-6})(x^{12}+2+x^{-12})$$
$$= x^{-18}+x^{-16}+x^{-14}+2x^{-12}+x^{-10}+x^{-8}-x^{-6}-2x^{-4}-2x^{-2}-4-2x^2$$
$$-2x^4-x^6+x^8+x^{10}+2x^{12}+x^{14}+x^{16}+x^{18}$$

The right hand side indicates that all the odd terms are zero and if we take central observation for x^0 then the filter spreads out to 18 hourly observations on either side. Therefore x_1 stencil will be:

x_1
(In stencil form, k^{th} term taken as central observation)

Multipliers	terms
0	All odd
-4	k^{th}
-2	$(k\pm2)^{th}, (k\pm4)^{th}$
-1	$(k\pm6)^{th}$
1	$(k\pm8)^{th}$, $(k\pm10)^{th}$, $(k\pm14)^{th}$ $(k\pm16)^{th}$, $(k\pm18)^{th}$
2	$(k\pm12)^{th}$
0	Remaining

Example 2

To make stencil of filtering operator $\frac{1}{6^2 \cdot 7}\alpha_6^2\alpha_7$ for observations 10 minutes interval apart.

The constant $\frac{1}{6^2 \cdot 7}$ can be applied later to the result.

Representing the filtering operator by polynomials.

$$\alpha_6^2\alpha_7 \longleftrightarrow \left[\sum_{k=-\frac{5}{2}}^{\frac{5}{2}} x^k\right]^2 \sum_{k=-3}^{3} x^k$$

$$= (x^{-\frac{5}{2}}+x^{-\frac{3}{2}}+x^{-\frac{1}{2}}+x^{\frac{1}{2}}+x^{\frac{3}{2}}+x^{\frac{5}{2}})^2\,(x^{-3}+x^{-2}+x^{-1}+x^0+x^1+x^2+x^3)$$

$$= x^{-8}+3x^{-7}+6x^{-6}+10x^{-5}+15x^{-4}+21x^{-3}+26x^{-2}+29x^{-1}+30x^0+29x^1$$

$$+26x^2+21x^3+15x^4+10x^6+3x^7+x^8$$

The filter spreads out to 8 observations on either side of central observations which is at x^0.

$\alpha_6^2\alpha_7$

(In stencil form where k^{th} term is taken as central observation).

Multipliers	terms
30	k^{th}
29	$(k\pm1)^{th}$
26	$(k\pm2)^{th}$
21	$(k\pm3)^{th}$
15	$(k\pm4)^{th}$
10	$(k\pm5)^{th}$
6	$(k\pm6)^{th}$
3	$(k\pm7)^{th}$
1	$(k\pm8)^{th}$
0	Remaining

Factors determining the application of filters in basic operation form or in stencil form.

Application of a filter in basic operation form involves as many steps as the number of factors are in the filtering operator. As an example, Doodson Diurnal Filter ($A_2 A_4 A_6 S_{12}^2$) has five factors and its application in basic operation form involves five steps, although each step comprises of either simple addition or subtraction. This filter in stencil form (Example 1 on page (6) requires one step in which multipliers are applied to the respective data. It is therefore inferred that the basic operation form becomes cumbersome for hand calculations as every step requires to be recorded for use in the next step and so on until the final filtered value is obtained. But this method can be conveniently used in computers without printing out values after each step until the final step is completed. Thus basic operation form is recommended for computers and stencil form for hand calculations.

For filters containing α, basic operation form is the most suitable as these filters usually do not have more than 3 factors in them. Moreover, in stencil form α filters generate large multipliers (see example 2 on page 6) making hand calculations slightly difficult.

A multiple filter program (POLYFIL) has been written, using basic operation form, in Fortran for KDF9. It is a general program involving basic filtering operators α_n, A_n, S_n and I where n = 1, 2 ... 99. For details see the instructions for POLYFIL.

