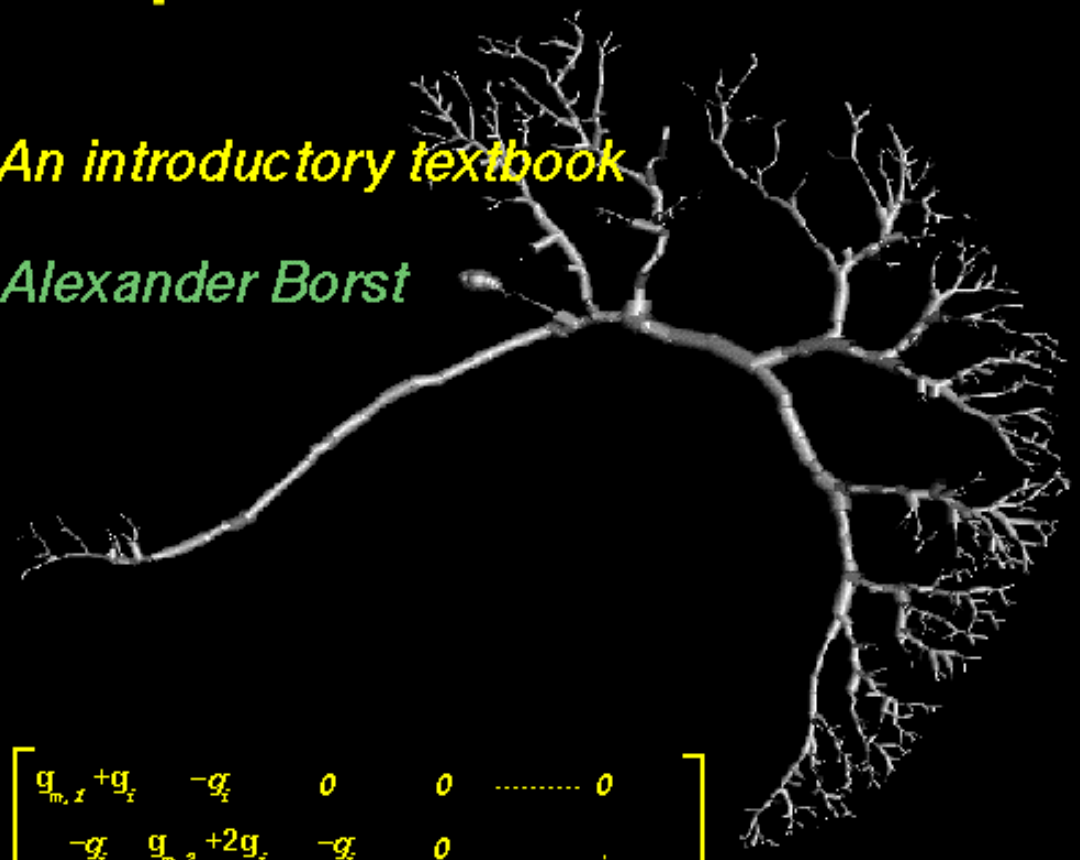
Mathematical Techniques in Computational Neuroscience



An introductory textbook

Alexander Borst

$$\begin{bmatrix} g_{m,1}+g_i & -g_i & 0 & 0 & \cdots & 0 \\ -g_i & g_{m,2}+2g_i & -g_i & 0 & & \\ 0 & -g_i & g_{m,3}+2g_i & -g_i & & \\ \vdots & & & & & \\ 0 & 0 & 0 & & g_{m,N}+g_i & \end{bmatrix} \quad \vec{V} = \vec{R}$$

TABLE OF CONTENT:

Chapter 3: Information Theory

Chapter 4: Vectors and Matrices

Chapter One: Linear Filter Theory

1.1. The Convolution Theorem

1.1.1. Everybody has done a convolution

Everybody has done a convolution. Yet most people do not know about it and would heavily contradict when confronted with this statement. Suppose you had collected some data, maybe one after the other, and put this data in a spreadsheet (Tab.1). The leftmost column holds the numbers which order the data in the column next to it. So, we can now refer each of them individually as data[i]. Now, we plot them and this gives us the black line in the graph of Fig.1.1. As one can easily realize, these data 'have something in them', but before we show them to our instructor, we chose to make them look a bit nicer, in order to make our finding clearer. So, we click on the 'smooth' button, select '3 point smoothing' and we get the red line in fig.1 and the respective data values in the right next column. That looks already better, and we try '5 point smoothing'. This gives us the green line in the graph shown in Fig.1.1 and the right most column in the spreadsheet of Tab.1. Now, we are satisfied, because we have removed almost all of the glitches in the data. The step-like increase in the middle of the graph becomes clearly visible.

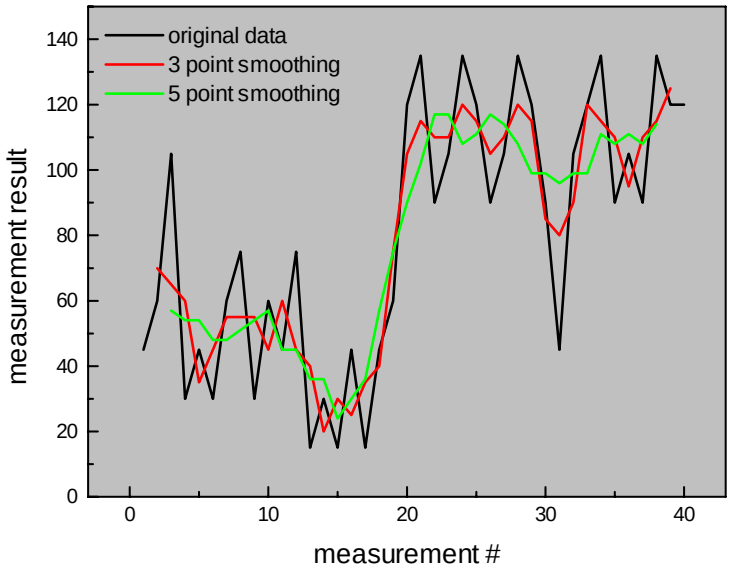
What have we done? For each index i, we have taken three neighboring data values x(i), added them together, divided the result by three, and put the result of that, y(i), in the next column at the index around which our three points were centered.

$$y(i) = (x(i-1) + x(i) + x(i+1)) / 3;$$

To obtain the result of the 5 point smoothing, we did the same, but this time we added 5 data values together instead of three, and, consequently, divided the result by 5, before entering it in the rightmost column at the corresponding row given by the center index.

$$z(i) = ((x(i-2) + x(i-1) + x(i) + x(i+1) + x(i+2)) / 5;$$

i	data	sm3	sm5
1	45	--	--
2	60	70	--
3	105	65	57
4	30	60	54
5	45	35	54
6	30	45	48
7	60	55	48
8	75	55	51
9	30	55	54
10	60	45	57
11	45	60	45
12	75	45	45
13	15	40	36
14	30	20	36
15	15	30	24
16	45	25	30



Tab1. *Some data have been collected which obviously have some obvious noise in them. The table shows indices of the data, the data values themselves and the result of 3 and 5 point smoothing of the data, respectively, in the consecutive columns.*

Fig.1.1 *Graph of the three columns of table 1 plotted as a function of the index i in the leftmost column.*

We did that for all data values data(i) from i=2 to 39 in case of 3 point smoothing and from i=3 to i=38 in case of 5 point smoothing. The reason why we did not work at the border values was that we did not know what to do when the data values, e.g. at i-2, were missing. So we left them blank. We can say, that we used something like a 'moving window' of the form

$$g(-1) = g(0) = g(+1) = 1/3;$$

to obtain the result in the 3 point column. The values g(k) gave us the weight by which the data values data(i) entered the sum. But now we want to try a more general tripod, e.g. a triangular one which gives more weight to the central data value:

$$g(-1) = 1/5; g(0) = 3/5; g(+1) = 1/5;$$

Using these weights g(k), without explicitly saying how they look like, i.e. what numbers they hold, we can rewrite our procedure as follows:

$$y(i) = x(i+1) \cdot g(-1) + x(i) \cdot g(0) + x(i-1) \cdot g(1);$$

This expression is already a bit long, so we want to use an abbreviation. The formula can be rewritten in a more compact form by using the Σ sign to indicate that we have a sum of terms. Each term is a product, and the index has to go from -1 to +1 in this explicit case. This reads:

$$y(i) = \sum_{j=-1}^{j=+1} x(i-j) \cdot g(j);$$

Suppose now that we would like to have a formula 'for all seasons', that we do not have to change every time we are playing with some other smoothing algorithm. One difference between the 3 and the 5 point smoothing algorithm was the number of terms in the sum. If we say that the weighting function holds zeros unless otherwise stated, we can as well go from minus to plus infinity to have to most general form of the operation we performed. Thus, our function now becomes: (for arbitrarily long window functions):

$$y(i) = \sum_{j=-\infty}^{j=+\infty} x(i-j) \cdot g(j)$$

As the reader can guess by now, what we arrived at is called an convolution in discrete steps. For smooth functions, this expression turns into the convolution integral:

$$y(i) = \int_{j=-\infty}^{j=+\infty} x(i-j) \cdot g(j) dj \qquad (1.1)$$

1.1.2. Some examples

Let´s play a little bit to see what we can do with it. First, let us consider a step function as input. This function is defined as being 0 for x<=0, and equal to 1 for x>0. It is also called 'heavy side function'. As filters we define the following:

1. $g_1(x) = \begin{cases} 1/(2k) \; for -k < x < +k \\ 0 \quad else \end{cases}$

2. $g_2(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-x^2}{2\sigma^2}}$

3. $g_3(x) = \frac{-x}{\sqrt{2\pi\sigma^6}} e^{\frac{-x^2}{2\sigma^2}}$

4. $g_4(x) = \frac{-1}{\sqrt{2\pi\sigma^6}} \left(1 - \frac{x^2}{\sigma^2}\right) e^{\frac{-x^2}{2\sigma^2}}$

Our first filter is a box filter, or window and has the black weighting function in Figs.1.2 and 1.3. The second filter is a gaussian function and is shown in red in the following. It has some important features or properties: i) It is symmetric around 0. ii) The integral from - to + infinity over a gaussian equals 1, and iii) the width of the gaussian function is given by σ. The factor in front of e takes care of the fact, that, no matter what value σ has, the integral will always be 1 (the broader the function, the smaller its peak amplitude). Loosely speaking, it can be looked upon as a smooth version of the box filter. The third filter, $g_3(x)$ is the first derivative of the gaussian function and is shown in green. In contrast to the gaussian filter, it is asymmetric. The fourth filter $g_4(x)$ is the second derivative of the gaussian filter. It again is symmetric. In appendix a.1., the reader can look up how the derivatives are calculated. The filter g_3 can also be approximated by the taking difference between two gaussian functions shifted against each other by some small amount ε: $g_3(x) \sim g_2(x-\varepsilon) - g_2(x+\varepsilon)$; In a similar way, the filter g_4 can be approximated by the difference of two gaussians, both centered at 0, but with different widths σ: $g_4(x) \sim g_{2,\sigma 1}(x) - g_{2,\sigma 2}(x)$. The later filter, therefore, is also called a *d*ifference-*o*f-*g*aussians or DOG filter.

Now, when applied to the step input function, we see the following (Fig.1.2): The box filter turns the step input into a ramp. The gaussian filter acts in a similar way, but instead of a

ramp with sudden on- and offsets, the resulting output function exhibits a gradual or smooth increase in slope on the left and again a smooth decrease of slope at the right side. The output function of the g_3 filter in response to the step input is a gaussian. We may note that reflects somehow the spatial derivative or slope of the smoothed version of the input function. If we follow this kind of description, we can say that the output of filter g_4 then reflects the slope of the output of the filter g_3 and, consequently, the curvature of a smoothed version of the input function. Now, we use, instead of a step input, an impulse input (Fig.1.3). When considering the result of the convolution with the very same filters, we realize, that the responses of the filters to an impulse input looks identical to their weighting functions. Therefore, the weighting function is also called the impulse response of a filter.