Application of Fig. 4 to practical problems

Four Low Pass filters are illustrated in Fig. 4. These filters are useful for eliminating higher frequencies, like seiches from a tidal record, and for smoothing current observations for tidal stream analysis. Following examples illustrate the application of Fig. 4 to such problems:

Example 1

A tidal record of 5 minutes interval is contaminated with seiches of 45 minutes period. Select the best suitable filter from Fig. 4 to smooth out the seiches.

Period of seiches is 45 minutes, therefore the frequency of seiches = 1·33 c.p.h.

In Fig. 4 read abscissa on scale Δt = 5 Min. (Sampling Interval).

1·33 c.p.h. on this scale is marked with dotted arrow; produce a line from this point on curve C and D and read the responses on the vertical scale.

Response for filter C = ·02

Response for filter D = ·36.

Therefore the reduction in seiches amplitude for filter C is 98% and for filter D is 64%.

Hence, filter C is suitable.

From the figure it is clear that A and B are also suitable filters as their cut off points lie on the left of the seiche frequency. But there are certain advantages in using filter C over A and B, such as:

(i) The reduction in the amplitude of the tidal species is less in C as compared to A and B. The most reduction is in A.

(ii) Fewer data points are required for the application of filter C as compared to A and B and hence less computation. The largest number of data points are required for filter A.

Example 2

Current observations taken at 10 minute intervals are required to be smoothed and converted into hourly observations for tidal stream analysis. Select a suitable filter for this purpose from Fig. 4.

In Fig. 4 read abscissa on scale marked c.p.d. (left hand corner). Against Δt = 10 min. (sampling interval) various harmonics are marked up to 12. If the tidal stream analysis is restricted to species up th 6th diurnal, produce a line from 6th diurnal point on the scale on curves B, C and D.

Compute responses from vertical scale which are:

B = ·18

C = ·72

D = ·88

Therefore reduction in the amplitude of 6th diurnal species will be:

B = 82%

C = 28%

D = 12%

Hence filter B is not suitable. The choice is between filters C and D. It can be seen from the Fig. 4 that filter D would allow more high frequencies to pass than filter C. Filter D cuts off at 2·3 c.p.h. and filter C at 1·4 c.p.h. Otherwise there is not much difference in both these filters, as far as computation effort is concerned. Only two data points more are required for computation of filter C as compared to D. Therefore filter C should be used for this problem.

In order to convert the data into hourly intervals, apply the filter starting with 1st observation and record the filtered value which is appropriate to the 9th observation. Next apply the filter starting from 7th observation i.e. after six 10 minute intervals, and so on.

RESPONSE OF FILTERS A (LOW PASS) & S (HIGH PASS).