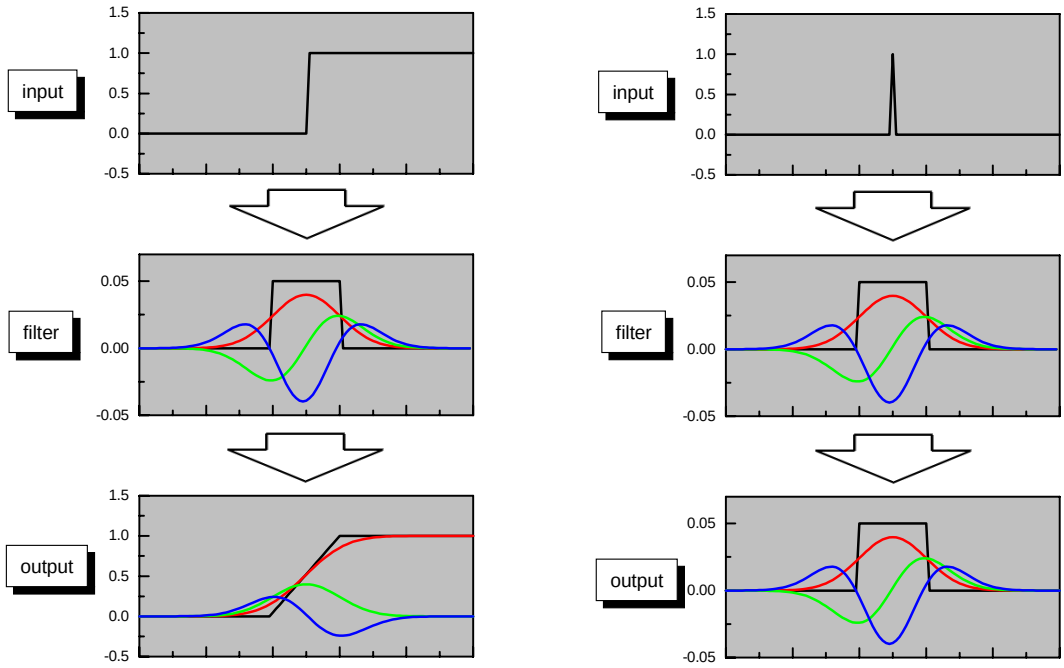


Fig.1.2: *Responses of various filters (box = black, gaussian = red, 1st derivative gaussian = green, 2nd derivative gaussian = blue) to a step input function.*

Fig.1.3 *Responses of various filters (box = black, gaussian = red, 1st derivative gaussian = green, 2nd derivative gaussian = blue) to an impulse function. Note that the output functions look identical to the weighting functions of the filter.*

Before discussing some important properties of the convolution, we might ask ourselves: What is weighting function or the impulse response of a filter that preserves the input function exactly? The answer is: The Dirac pulse. Without knowing anything about the Dirac function, we might intuitively say, that the filter should be equal to 1 at position 0 and zero anywhere else. This is absolutely correct, as long as we work with discrete functions like in the examples above. We might, e.g. assign measures to the index values, like e.g. degrees of visual angle. What if another function comes along where the precision is in 10^{th} of degree? If we keep our function in degrees and if our function should still preserve the input image, it should have a value of 10 but only for 0<x<0.1. We see that the smaller the range gets, the higher the value of the pulse. The Dirac function expresses this fact by saying that

$$\delta(x) = \begin{cases} \lim_{\Delta x \to 0} \frac{1}{\Delta x} & for 0 < x < \Delta x \\ 0 & else \end{cases} \qquad (1.2)$$

This Dirac pulse is of great importance for the formal treatment of linear filters. For pratical purposes, i.e. when considering functions of discrete variables, we do not have to worry about the limit value of δ(x), but only have to keep in mind that the integral, i.e. the area under the pulse equals 1 no matter how small we pick the pulse to be.

1.1.3. Linearity of convolution

We now come to an important feature of the convolution, and that is its linearity. In general linearity is defined in the following way: the function value given the sum of two inputs a and b is equal to the function value given one input a plus the function value given the other input b: f(a+b) = f(a) + f(b). Furthermore, the function value given n times an input a is the same as n times the function value given this input a, or: f(na) = nf(a). Applied to the convolution, these two relations exactly hold true:

$$y(i) = \int_{j=-\infty}^{j=+\infty} (a(i-j) + b(i-j)) \cdot g(j) dj = \int_{j=-\infty}^{j=+\infty} a(i-j) \cdot g(j) dj + \int_{j=-\infty}^{j=+\infty} b(i-j) \cdot g(j) dj ; \quad (1.3)$$

Therefore, since any function can be represented as an infinite sum or series of impulses, the impulse response of the filter is a complete description of the system. Given the system is linear, once we know the impulse response, we know everything. There is nothing more that can be learned.

In a similar way, linearity also holds true with respect to the filters' weighting functions:

$$y(i) = \int_{j=-\infty}^{j=+\infty} a(i-j)\cdot(g_1(j) + g_2(j))\,dj = \int_{j=-\infty}^{j=+\infty} a(i-j)\cdot g_1(j)dj + \int_{j=-\infty}^{j=+\infty} a(i-j)\cdot g_2(j)dj\,;$$

1.1.4. Differentiating Filters

The reader might have noted previously, that interestingly, the response of the filter g_3, i.e. the 1st derivative gaussian, looked like the 1st derivative of the smoothed input, and the response of the 2nd derivative gaussian g_4 was pretty much like the 1st derivative of the output of g_3. What is behind is the following relation: Given

$$y(i) = \int_{j=-\infty}^{j=+\infty} x(i-j)\cdot g(j)dj \qquad \text{then}$$

$$\frac{dy(i)}{di} = \int_{j=-\infty}^{j=+\infty} \frac{dx(i-j)}{di}\cdot g(j)dj = \int_{j=-\infty}^{j=+\infty} x(i-j)\cdot\frac{dg(j)}{di}dj\,; \qquad (1.4)$$

Therefore, the 1st derivative of a gaussian is said to calculate the 1st derivative of the smoothed input, and the 2nd derivative gaussian calculates the 2nd derivative of the smoothed input.

1.1.5. Filters in Sequence

Imagine having two filters in a row with g_1 and g_2 denoting their weighting functions. Thus, the output of the first filter y(t) is the input to the second filter. The output of the second filter z(t) can then be calculated according to:

$$y(i) = \int_{j=-\infty}^{j=+\infty} x(i-j)\cdot g_1(j)dj\,;$$

$$z(i) = \int_{k=-\infty}^{k=+\infty}\int_{j=-\infty}^{j=+\infty} x(i-j-k)\cdot g_1(j-k)g_2(k)d(j-k)dk\,;$$

This is, however, equivalent, to

$$z(i) = \int_{k=-\infty}^{k=+\infty} x(i-k) \int_{j=-\infty}^{j=+\infty} g_1(j-k) g_2(k) d(j-k) dk ; \qquad (1.5)$$

The consequence is that instead of taking the output of the 1st filter as the input to the second one, we can also convolve the impulse response of the first filter with the one of the second and, thus, get a new impulse reponse g12 with which we now can convolve the input x(t) directly.

1.1.6. Filters in the frequency domain

We now consider the different filters in the frequency domain without explicitly introducing the Fourier transform at this point. For this purpose, we now use a sinusoidal input function of a given circular frequency ω which is defined as 2π times the frequency f in our coordinate system. So, whenever our measure of the axis equals one, the sinus function has gone through one full cycle. The period or cycle length T of this sinusoid is inversely proportional to the frequency f: f = 1/T.

Consider now a sinusoid being convoluted by a box filter. When the frequency is low, i.e. the filter 'sees' only a small proportion of a cycle, the output will be modulated with about the same amplitude as the input is. Thus, low frequencies pass through a box filter well. In contrast, consider a sinusoid with a high frequency so that many cycles are comprised in the window of the filter. The filter output will always be zero or close to zero, no matter which part of the input is presented. Thus, high frequency input signals have a small output amplitude. The same considerations hold true for a gaussian filter. That is why a box or a gaussian filter are called *'low-pass filters'*.

Let us now consider the same situation for a DOG filter. When the frequency of the input is low, the negative and the positive parts of the weighting functions will contribute about equally well to the output. When the negative and positive areas are balanced, i.e. when the integral over the weighting function is zero, the output to low frequency input signal will be zero, too. When, on the other hand, the input frequency is high, so that many cycles are within the central trough of the filter, these will be averaged out in the same way as they become zero in case of the above considered low-pass filters. Thus, the output of the DOG filter is zero for high as well as for low frequencies. So, somewhere in between, there should be a

frequency range where we expect to see some responses, and this is exactly at this frequency where the sinusoid matches the positive and negative modulations of the weighting function. Since the DOG filter allows only a certain frequency band to pass through, it is called a *'band-pass filter'*.

We could also make a formal argument to arrive at this conclusion using the above given formula about the linearity of convolution with respect to different filters. When the DOG filter is equal to the difference between two gaussians having different widths, then the frequencies represented at the output should be exactly equal to the difference between the output functions of the two gaussians. Since they both differ with respect to their width, their onset of action will be at different frequencies (going from low to high frequencies). This window will exactly describe the frequencies that pass through the DOG filter. In the example shown in Fig.1.4., these various points are illustrated using a sinusoidal frequency ramp as input signal.

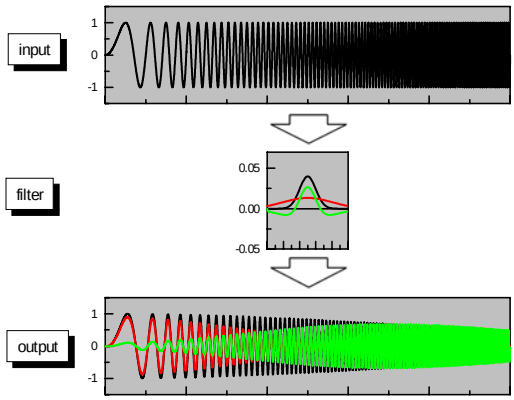


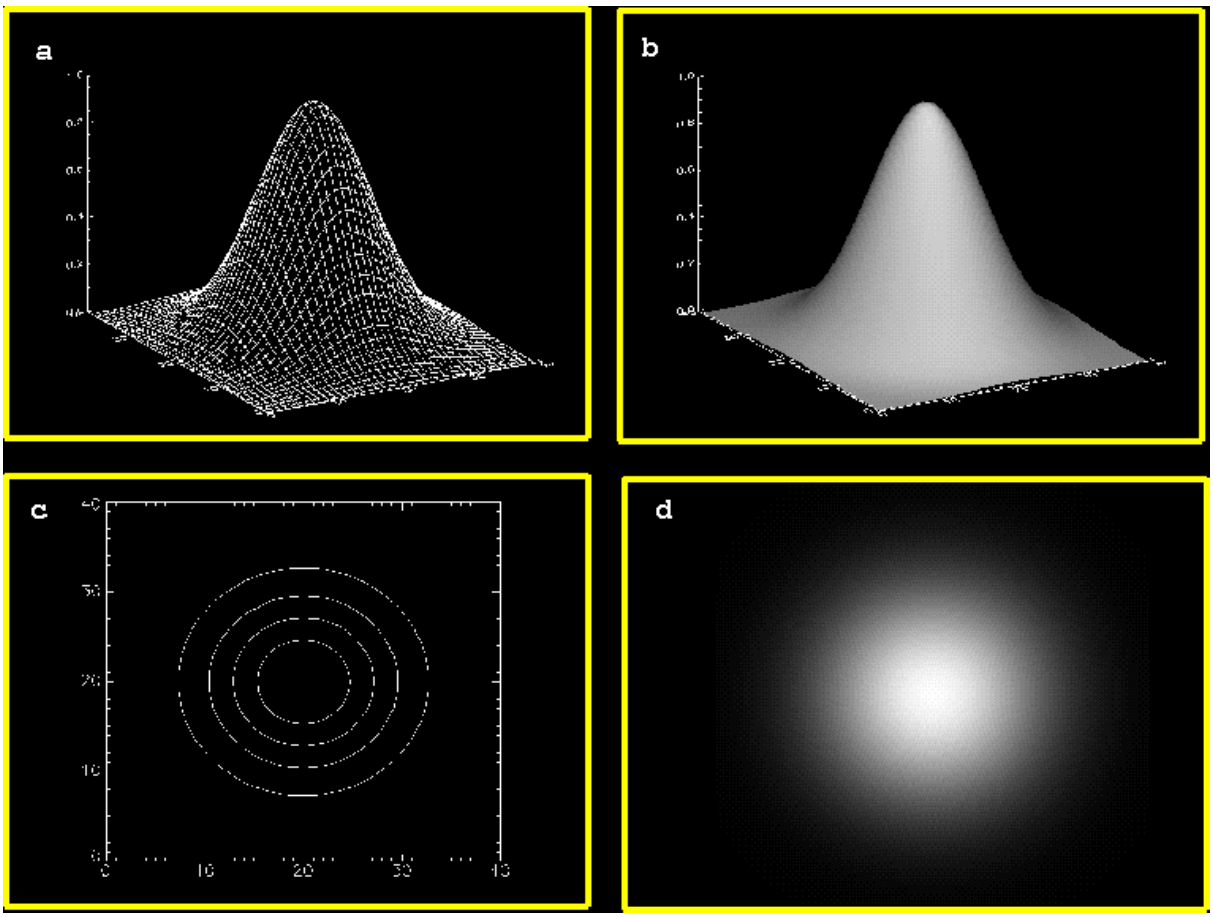
Fig.1.4 *The input function consist in a frequency ramp and is defined as sin(ωi) with ω increasing linearly from 0 to twice the window size of the spatial filters. Note that spatial filters are shown expanded in the x-direction by a factor of 20 with respect to the input and output functions. Three different filters are used: a narrow and a broad gaussian filter, and a difference-of-gaussians (DOG) filter. The output of the first two filters can be seen to be damped only in the higher frequency range, whereas the output of the DOG is affected both at low and at high frequencies.*