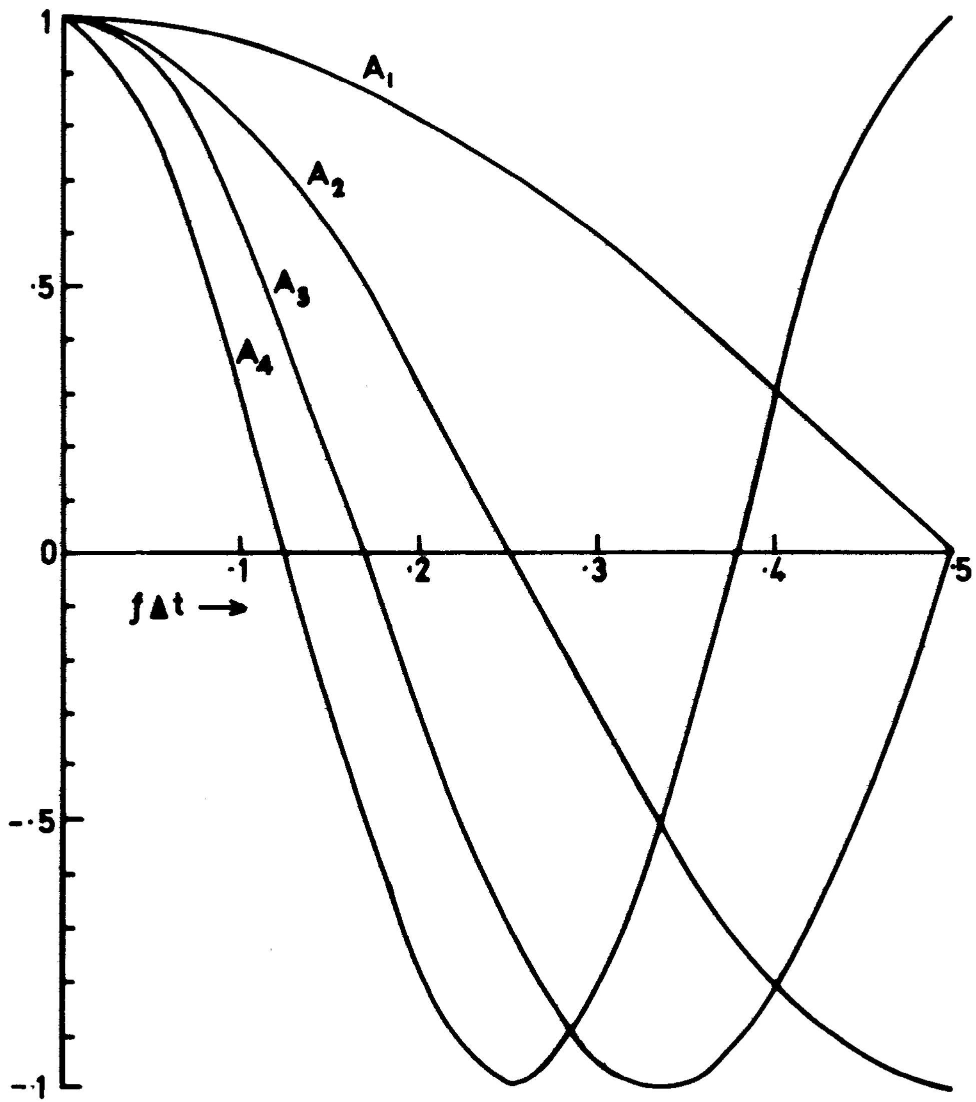


FIG. 1A.