1.2. Spatial Filters and Neural Receptive Fields

1.2.1. How to visualize functions of two variables

The concept of filters will now be expanded from one to two dimensions. Before considering various of such filters however, different ways to graph or visualize such functions of two variables f(x,y) shall be explained. One way, often used, is to plot the function as a three dimensional grid deformed in accordance with the actual function values (Fig.1.5a). Thus, the height of each grid node represents the function value of the x-y pair lying underneath it in the x-y plane. In a similar way, the function can also be displayed as a solid landscape (Fig.1.5b). The surface between the function values is interpolated and shaded to give a better 3D impression. While this is certainly a fairly esthetic way to show a function of two variables, it is only useful as long as the function is not complicated, like the example used in Fig.1.5. However, when the function has several peaks and troughs, the latter might become hidden behind the peaks. In this case, a contour plot with several height lines might be a better way to visualize the function (Fig.1.5c). All readers will be familiar with this kind of graph from hiking maps. The disadvantage from this kind of illustration, however, is the poor resolution with respect to the actual function values due to the limited number of contour lines. This limitation is overcome by the representation of the function as an image where the gray level at each pixel is proportional to the function value at this x-y location (Fig.1.5d). Often, a gray level scale is given along with such images indicating the exact function values corresponding to each gray level. Since gray level images are very sensitive to the quality of the actual reproduction in a publication, many authors tend to use so-called ‘false color images’, where instead of a gray level coding of the function values, a color map is being used (e.g. from cold, i.e. blue, to warm, i.e. red, colors).

Fig.1.5 *Different ways to graph a function of two variables f(x,y). Here, a 2D gauss function is plotted in 4 different ways.* ***a*** *3D-surface* ***b*** *3D shaded surface* ***c*** *contour lines* ***d*** *image*



1.2.2. Spatial filters in two dimensions

From what we have learned about linear filters in our previous examples, we can now go straight to filters in two dimensions. There is only one new distinction that we have to make that does not exist in case of one-dimensional filters, that concerns the spatial layout of the filters. This distinction is between isotropic and anisotropic filters.

Spatial filters are called ‘isotropic’ when they are radially symmetric. As a consequence of the radial symmetry of isotropic filters, there is no preferred orientation in the input function (which we will also call ‘input image’) with respect to the filter operation. Thus, isotropic filters are ‘orientation blind’ and treat vertical contour lines the same as horizontal ones or those of any other orientation. In contrast, anisotropic filters discriminate between the two spatial axes and all possible orientations.

Let us consider the following example of a spatial gauss-filter with an impulse response of the form:

$$f(x,y) = \frac{1}{\sqrt{2\pi\sigma_x^2}} e^{-\frac{x^2}{2\sigma_x^2}} \cdot \frac{1}{\sqrt{2\pi\sigma_y^2}} e^{-\frac{y^2}{2\sigma_y^2}} \;;$$

This is simply the product of two gaussian filters, one for the x-axis, the other for the y-axis. The formula can be further simplified to the following expression:

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y} e^{-\left(\frac{x^2}{2\sigma_x^2}+\frac{y^2}{2\sigma_y^2}\right)};$$

If $\sigma_x = \sigma_y$, then this filter is radially symmetrical and can be rewritten as:

$$f(x,y) = \frac{1}{2\pi\sigma} e^{-\frac{x^2+y^2}{2\sigma^2}} \;; \qquad (1.6)$$

Such a filter is shown in Fig.1.5.

In contrast, if $\sigma_x \neq \sigma_{y,}$ then the filter will be anisotropic. If one, e.g. chooses σ_x much larger than σ_y, then the blurring effect such a filter has in the x dimension would be much more pronounced than it has on the y dimension.

How do the higher derivatives of the radially symmterical gauss function look like? For that, we need to partially differentiate the above given formula with respect to x or y:

$$\frac{\partial f(x,y)}{\partial x} = \frac{1}{2\pi\sigma} e^{-\frac{x^2+y^2}{2\sigma^2}} \cdot -\frac{2x}{2\sigma^2} = \frac{-x}{2\pi\sigma^3} e^{-\frac{x^2+y^2}{2\sigma^2}}; \qquad (1.7)$$

$$\frac{\partial f(x,y)}{\partial y} = \frac{1}{2\pi\sigma} e^{-\frac{x^2+y^2}{2\sigma^2}} \cdot -\frac{2y}{2\sigma^2} = \frac{-y}{2\pi\sigma^3} e^{-\frac{x^2+y^2}{2\sigma^2}}; \qquad (1.8)$$

These are the 1st derivatives along either the x- or the y-axis. The next formula gives the 2nd derivative of a gauss function which is also called the 'Laplacian' of a gauss:

$$\frac{\partial^2 f(x,y)}{\partial x \partial y} = \nabla f(x,y) = \frac{-2x}{4\pi\sigma^3} e^{-\frac{x^2+y^2}{2\sigma^2}} \cdot \frac{-2y}{2\sigma^2} = \frac{2xy}{3\pi\sigma^5} e^{-\frac{x^2+y^2}{2\sigma^2}}; \qquad (1.9)$$

As an example for how such filters affect an input function, an image of a friendly looking young man is shown in fig.1.6 together with 4 different spatial filters and the resulting output functions. Keep in mind, that all these images are in fact simply discrete functions f of x and y, numbers written in a spread sheet. To some of the readers this might sound trivial, but images mean so much to our visual system and are so prone to all kinds of illusions that one should always be careful. The input image is shown to be filtered by a isotropic gauss filter, two differently oriented anisotropic 1st derivative gauss filters and one 2nd derivative gauss or 'Laplace' filter. The resulting output images are strikingly different. As expected, the isotropic filter to the left has a strong blurring effect on the image. The horizontally oriented 1st derivative gauss seems to emphasize the horizontal edges preferentially. Thus, vertical edges in the input image seem to vanish in the output image. Furthermore, horizontal edges in the output image also differ with respect to their sign: when local function values decrease from top to bottom, i.e. where white areas are at the top of black ones, the resulting edges appear dark in the output image. When local function values increase, the resulting edges

appear white. The same holds true for the resulting output image for the vertical filter, respectively. No such distinction between differently oriented edges is made by the 2nd derivative isotropic gauss filter, shown to the right. As in the previous two examples, the absolute function levels are lost at the output image wherever the input areas are fairly uniform. In other words, the DC luminance level is thrown away by all these filters except for the gaussian low-pass, shown to the left. This is, why these filters are said to possess band-pass characteristics. In the rightmost example, the second derivative gauss turns edges in the original image into zero-crossings of the output image (the reader might go back to Fig.1.2 to see this more clearly in one dimension). It does this for all orientations of contour lines. Depending on the particular spatial outline of the filter, such filters can also be used to improve the contrast of an input image and, therefore, are used for 'sharpening' images in various image processing software. If such software is available, the reader might want to play a bit with various spatial filters and also might want to subtract, e.g., a blurred image from the original one in order to see what happens with the output image.

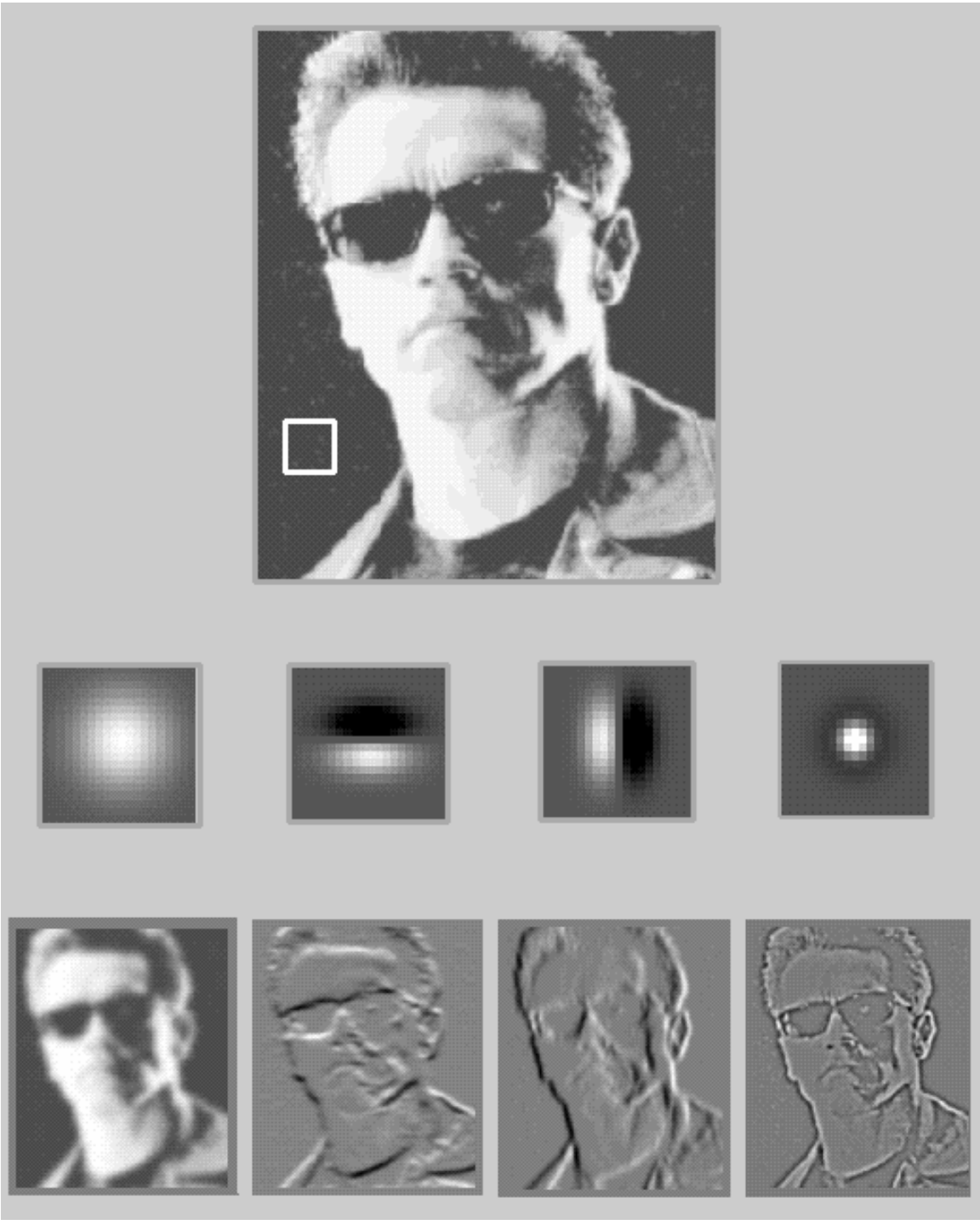


Fig.1.6 *The effect of different spatial filters on an image. The top image (274x326 pixels) serves as input functions for four different filters displayed in the second row. The filters are shown enlarged relative to the original image and have a size of 20x20 pixels. The true filter size is indicated by the white square in the input image. The result of the different filters are shown in the bottom row(from left to right: gaussian filter, a horizontal 1st derivative filter, a vertical 1st derivative filter, a 2nd derivative or DOG filter).*

1.2.3. Impulse responses as neural receptive fields

We now come to the crucial interpretation of the impulse response as the receptive field of a nerve cell. In the original derivation of the convolution, we stated that the result of the convolution can be understood as shifting the impulse response over the input function and calculating, for each location, the sum or integral of the point-wise product of the input function with the impulse response or weighting function. Now, instead of 'shifting' the impulse response over the input function, we can also imagine having many of such impulse responses available at one time and calculating the result at all different locations at once. This is shown in Fig.1.7 for a gaussian impulse response where we know already that a step input function will result in a smoothed or gradual step at the output.