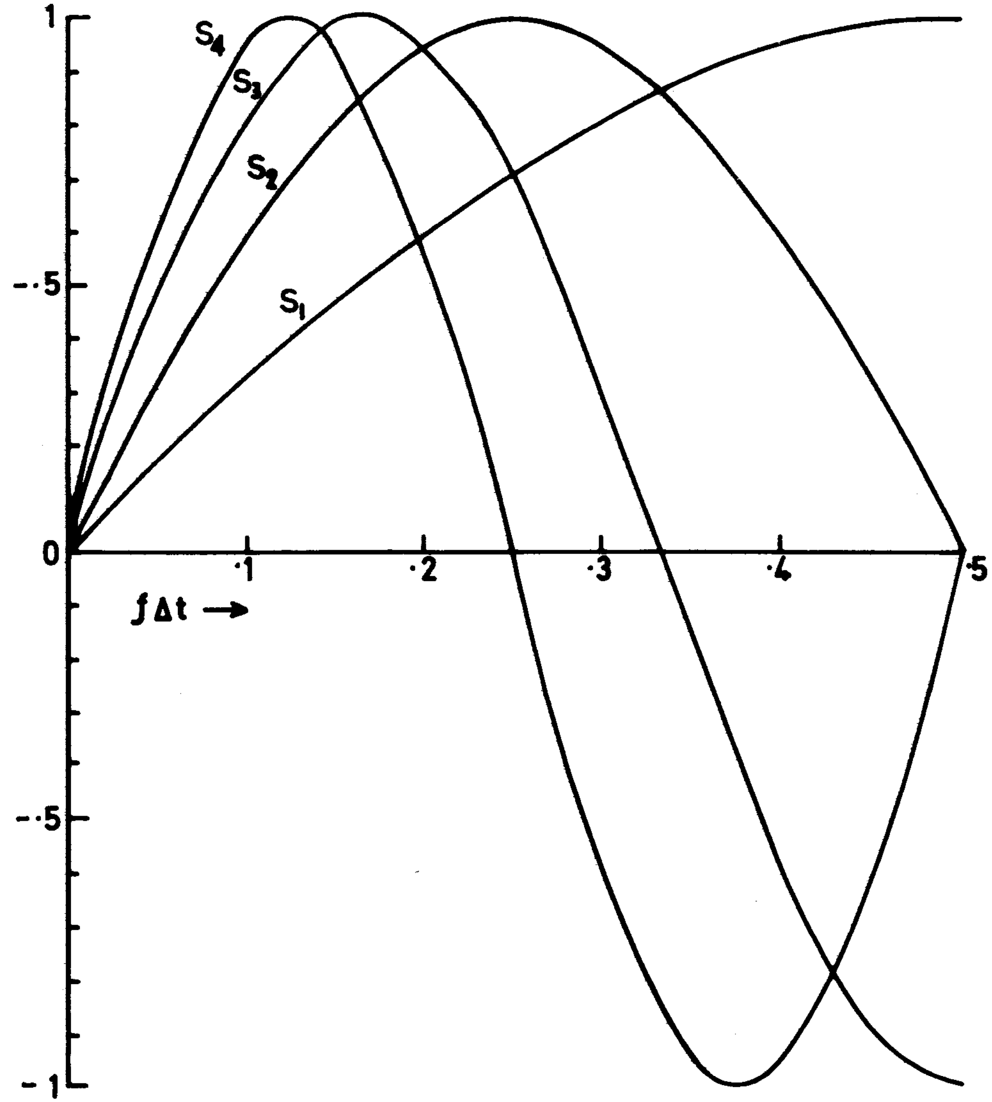


FIG. 1B.

RESPONSE OF THE POWERS OF FILTERS A & S

FIG. 2A.

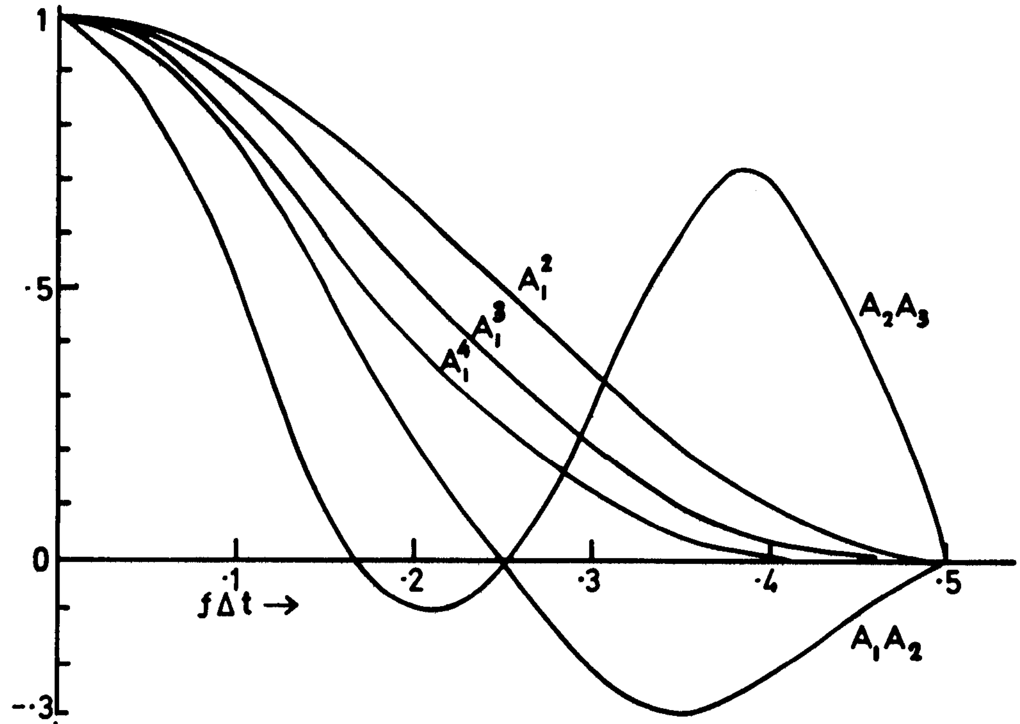


FIG. 2B.