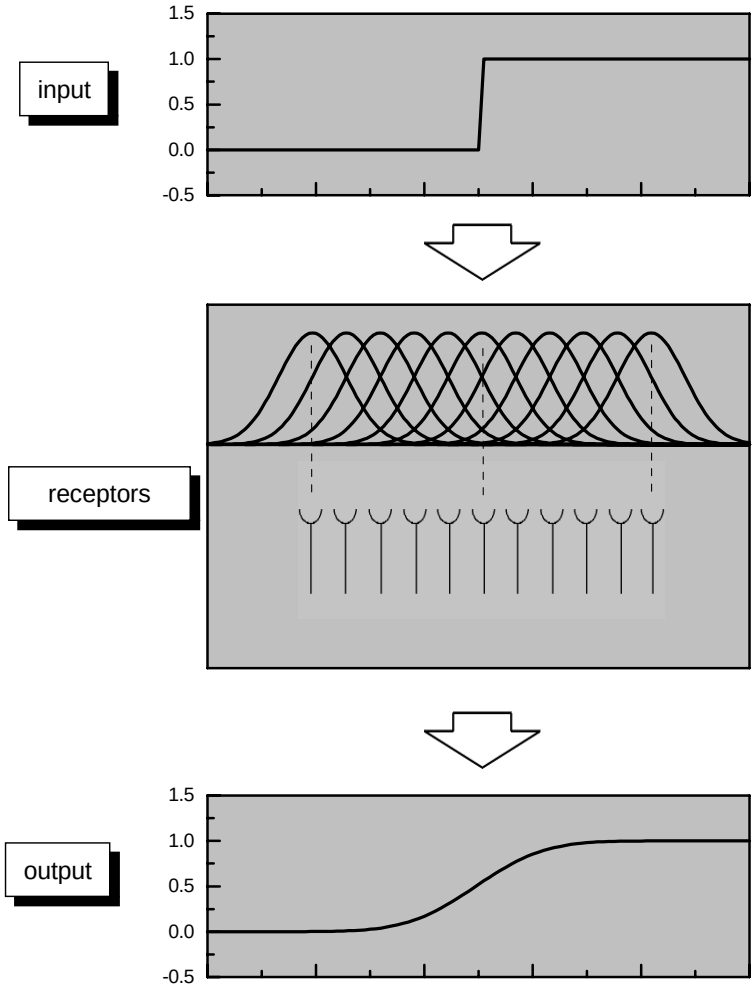


Fig.1.7
Interpretation of impulse responses as receptive fields. An array of gaussian impulse responses are shown at the input of an array of receptors. The input function becomes represented at the output level of the receptors as a smoothed step function just like in the example shown in Fig.1.2.

There are several points to be kept in mind here. First of all, in order to account for the minus sign in the convolution formula between the variable of the input function and the one of the filter function, we have to reverse the impulse response according to the variable when we want to calculate the output function graphically. The second point is that, given we see in some textbook a receptive field of a neuron, we should read carefully about how this measurement was done. Has it really been measured with an impulse function as stimulus, or were some additional manipulations done in order to make the cell respond to the stimulus? The third point concerns the spacing of the neural elements. When higher spatial frequencies are inherent in the input function, the spacing becomes a limiting factor on the resolution by which this function will be represented at the output. We will come to this point later when considering spatial receptive fields of insect photoreceptors where this point has been examined in great detail. The final point concerns the linearity of the convolution. If, and only if, the experimental system that is considered has been shown to be linear within the range of input amplitudes used in the experiments can we expect a good match between its real response to any kind of input function and the response as predicted from the knowledge of the impulse responses of its constituting elements. Nevertheless, we will be always allowed to predict the response from applying the convolution theorem, even if we know that there are some nonlinearities in the system. From a comparison between real and predicted response, we will learn something more about the whereabouts of this nonlinearity in each case.

1.2.4. Discrete sampling and aliasing

There is another difference between a convolution and the transmission of an input function through a receptor array. That difference is the fact that in contrast to the smooth gliding of the weighting function in case of a convolution, there exists only a finite number receptor cells or any other neurons. Therefore, the distance between them is a discrete number and the operation performed by such an array of neural elements is better described by the formula about the discrete convolution using a sum of elements rather than the integral over an infinite number with infinitely small distances between them. The additional problem we have to consider in such a case is the limitation with respect to the smallest wavelength (or the highest frequency) of input function that can be resolved by such an array.

Imagine you were told to switch the light on and off as you like but the fastest you can do is operate the switch every second. What is the highest frequency of light modulation you can produce under these conditions? Obviously, you will change the status of the bulb every second. This results in a cycle length or period of 2 seconds, since the light will stay on for one second and then off for another second. The highest frequency you can produce is 1/2s or 0.5 Hz. It is important to note that this frequency is just half the frequency at which you were able to operate the switch. Or, to talk about time intervals, the period of light modulation is twice as long as the inter-switch interval. The same holds true when the task, instead of producing a modulation of some signal, is to read a signal. Again, the highest frequency you can resolve properly is half the frequency at which you read the signal. This conclusion we arrived at is the famous 'sampling theorem' stated by Shannon in the late 40ies. The frequency limit is often also called the 'Shannon' or 'Nyquist' limit in the literature.

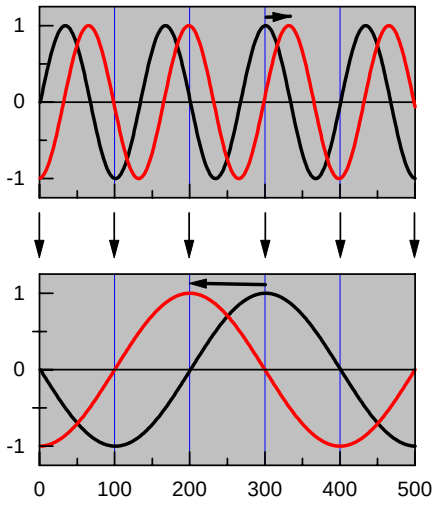


Fig.1.8 *Aliasing effects due to undersampling. A sine wave is illustrated moving in front of a receptor array to the right. The function values of the sine wave are read by an array of receptors spaced at ¾ of the sine wavelength (receptor locations indicated by the array of vertical arrows). The resulting output functions are displayed in the bottom part as continuous lines, although having discrete values only. The aliasing leads to an increased apparent wavelength which moves in the opposite direction at higher speed.*

In order to illustrate what happens when going beyond this limit is illustrated in Fig.1.8. Here, a sine wave is shown in the top part of the figure traveling to the right. Two snapshot of this wave are taken, one, shown in black, at time t_0, and another one in red taken some time Δt later. The horizontal arrow indicates the displacement the sine wave has undergone within Δt. This sine wave is traveling in front of a receptor array having a discrete spacing $\Delta\varphi$ of 3/4 of the wavelength λ of the sine. The locations of the receptors are indicated by the black vertical

arrows underneath the top part of the figure. Obviously, this situation is beyond the Shannon limit, since λ is equal to 4/3 times $\Delta\varphi$ and, thus, smaller than twice the receptor spacing. The resulting output signals of the receptors are shown in the bottom part of the figure, again in black at time t_0 and in red at time Δt later. There are two points to be noticed: i) The apparent wavelength of the input has changed by passing through the array. ii) The direction in which the sine wave has moved, as judged from a nearest neighbor match, is reversed. This example may suffice to illustrate the 'aliasing' effects arising from 'undersampling', i.e. going beyond the Shannon limit.

1.2.5. Receptive fields of neurons

1.2.5.1. The visual system of insects

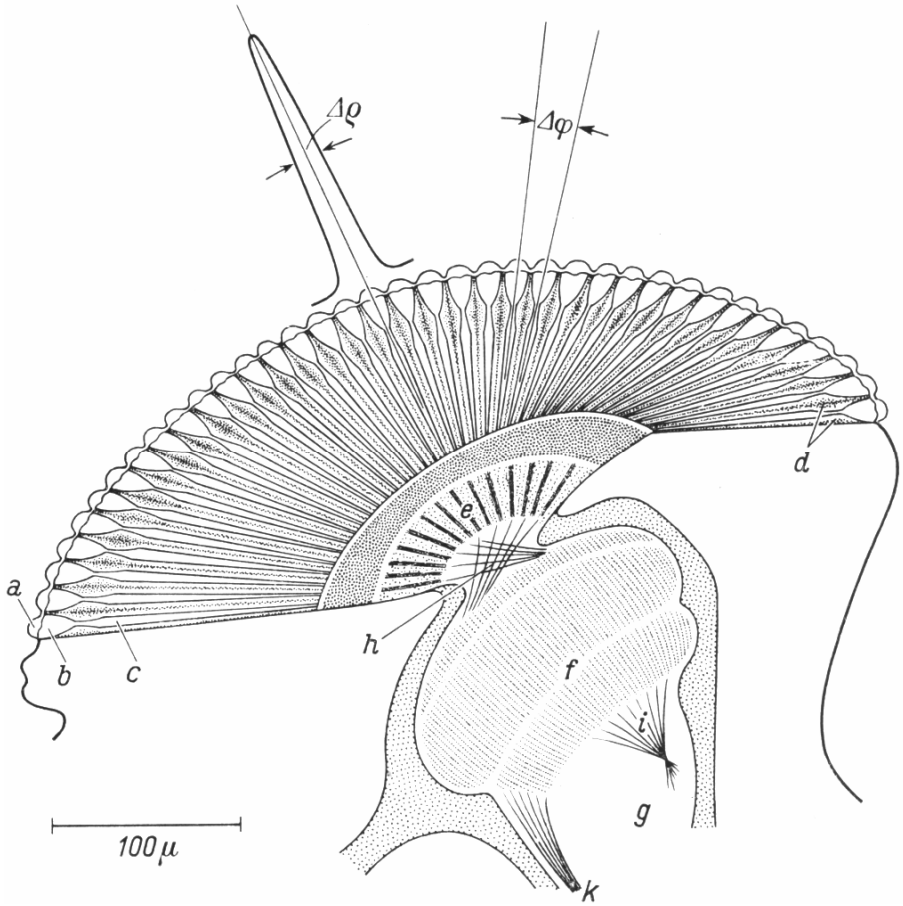
We will now take a look at the receptive fields of nerve cells. Since we have talked about spatial filters so far, the visual system lends itself as the prime sensory system to be examined in this respect. Let us consider the insect eye as our first example. In general, the insect eye consists of a two-dimensional array of facets or ´ommatidia´ each of which houses its own optical apparatus and a set of photoreceptors. In case of dipteran flies, there exist eight photoreceptor cells per ommatidium. In general, insects have poor spatial vision due to their limited number of photoreceptors per visual angle. In case of the blowfly the receptive fields of single photoreceptors have been measured with great accuracy. This has been done with a small spot of light presented in different positions with respect to the photoreceptor that was recorded from. The result of such a measurement is shown in Fig.1.9a. What we see is a bell-shaped curve which represents a one-dimensional slice cut through a radically symmetric 2D-function. This should look familiar to us. Indeed, such a function can well be approximated by a gaussian. The half-width, which roughly corresponds to the σ of a gauss is about 1° in this case. It is also called the receptors 'acceptance angle' in the literature.

We can readily imagine what such a receptive field does with respect to image transformation: It blurs the image. And, it does so quite dramatically. When this acceptance angle ρ is measured as a function of the horizontal position within the eye we see hat this angle even increases to more than 2° (Fig.1.9b). The angle ρ is even larger in case of the small fruitfly Drosophila. Here, the receptor's acceptance angle amounts to more than 4°. There is an interesting trend which is, the smaller the insect, the larger, in general, the photoreceptor acceptance angle. The effect of the spatial filter is called spatial acuity of the visual system.

Before going into the details of how much this affects the input image we need to know the sampling base of the visual apparatus. In case of the insect eye, this angle is the interommatidial angle $\Delta\varphi$ which can be measured optically by shining light from the inside through the optics of the eye. It determines the spatial resolution of the visual system. Interestingly, this interommatidial angle $\Delta\varphi$ matches the acceptance angle ρ both between different areas within one eye and between different species. The ratio of $\Delta\varphi$ and ρ always amounts to approximately 1. Why is this so? The answer relates to the sampling limit given a

certain separation of receptors. Spatial frequencies of a wavelength smaller than 2 $\Delta\varphi$ will lead to aliasing effects. In order to avoid these aliasing effects to affect the visual performance, these high spatial frequencies should be reduced in amplitude. This is done by the low-pass filter effect of the gaussian receptive field. The cut-off frequency will be the larger, the smaller the half width ρ. Therefore, it is a wise idea to adjust these two values so that aliasing is always brought to minimum depending on the exact spatial separation of receptors. Yet, aliasing can still be measured in the laboratory. Fig.1.26 shows a turning response of Drosophila as a function of spatial wavelength of a rotating pattern surrounding the fly. At spatial wavelengths below 9.2°, the response changes its sign: while before, the fly was turning along with the motion of the drum in order to stabilize itself in space (it mistakes the rotation of the drum with self-motion in the opposite direction and tries to counteract it and everybody who ever was in an IMax movie theaters knows this effect), it now turns into the opposite direction. However, when further decreasing the spatial wavelength, the response goes to zero. This is due to the effect of the spatial low-pass filter provided by the large acceptance angle of the photoreceptors. As mentioned above there is a rule of thumb which says the smaller the insect the larger the receptor angle. Since this angle and the receptor spacing match, the rule also says that $\Delta\varphi$ increases with decreasing size of the eye carrier. Since the total amount of visual angle is fixed, this rule also says that smaller insects have less photoreceptors. All this has to do with the photon limited performance of any visual system. When the eye becomes smaller, keeping the number of ommatidia the same would mean a respective reduction of the lens diameter and in concomitant loss of quantum catch. In order to avoid that photoreceptors are completely dominated by photon noise, the diameter of the lenses is not scaled to the head or body size of the insect, but rather the number of ommatidia is reduced. Thus, a blowfly has about 4-5000 ommatidia per eye with an angular separation of about 1° depending on the particular eye region. The receptive field half-width is in about the same range. In contrast, Drosophila, the little fruitfly, has only about 900 ommatidia per eye with a angular separation of about 4° and about the same receptive field half width. In order to appreciate these numbers, one has to know that humans have a spatial resolution in the range of arc-seconds, i.e. less than or better than 0.001°. Imagine putting a postcard with the friendly young man of Fig.1.9b in front of fly at a distance of about 20 cm. This results in about 6 times 4 ommatidia being affected by light coming from the postcard, in the case our fly is a Drosophila, and about 24 times 16 ommatidia, in case it is a blowfly. The way how the

front-end of the visual system of both of these insects transforms the image is shown in Fig.1.9b.



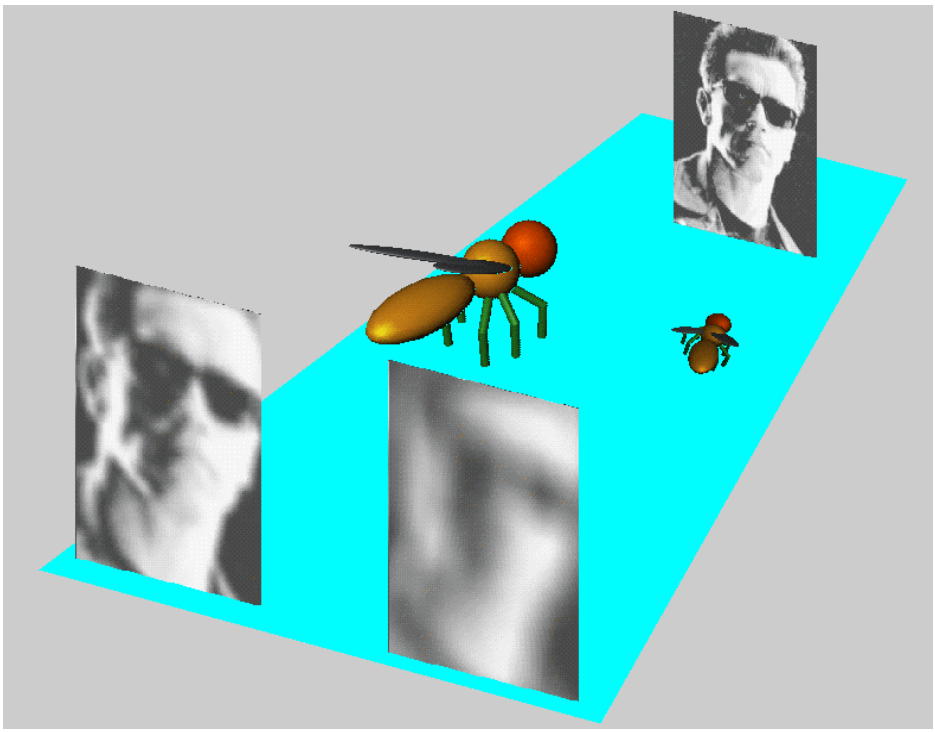


Fig.1.9 a *An array of photoreceptors is located in the facets of the insect eye. Each of these photoreceptors has a pronounced bell-shaped spatial sensitivity profile or receptive field. The half-width ρ of these receptive fields are scaled to the inter-ommatidial angle φ of the eye and, thus, to the sampling base of the visual system.* **b** *In case of the blowfly this angle φ amounts to about1-2°, in case of the fruitfly Drosophila, the angle amounts to only about 5°. With their total of about 900, i.e. 30 times 30, ommatidia per eye, this results in a rather blurred image of the environment.*

1.2.5.2. The visual system of mammals

The Retina

One of the most exciting discoveries was made in the early 50ies by Stephen Kuffler. When recording from cat retinal ganglion cells (which form the output of the retina towards the brain), he noticed that these cells did not change their activity when he illuminated a rather large area but only when a very small spot light was used. Furthermore, he discovered that there exist two kinds of ganglion cells one which increase their activity in response to a bright spot, and one which decrease their activity in response to a bright spot. Further experiments revealed the distinct structure of these ganglion cells: Using a ring of light which just spares the aforementioned spot area had the opposite effect on the cells: the ones which were activated by a bright spot were inhibited by a bright ring, and the ones which were inhibited by a bright spot were activated by a bright ring. Using an ambient illumination, dark spots and rings had exactly the opposite effects than did bright spots and rings. According to these receptive field properties, the ganglion cell populations were named on-center off-surround and off-center on-surround cells.

Equipped with our knowledge about spatial filters, we readily see that these receptive field properties characterize a spatial band-pass filter of the form what we called a DOG filter or 2^{nd} derivative gauss before. The effect of these filters is to neglect uniform light levels and other low frequency signals and to filter out very high spatial frequencies as well. We have also seen that, by this way, they enhance edges. We, thus, can easily predict the kind of images which are transmitted to the brain by the output signals of arrays of such retinal ganglion cells.

LGN

The retinal ganglion cell axons of higher vertebrates either run to a dorsal region of the mid brain called the superior colliculus which corresponds to the tectum opticum in lower vertebrates, or to the lateral geniculate nucleus, called LGN for convenience, in the thalamus. The LGN can be regarded as the entry to the visual cortex, since from there, cells project directly to the primary visual area, the striate cortex. For our purpose, it is legitimate to say, that as a first approximation, the receptive field properties of LGN neurons are just like the ones of their input fibres, i.e. the retinal ganglion cells. The have the same overall properties,

i.e. are isomorph, circular, with the typical center-surround antagonism. Thus, nothing new is happening at this stage of the visual pathway.

Striate Cortex

This is different when receptive field properties of cells of the primary visual cortex are considered. These cells can be grouped in two classes: linear and nonlinear ones. The first class is usually called ‘simple cells’, because most of their response behavior can be predicted from their spatial impulse response, i.e. their response to a small spot of light. The second class does not respond at all to such stimuli: Only when stimulated by moving gratings of a particular spatial frequency, orientation and velocity will they increase their firing rate substantially. Since this type of response behavior implies severe nonlinearities, we will not further consider them in the present context. Simple cells of the striate cortex have, in contrast to their LGN input cells, elongated receptive fields. Like the retinal ganglion cells, their receptive field can be subdivided into antagonistic subregions, but, unlike the former, these are not arranged in a circular manner but rather form elongated ‘cigars’ of neighboring on- and off regions (Fig.1.10). This brings along the following properties: i) They are anisomorph, i.e. their response to a visual pattern will depend heavily on the orientation of the pattern. ii) They have band-pass characteristic, i.e. a grating of a certain spatial frequency excites these cells optimally. Frequencies above or below this optimal frequency leads to smaller responses. Since there exists a variety of simple cells of all different orientations and receptive field sizes, which determines their spatial frequency optimum, the visual input, thus, is being represented according to its local orientation of contours and grain. We can think of this as a multiple map where for each retinal location the prevailing orientation and spatial frequency can be read out by some postsynaptic mechanism.

But how are this maps laid out in the striate cortex? Is it all intermingled, or is there a strict separation between cortical areas preferring different orientations and frequencies? In their pioneering work on the visual cortex, Hubel and Wiesel discovered that simple cells having the same orientation are located above each other in a columnar structure. Cells in different columns have differently oriented receptive fields. Thus, these receptive field features were clearly separated. With the available technique at their time, Hubel and Wiesel could only provide indirect evidence for the spatial arrangement of these different orientations. From their encounters when driving a single electrode through the cortex in an oblique way, they

concluded that there exist parallel bands along which the orientation changes smoothly. To definitely resolve this issue, it took the development of new techniques which allowed to see the spatial layout of activity in the cortex. This technique came up in the late 80ies by A.Grinvald and is based on intrinsic fluorescence changes occuring in the fine blood capillaries when hemoglobin is less and less oxygenated due to the energy consumption of the neighboring nerve cells. Using this optical signal, Bonhoeffer and Grinvald found, that in contrast to the original proposal of Hubel and Wiesel, different orientations are radially grouped about what they called ‘pin-wheel’ centers (Fig.1.11). Going around such a center, each orientation is represented once.

Fig.1.10 *Receptive field properties of retinal ganglion cells and simple cells of the striate cortex (from Reid and Alonso, 1995).*

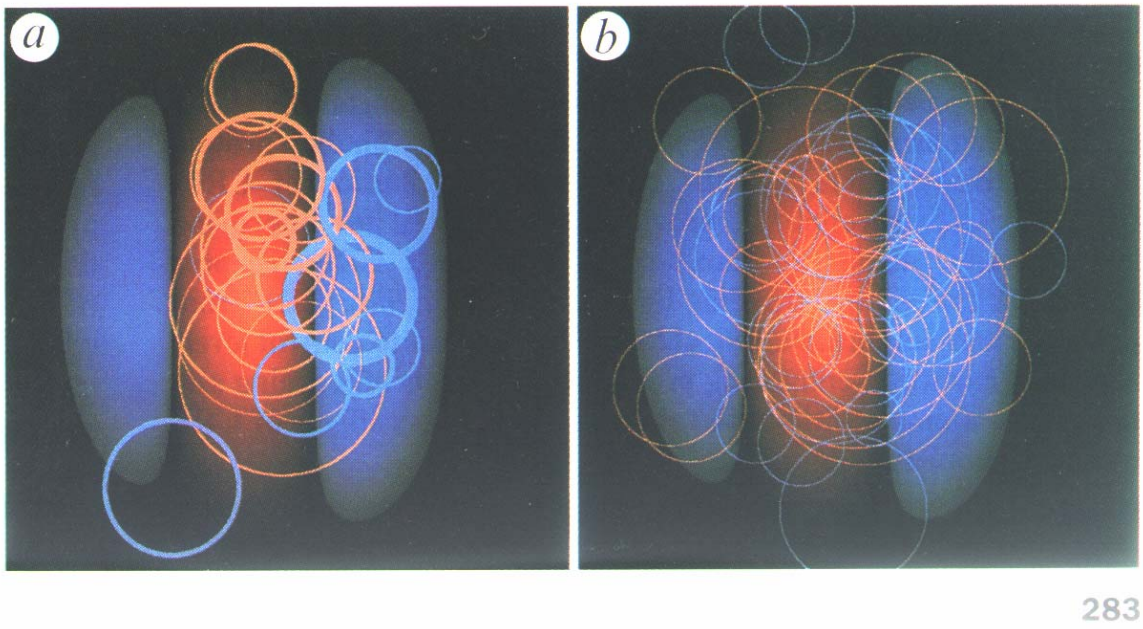
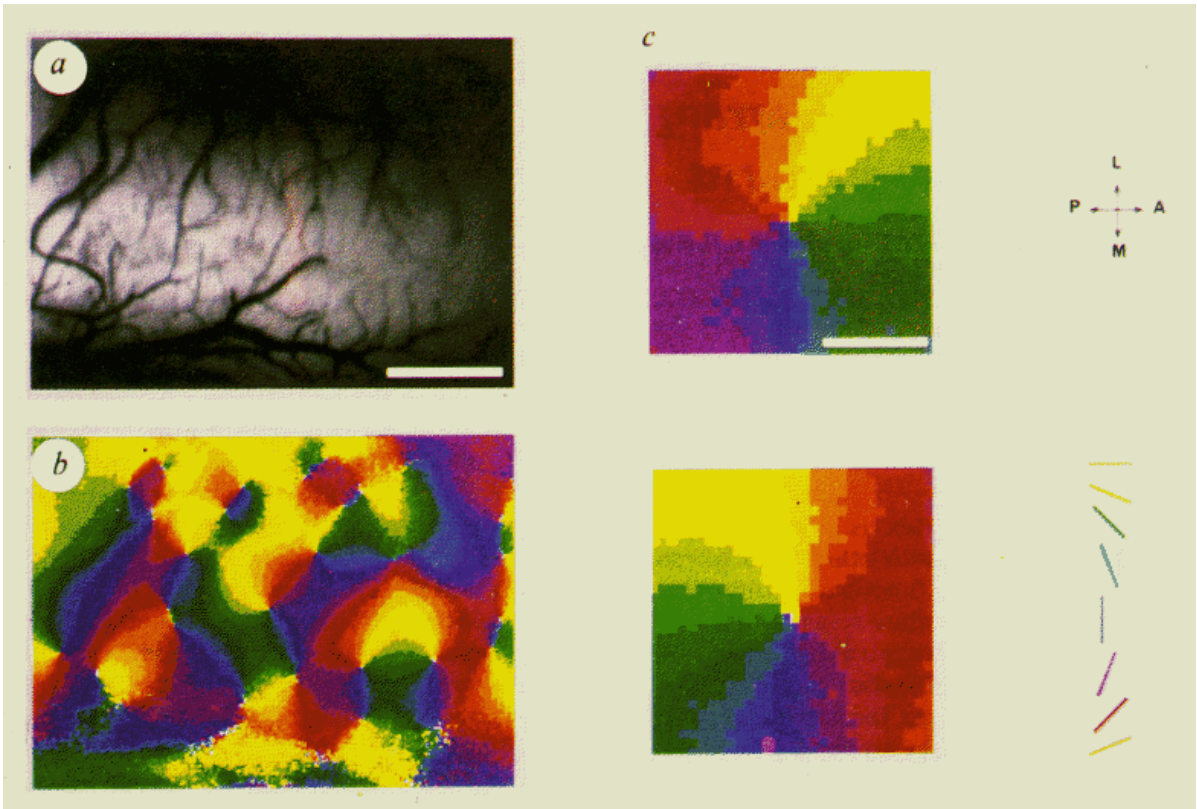