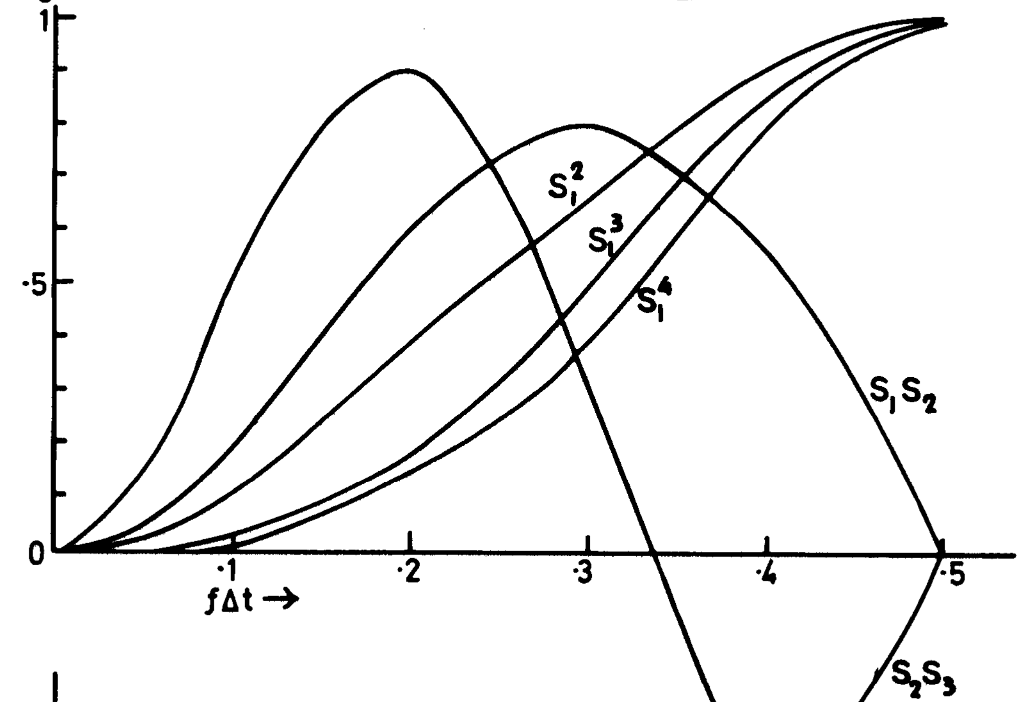
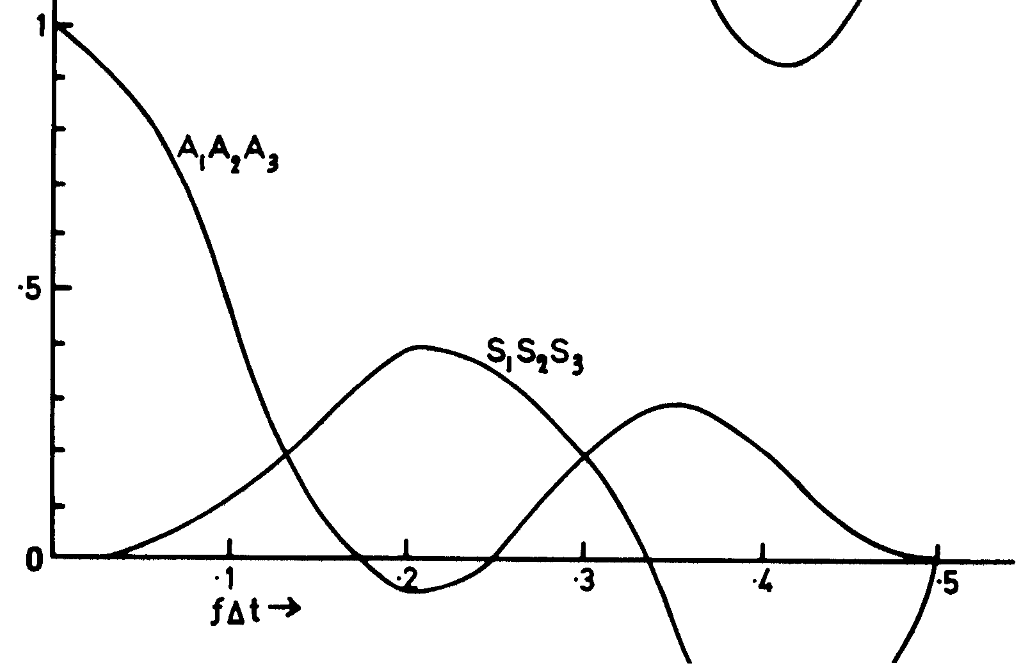


FIG. 2C.



RESPONSE OF THE PRODUCTS OF FILTERS A & S

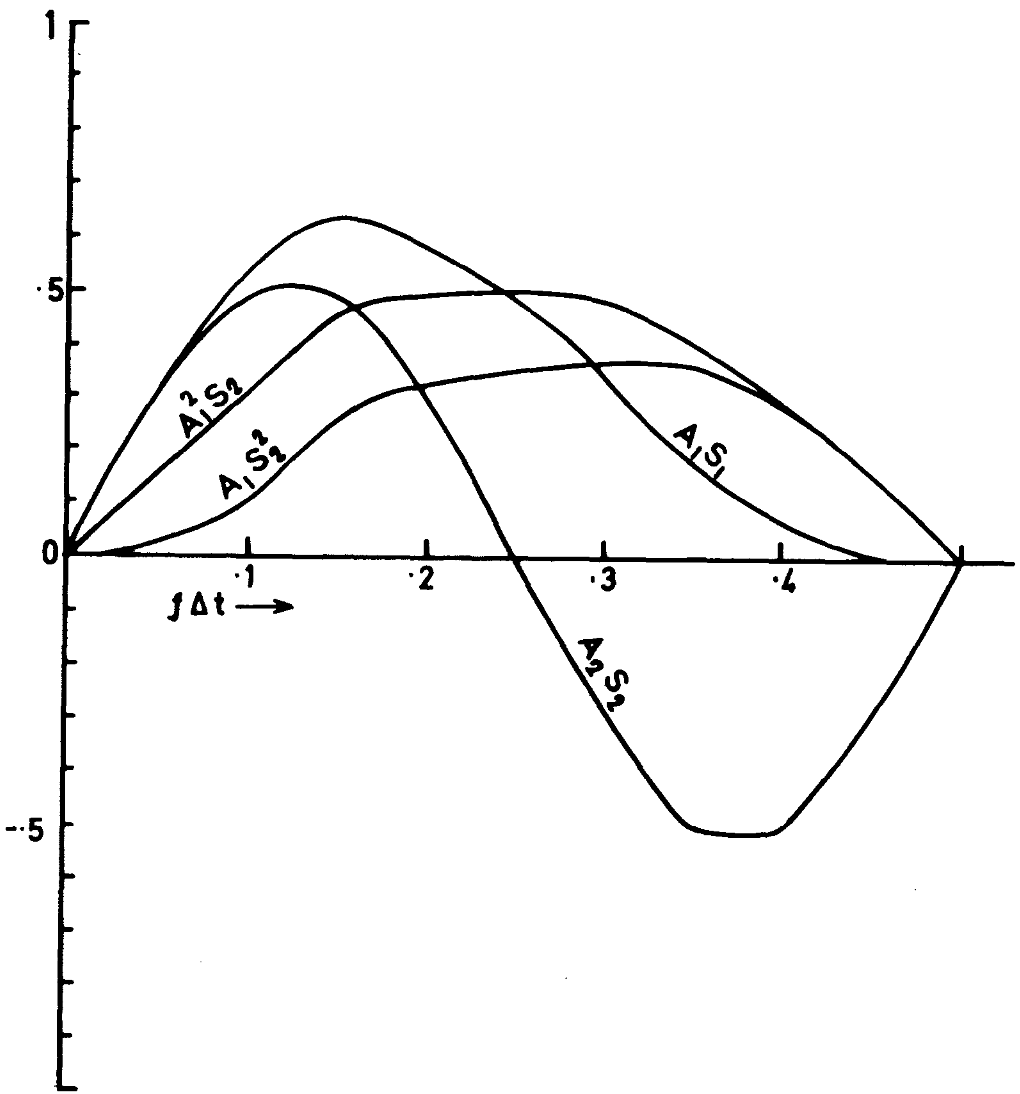


FIG. 3A.