Fig.1.11 *Spatial arrangement of orientation-selective cells in cat striate cortex (from Bonhoeffer and Grinvald, 1988).*



1.2.6. Brightness illusions

What is the color of a TV screen, when it is turned off? It is gray. Nevertheless, when we watch a movie, we would bet that we have seen a lady in a black robe or whatever the reader pays attention to. Being aware of the fact, that a cathode ray tube will only emit but never suck photons, we conclude that our perception of brightness is not very reliable with respect to absolute luminance levels. This is well-known fact and appreciated over the centuries by psychophysicists throughout the world. What most of these brightness illusions have in common is that a gray area neighboring a black area seems brighter than in the neighborhood of a white area. This is called simultaneous luminance contrast and two examples are the displayed on the left hand side in Fig.1.12. In the upper part, the left central area looks brighter than the right area which is embedded in white, and in the lower part, there appear gray squares at the intersection points of the white stripes except for the central one where we focus on. Moving our eyes over the image, the gray squares flicker on and off and give an overall unstable perception of what there really is. Equipped what we learned so far about the action of spatial filters and how receptive fields of peripheral elements of the primate visual system look like, we can readily explain these kind of brightness illusions by the contrast enhancing effect of on-center off surround cells etc. However, although this is certainly true to some extent, one should be careful with that kind of physiological explanations of psychophysical phenomena. First of all, neurons at different processing levels of the visual system have different receptive field properties and it is not at all clear the response of which processing level corresponds to our subjective perception. Second, the fact that all retinal ganglion cells forming the sole input to our visual centers in the brain do not respond at all to an area of homogeneous luminance, should let us take quite the opposite point of view and ask, as the real problem here, why not all areas look exactly the same in their center part no matter what actual gray level they have. The adelson illusion (Fig.1.13) is an example that warns us that not everything about the perception of local brightness can be accounted for by simultaneous luminance contrast. Here, the geometrical, 3D-interpreation of the visual scene influences the perception of local brightness. While in the top figure, the center field of the 2^{nd} row appears darker than the center field of the 4^{th} row, this is not the case in the bottom figure. Yet, neighborhood relations are all the same: Only a pseudo 3D impression is generated by the sheering.

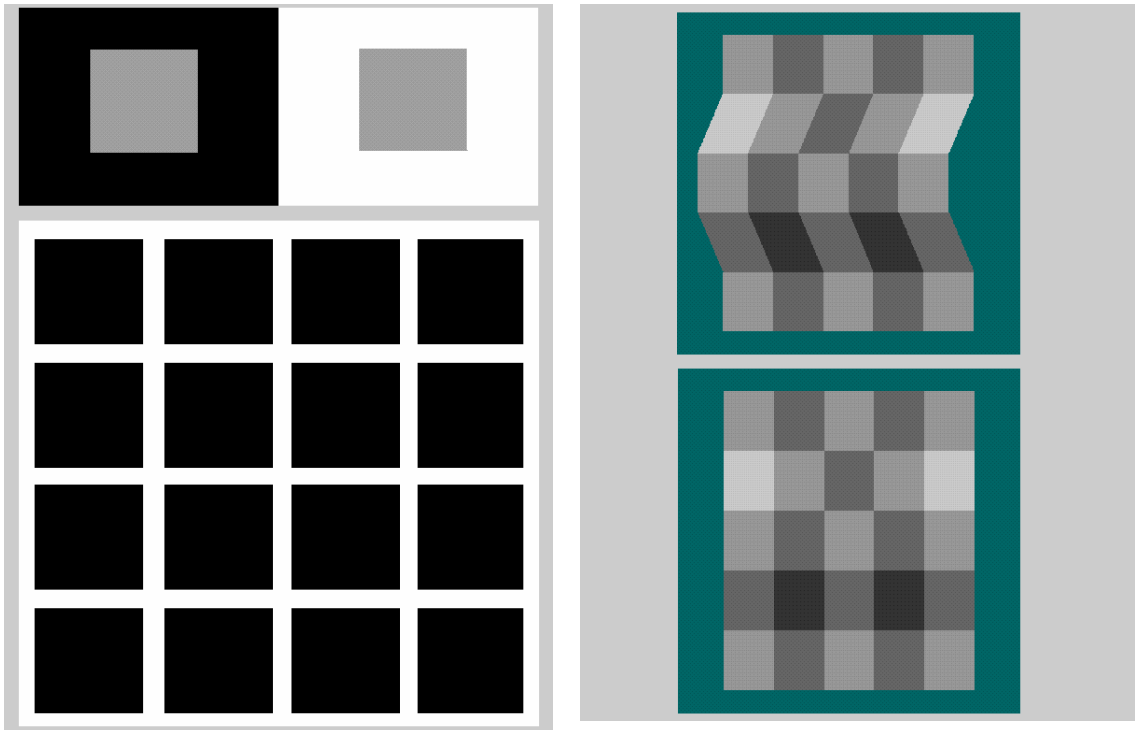


Fig.1.12 *Simultaneous contrast enhancement.*

Fig.1.13 *The adelson illusion teaches that not all brightness illusions can be accounted for on the basis of simultaneous contrast enhancement (from Adelson, 1993).*

This geometrical interpretation, however, seems to have a strong effect on the brightness perception. Keeping in mind that the number of photons coming from a certain location is the product of the strength of illumination and the reflectance properties of the illuminated area, the perceived brightness seems to result from an interpretation of the brain about the reflectance properties only. The different mean luminance values of the 2nd and the 4th row suggests that light is coming somewhere from the top of the figure, and the sheering suggests a wall standing out towards the viewer. Since both these areas have exactly the same gray level, the one standing in the light, therefore, must reflect less light than the one standing in the shadow. Something like this string of arguments is what the brain tries to tell us when we look at this figure. It is an ‘inference’ about the real world properties that the brain is doing here.

1.3. Temporal Filters and Dynamic Response Properties of Nerve cells

1.3.1. Low-pass filters

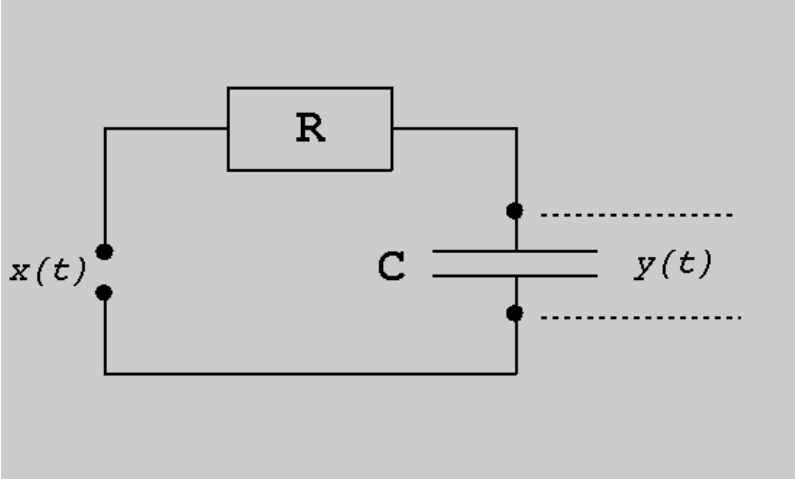


Fig.1.14 *Electrical circuit consisting of a resistor with resistance R and a capacitor with capacitance C. The voltage over the capacitor of such an RC-circuit is equivalent to the output of a temporal 1st-order low-pass filter given an input function to the left.*

Consider the electrical circuit shown in fig.1.14. It consists of a resistor R and a capacitor C. Suppose the input voltage has been down (x(t)=0) for a long time, so that no charge is on the capacitor (y(t)=0). Suppose now, that the input suddenly is stepped to some constant value, e.g. 1. Obviously, the capacitor will charge up, too, but this charging up will take some time before y(t) will finally reach x(t). If the input is regularly switched back to zero and back to 1 with long time intervals in between, y(t) will always reach x(t) before x(t) changes again. So, for low input frequencies, the peak amplitudes of y(t) will be the same as for x(t) and, hence, low frequencies are allowed to pass through the circuit. If, however, x(t) changes its amplitude before y(t) has reached x(t), the ratio of peak amplitude A(y)/A(x) will be smaller than 1. It will actually be the smaller, the faster x(t) changes, i.e. the higher the input frequency. Therefore, high frequencies do not pass through the circuit with full amplitude. Therefore, the circuit is called a low-pass filter.

Up to which input frequency can a low-pass filter follow? This so-called corner frequency is determined by both the value of the resistance R and the capacitance C. The larger R, the less charges pass through the resistor per unit time, and the larger C, the longer it takes to fill the capacitor. The product of R and C is called the time constant τ of the filter: $\tau = RC$. The time constant determines the cut-off frequency of the filter. We will later derive a quantitative relation between these two values.

Let us move, step by step, into a formal description of the circuit. We need to know only a few basic principles of electrical engineering. The current i flowing through a resistor is proportional to the potential difference over the resistor (x(t)-y(t)) divided by R. The same current has to flow across the capacitor, since there is no node between them. The current across a capacitor is given by Cdy(t)/dt, i.e. its proportional to the change of potential over time times the capacitance. So the following equality must hold: Cdy/dt=(x(t)-y(t))/R. Dividing this equation by C and substituting R times C by τ, we arrive at the following differential equation:

$$dy(t)/dt = \frac{1}{\tau}\big(x(t) - y(t)\big); \qquad (1.10)$$

What does it mean? It says that the output increases the faster the stronger the larger the difference is between the actual input and output value, and the smaller the time constant of the filter. It also says that the output does not change anymore, once the output has reached the input level (when (x(t)-y(t))=0, then dy(t)/dt = 0). If this is not obvious, the reader should practice a bit to actually read differential equations, by formulating phrases about formulas.

We now will solve this differential equation by saying that the input x(t) equals zero. The equation then becomes very simple:

$$dy(t)/dt = -\frac{1}{\tau}y(t);$$

This is also called the homogeneous versions of a differential equation. Dividing by y(t) and multiplying by dt yields:

$$dy(t)/y(t) = -\frac{1}{\tau}dt;$$

If this equation becomes integrated on both sides, we get:

$\int_{y(0)}^{y(t)} 1/y(t)dy = -\frac{1}{\tau}\int_{0}^{t} dt;$ Solving the integrals on both sides yields:

$\ln(y(t)) - \ln(y(0)) = -\frac{t}{\tau}$; Taking both sides as the exponent of the natural number e, we arrive at:

$y(t) / y(0) = e^{-t/\tau}$ or

$y(t) = y(0) \cdot e^{-t/\tau}$;

Without much calculations, we could have also said that the solution of the homogeneous equation would be a function y(t) the derivative of which equals the function times -1/τ. Knowing that the e^t has the property of being 'self similar' with respect to its derivative and applying the chain rule of differentiation, we would have readily come up with the same solution. Let us read the equation to see what it means. From whatever level y has at time zero, it will decay exponentially given the input x(t) equals zero from time zero on. If x(t) was 1 for all times before 0, then y(0) will be 1, too. Then, the above equation turns into:

$y(t) = e^{-t/\tau}$

This is the response to a negative step input. The derivative of a negative step should be a negative pulse. To obtain the response to this input, we simply take the derivative:

$dy(t) / dt = -\frac{1}{\tau} e^{-t/\tau}$;

Multiplying the expression on the right hand side by -1 leads to the solution we aimed form, i.e. the impulse response of the electrical circuit shown in Fig.1.14. It is of the form:

$g(t) = \frac{1}{\tau} e^{-t/\tau}$; (1.11)

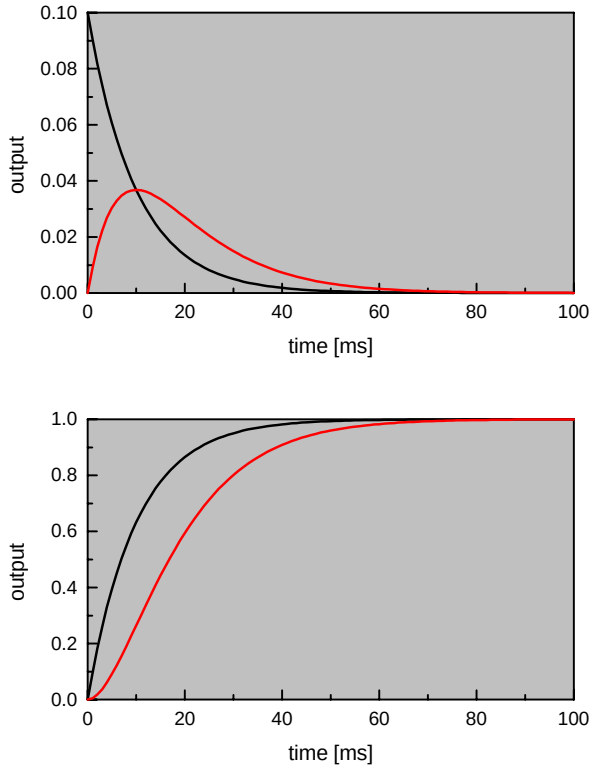


Fig.1.15 *Characteristics of low-pass filters of 1st (black) and 2nd order (red) having a time constant τ of 10 ms. The top graph illustrates the impulse response, the middle one the step response, and the bottom graph shows the amplitude response as a function of frequency.*

Now, we can calculate the step response of such a filter by convolving a function which equals 1 for positive times and zero for negative ones. We thus only have to integrate from zero to t:

$$y(t) = \int_0^t 1 \cdot \frac{1}{\tau} e^{-t'/\tau} dt' = \frac{1}{\tau} \left| e^{-t'/\tau} \cdot (-\tau) \right|_0^t = \frac{1}{\tau} ((-\tau) \cdot e^{-t/\tau} + \tau) = 1 - e^{-t/\tau}; \qquad (1.12)$$

We see that this is indeed compatible with our response to a negative step: instead of decaying to zero when x(t) turns 0, y(t) goes up to 1 when x(t) turns to 1.

A feed-back free sequence of two low-pass filters of 1st-order is called 2nd order low-pass filter. To calculate its impulse response, we take the impulse response of the 1st-order filter and pass it through itself again:

$$x(t) = \frac{1}{\tau} e^{-t/\tau};$$

$$y(t) = \int_0^t \frac{1}{\tau} \cdot e^{-(t-t')/\tau} \cdot \frac{1}{\tau} e^{-t'/\tau} dt' = \frac{1}{\tau^2} \int_0^t e^{-t/\tau} dt' = \frac{1}{\tau^2} \left| e^{-t/\tau} \cdot t' \right|_0^t = \frac{t}{\tau^2} e^{-t/\tau}; \qquad (1.13)$$

This function is shown in the top panel of Fig.1.15 as the red line. In contrast to the 1st order impulse response, which has an infinite steep rise at time zero, the 2nd order impulse response goes up more slowly and, therefore, is often better adapted to describe some physiological behavior in biology.

To calculate the step response of a 2nd order low-pass filter, we can use the step response of one 1st order filter as input to the second one:

$$x(t) = 1 - e^{-t/\tau};$$

$$y(t) = \int_0^t (1 - e^{-(t-t')/\tau}) \cdot \frac{1}{\tau} e^{-t'/\tau} dt' = \int_0^t \frac{1}{\tau} e^{-t'/\tau} dt' - \int_0^t e^{-(t-t')/\tau} \frac{1}{\tau} e^{-t'/\tau} dt' =$$

$$= \frac{1}{\tau} \left| e^{-t'/\tau} \cdot (-\tau) \right|_0^t - \frac{1}{\tau} \int_0^t e^{(-t-t'+t')/\tau} dt' = \frac{1}{\tau} (-\tau \cdot e^{-t/\tau} + \tau) - \frac{1}{\tau} \left| e^{-t/\tau} \cdot t' \right|_0^t =$$

$$= 1 - e^{-t/\tau} - \frac{t}{\tau} e^{-t/\tau} = 1 - e^{-t/\tau} (\frac{t}{\tau} + 1); \qquad (1.14)$$

We could as well used the step function directly and convolved it with the weighting function of the 2nd order low-pass filter directly. The reader may use this way to confirm the result we just obtained. The step-response of the 2nd order low-pass filter is shown in Fig.1.15, too.

Now, let us calculate the response of a 1st order low-pass filter to a sinusoidal input function x(t) = sin(ωt). For the sake of simplicity, let us exchange roles here and be the sine the weighting function and the exponential the input function. As we shall show in general, that makes no difference at all for the output (law of commutativity). The integral we have to solve then is:

$$y(t) = \int_0^t \frac{1}{\tau} \cdot e^{-(t-t')/\tau} \cdot \sin(\omega t') dt';$$

Here, we need to know the following formula:

$$\int e^{cx} \cdot \sin(bx)dx = \frac{e^{cx}}{c^2 + b^2}(c\sin(bx) - b\cos(bx));$$

Applying this formula to our problem, we arrive at the following expression:

$$y(t) = \frac{1}{\tau}\left|\frac{e^{-(t-t')/\tau}}{\frac{1}{\tau^2} + \omega^2}\left(\frac{1}{\tau}\sin(\omega t') - \omega\cos(\omega t')\right)\right|_0^t =$$

$$\frac{1}{\tau} \cdot \frac{1}{\frac{1}{\tau^2} + \omega^2}\left[e^{-(t-t)/\tau} \cdot \frac{1}{\tau}\sin(\omega t) - e^{-(t-t)/\tau} \cdot \omega\cos(\omega t) - e^{-(t-0)/\tau} \cdot \frac{1}{\tau}\sin(0) + e^{-(t-0)/\tau} \cdot \omega\cos(0)\right] =$$

$$\frac{1}{\frac{1}{\tau} + \tau\omega^2}\left[\frac{1}{\tau}\sin(\omega t) - \omega\cos(\omega t) + \omega e^{-t/\tau}\right] = \frac{1}{1 + \tau^2\omega^2}\left[\sin(\omega t) - \tau\omega\cos(\omega t) + \tau\omega e^{-t/\tau}\right];$$

(1.15)

Here, the third term will vanish over time. After that, in the so-called steady-state, the response of the low-pass consists of the superposition of a sine and a cosine function with the same frequency but different amplitude. Expressions of the form $f(t) = a\sin(\omega t) + b\cos(\omega t)$ may be also expressed as $f(t) = A\sin(\omega t + \varphi)$ where the coefficients A and φ are calculated according to the following rules:

$a\sin(\omega t) + b\cos(\omega t) = A\sin(\omega t + \varphi) = A\sin(\omega t)\cos(\varphi) + A\cos(\omega t)\sin(\varphi);$

Therefore, we know that

$a = A\cos(\varphi);$ (1.16)

and

$b = A\sin(\varphi);$ (1.17)

Squaring and adding (1.16) and (1.17) results in:

$a^2 + b^2 = A^2\cos^2(\varphi) + A^2\sin^2(\varphi) = A^2;$ because $\cos^2(\varphi) + \sin^2(\varphi) = 1;$

So, we can assign an amplitude A to our new periodic function.

$A = \sqrt{a^2 + b^2};$ (1.18)

Calculating the phase results in:

$$\frac{b}{a} = \frac{A\sin(\varphi)}{A\cos(\varphi)} = \tan(\varphi);\ \text{or}\ \varphi = arctg(\frac{b}{a}); \tag{1.19}$$

We thus obtain the following rule:

$f(t) = A\sin(\omega t + \varphi)$ with $A = \sqrt{a^2 + b^2}$ and $\varphi = arctg(b/a)$;

Applying this to the result we obtained above, we arrive at:

$$A(\omega) = \sqrt{\frac{1}{\left(1+\tau^2\omega^2\right)^2} + \frac{\tau^2\omega^2}{\left(1+\tau^2\omega^2\right)^2}} = \frac{\sqrt{1+\tau^2\omega^2}}{1+\tau^2\omega^2} = \frac{1}{\sqrt{1+\tau^2\omega^2}}; \tag{1.20}$$

$$\varphi(\omega) = -arctg(\tau\omega); \tag{1.21}$$

This is the amplitude and phase response of a 1st order low-pass filter. In a similar way as just shown, we can also calculate the respective responses of a 2nd order low-pass filter. The results are:

$$A(\omega) = \frac{1}{1+\tau^2\omega^2}\ \text{and}\ \varphi(\omega) = -2arctg(\tau\omega); \tag{1.22}$$

Again, these responses are shown as a function of the frequency in Fig.1.15. As we will show in the chapter about Fourier transform, amplitude responses of feed-back free sequential filters multiply while phase responses add. Therefore, we can easily write down the respective formulas for a low-pass filter of nth order:

$$A(\omega) = \frac{1}{\sqrt{(1+\tau^2\omega^2)^n}}\ \text{and}\ \varphi(\omega) = -n \cdot arctg(\tau\omega); \tag{1.23}$$

(Amplitude and phase response of an nth order low-pass filter).

Amplitude spectra of filters are usually displayed in so-called Bode-diagrams (Fig.1.16). There, the y-axis represents the output amplitude relative to the input amplitude in a logarithmic scale. The x-axis in a Bode-diagram represents the frequency, again in a

logarithmic scale. In technical language, the output y of a filter is often indicated in dezibel, a logarithmic unit where a ratio of 10 equals 20 dB, or in other words:

$$y[dB] = 20\log_{10}\frac{A_{out}}{A_{in}};$$

Often, a filter is characterized by its corner frequency. In a Bode-diagram, this is the frequency where the extension of the straight line on the filter's falling edge crosses the zero line. In case of a 1st order low-pass filter, the input is damped to $1/\sqrt{2}$ of the input amplitude, or by -3 dB, respectively. This is because

$$y[dB] = 20\log_{10}\frac{1}{\sqrt{2}} = 20\log_{10} 2^{-1/2} = 20\log_2 2^{-1/2}\cdot\log_{10} 2 = -10\log_{10} 2 = -3dB;$$

For large frequencies, the amplitude spectrum can be approximated by

$$\lim_{\omega\to\infty} A(\omega) = \frac{1}{(\tau\omega)^n};$$

In a double-logarithmic plot such as Bode-diagram, the spectrum therefore follows the line

$$\log(A(\omega)) = \log(\frac{1}{(\tau\omega)^n}) = -n\log(\tau\omega);$$

Thus, the slope of this line indicates the order of the filter. Another way to see that at the corner frequency of a 1st order low-pass filter, the output is damped to $1/\sqrt{2}$, we can use the approximation of the amplitude spectrum for large frequencies:

$$\lim_{\omega\to\infty} A(\omega) = \frac{1}{\tau\omega};$$

This values will cross the upper line (where A equals one, i.e. no damping at all) for $\omega=1/\tau$. At this frequency, the real (and not approximated) amplitude equals:

$$A(\omega = 1/\tau) = \frac{1}{\sqrt{1+\tau^2\omega^2}} = \frac{1}{\sqrt{1+\tau^2(1/\tau)^2}} = \frac{1}{\sqrt{2}};$$

Usually, one indicates the damping action of a filter by the dezibel per octave, it is by doubling the frequency.

$$20\log\frac{A(2\omega)}{A(\omega)} = 20\log\frac{1/(2\tau\omega)^n}{1/(\tau\omega)^n} = -20n\log 2 = -6ndB/octave;$$

One can also indicate the damping per decade:

$$20\log\frac{A(10\omega)}{A(\omega)} = 20\log\frac{1/(10\tau\omega)^n}{1/(\tau\omega)^n} = -20n\log 10 = -20ndB/decade;$$

The amplitude and phase spectra of low-pass filters of various orders are displayed in Fig. 1.16 to illustrate their filter properties.