RESPONSE OF FILTER α AND ITS PRODUCTS

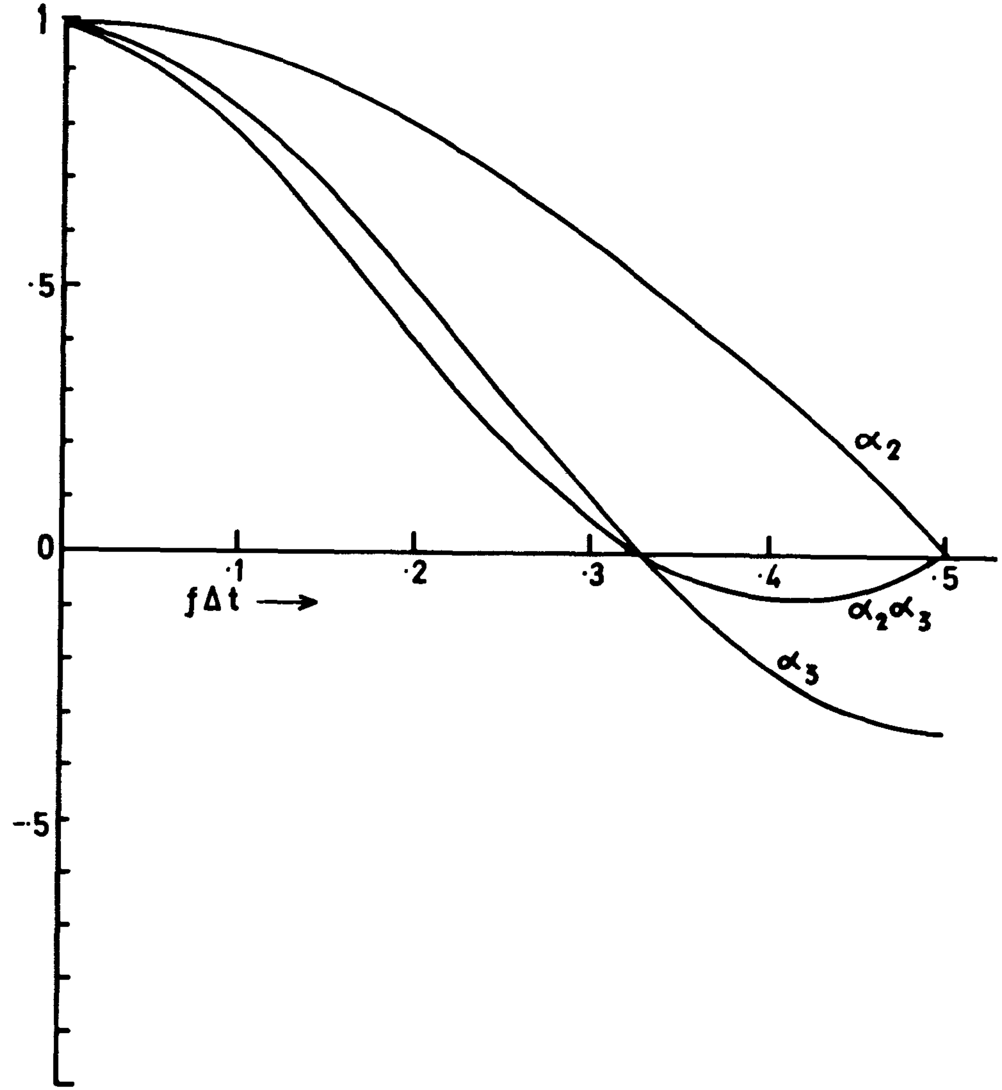


FIG. 3B.

SMOOTHING FILTERS FOR TIDE AND CURRENT OBS.

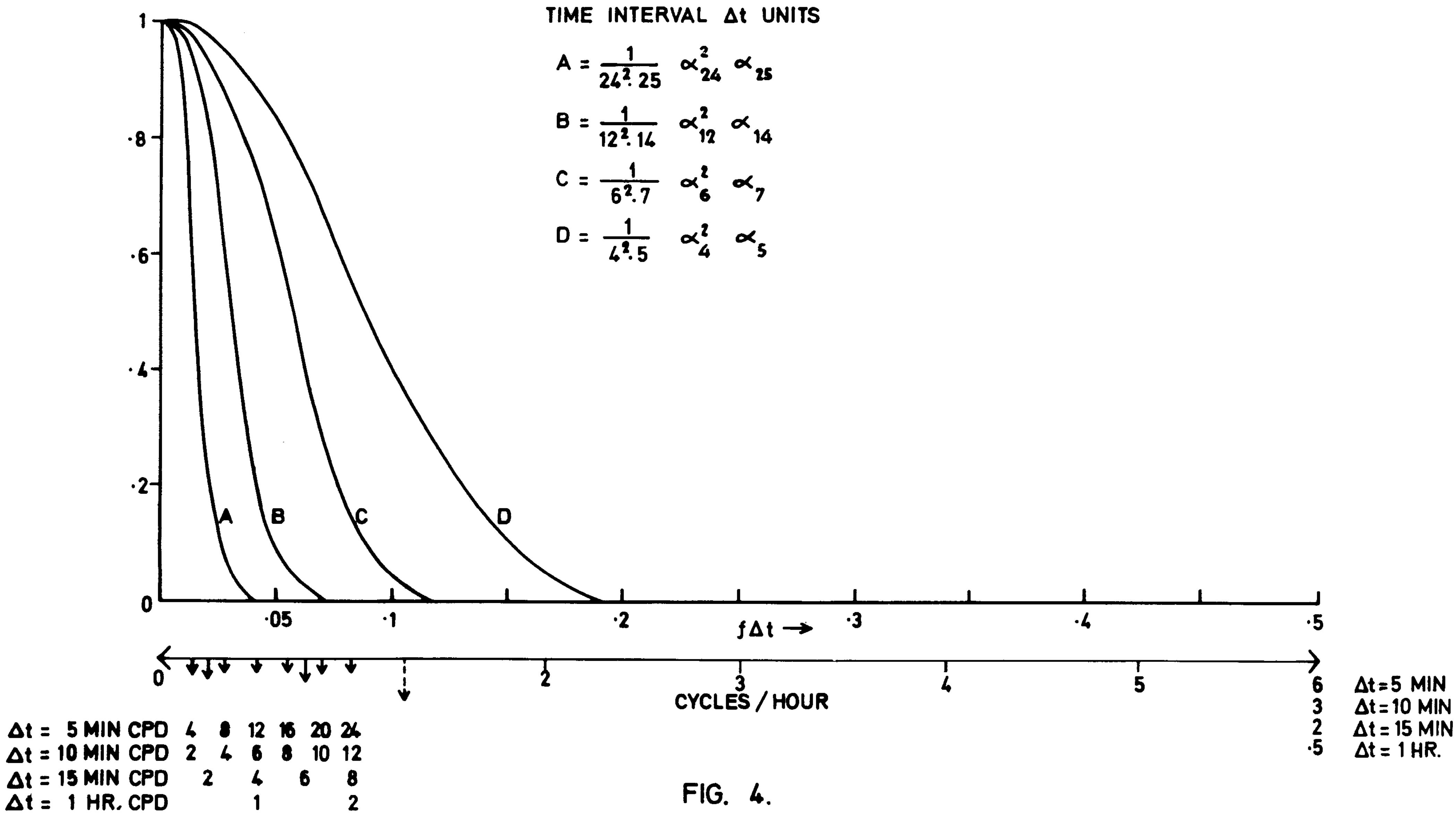


FIG. 4.