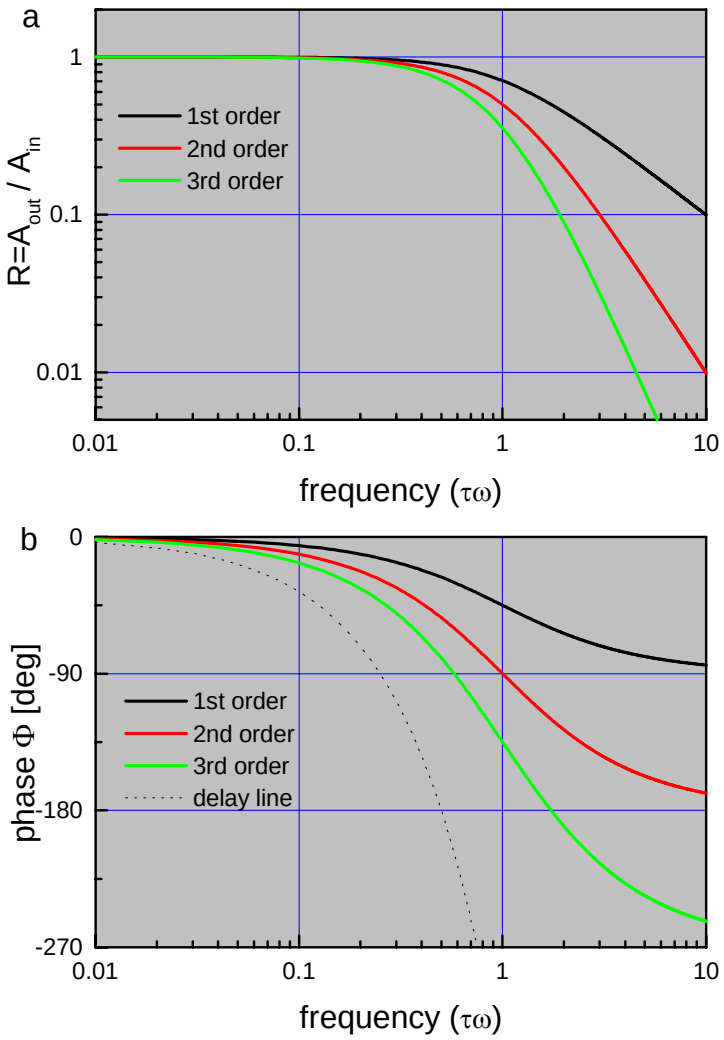


Fig.1.16 **a** *Double-logarithmic ('Bode') diagram of the amplitude spectra of low-pass filters of various orders. Using this logarithmic scaling, the slope of the filter curve turns out to be a straight line for higher frequencies. The higher the order of the filter, the steeper the slope. The x-axis has been normalized by τω. So when this value equals one, the circular frequency is just 1/τ.* **b** *Phase spectra of the same filters plus the phase spectrum of a delay line. The y-axis is scaled in linear units of phase angle in degree.*

We have not talked about a special type of filter which unlike low-pass filters does not influence the amplitude of the signal in a frequency-dependent way, but which shares the property to shift the signals back in time: The so-called delay time. Its impulse response is $\delta(t-\varepsilon)$ with ε being the time by which the filter delays the incoming signal. Its amplitude spectrum is simply a constant: $A(\omega)=1$, for all frequencies. What about its phase spectrum? We know that the filter introduces a constant delay ε. We now have to express this delay in terms of the phase, i.e. as a fraction of the cycle: $\Phi(\omega)=-\varepsilon/T=-\varepsilon\cdot\omega$. It is a straight line through the origin, i.e. starting at zero phase for zero frequency, with the negative slope -ε.

1.3.2. High pass and band-pass filters

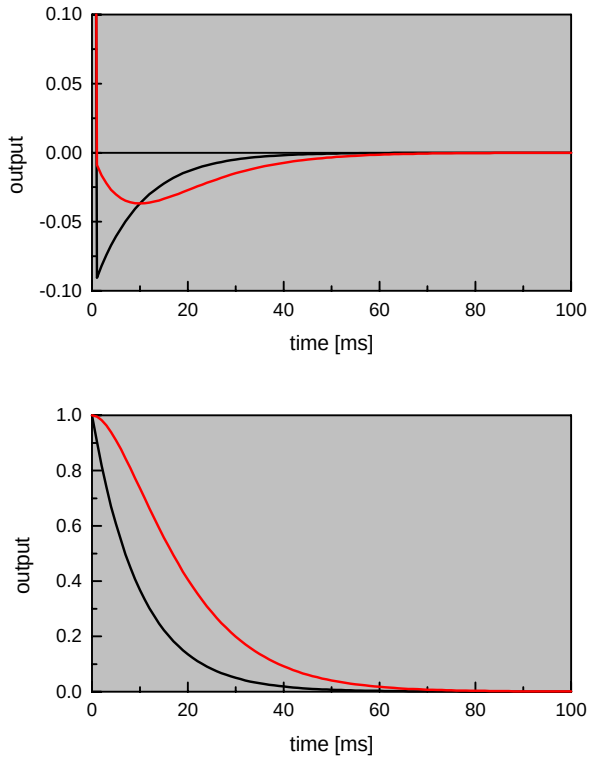


Fig.1.17 *Characteristics of high-pass filters of 1st (black) and 2nd order (red) having a time constant τ of 10 ms. The top graph illustrates the impulse response, the middle one the step response, and the bottom graph shows the amplitude response as a function of frequency.*

From looking at the circuit shown in Fig.1.14., we can instantly derive the impulse and step responses of a 1st order high-pass filter once we know the response of the low-pass filter. The reason why this is easy is that both voltages always have to add to equal the input voltage x(t). In case of an impulse response, x(t) = δ(t). Thus, the impulse response becomes:

$$y(t) = \delta(t) - \frac{1}{\tau} e^{-t/\tau} \qquad \text{(Impulse response of 1st order high-pass filter)} \qquad (1.24)$$

$$y(t) = \delta(t) - 1 + e^{-t/\tau}\left(\frac{t}{\tau} + 1\right) \qquad \text{(Impulse response of 2nd order high-pass filter)} \qquad (1.25)$$

In case of a step input, x(t) = 1 for all t>0. Thus, the step response y(t) of a 1st order high-pass filter is:

$$y(t) = 1 - (1 - e^{-t/\tau}) = e^{-t/\tau} \qquad \text{(Step response of 1st-order high-pass filter)} \qquad (1.26)$$

The step response of a 2nd order high-pass filter becomes, in analogy:

$$y(t) = 1 - (1 - e^{-t/\tau}(\frac{t}{\tau} + 1)) = e^{-t/\tau}(\frac{t}{\tau} + 1); \quad \text{(Step response of 2nd order high-pass filter)} \tag{1.27}$$

The amplitude and phase response of a 1st order high pass filter can be calculated in analogy to our derivation in case of a low-pass filter:

$$A(\omega) = \frac{\tau\omega}{\sqrt{1+\tau^2\omega^2}} \text{ and } \varphi(\omega) = arctg(\frac{1}{\tau\omega}); \tag{1.28}$$

Amplitude and phase response of an nth order high pass filter:

$$A(\omega) = \frac{(\tau\omega)^n}{\sqrt{(1+\tau^2\omega^2)^n}} \text{ and } \varphi(\omega) = n \cdot arctg(\frac{1}{\tau\omega}); \tag{1.29}$$

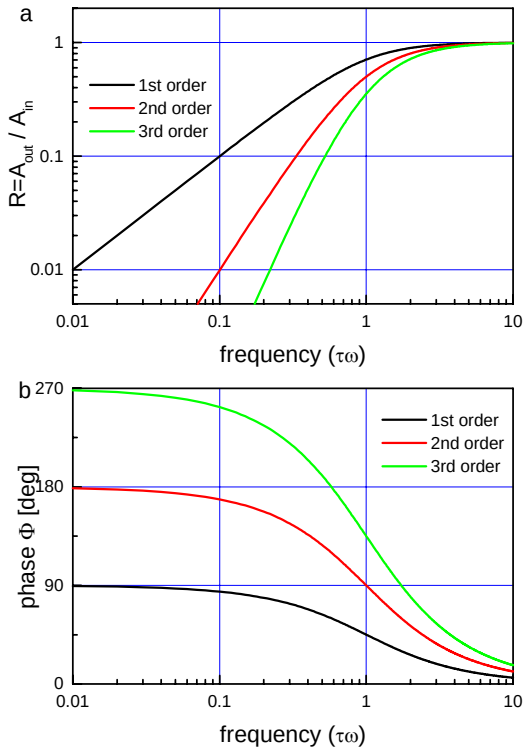


Fig.1.18 **a** *Double-logarithmic ('Bode') diagram of the amplitude spectra of high-pass filters of various orders. Using this logarithimc scaling, the slope of the filter curve turns out to be a straight line for higher frequencies. The higher the order of the filter, the steeper the slope. The x-axis has been normalized by* $\tau\omega$*. So when this value equals one, the circular frequency is just* $1/\tau$*.* **b** *Phase spectra of the same filters. The y-axis is scaled in linear units of phase angle in degree.*

From all we know so far about low- and high-pass filters, we can construct any series of low- and high-pass filters to create what is called a band-pass filter. Since we already know that amplitude responses of individual filters multiply and phase responses add when the filters are arranged in series, we can easily use the above formulas to calculate the respective responses of any kind of fancy band-pass. Since we meanwhile have also learned to use the convolution theorem to calculate impulse and step responses, it is left to the reader to get some practice by calculating the responses of interest of his favorite kind of band-pass filter. We will, therefore, only derive the formulas of combinations of 1st order low- and high-pass filters with identical time constants:

Impulse response of 1st order band-pass filter:

$$y(t) = \int_0^t (\delta(t-t') - \frac{1}{\tau} e^{-(t-t')/\tau}) \cdot \frac{1}{\tau} e^{-t'/\tau} dt' = \int_0^t \delta(t-t') \frac{1}{\tau} e^{-t'/\tau} dt' - \int_0^t \frac{1}{\tau} e^{-(t-t')/\tau} \cdot \frac{1}{\tau} e^{-t'/\tau} dt' =$$

$$= \frac{1}{\tau} e^{-t/\tau} - \frac{1}{\tau^2} \left| e^{-t/\tau} \right|_0^t = \frac{1}{\tau} e^{-t/\tau} - \frac{1}{\tau^2} t e^{-t/\tau} = \frac{1}{\tau} e^{-t/\tau} \cdot (1 - t/\tau); \qquad (1.30)$$

When deriving this formula we made use of the fact that a Dirac pulse convolved with the impulse response results in the impulse response because this is how it is defined. When looking at the impulse response in Fig.1.17, we note that the function has a decay for the first few milliseconds and then becomes negative before smoothly approximating zero with increasing time. We again try to get some intuition for such a filter and ask why this filter is a band-pass? If we imagine a high frequency sine wave as input, the filter will average over many periods and, thus, will lead to a zero result at the output. The same is true for very low frequency input where basically a constant value becomes multiplied with the positive as well with the negative part of the impulse response which then will cancel to zero. This leaves only an input with a certain frequency to pass through with reasonable amplitude, and this is such a frequency where the period matches the filters impulse response as much as possible.

We now calculate the step response of 1st order band-pass filter by feeding the step response of a 1st order high-pass filter into a low-pass filter:

$$y(t) = \int_0^t e^{-(t-t')/\tau} \cdot \frac{1}{\tau} e^{-t'/\tau} dt' = \frac{1}{\tau} \int_0^t e^{-(t+t'-t')/\tau} dt' = \frac{1}{\tau} \int_0^t e^{-t/\tau} dt' = \frac{1}{\tau} e^{-t/\tau} \cdot |t'|_0^t = \frac{t}{\tau} e^{-t/\tau}; \quad (1.31)$$

The step response is again shown in Fig.1.17. It looks very much like the impulse response of a low-pass filter (compare to Fig.1.15). The reason for that is that a high pass filter differentiates the input function. We may note that the impulse is the first derivative of the step function. Since the operation of differentiation can be applied anywhere, either to the input, to the weighting function or to the output, we conclude that it does not matter whether we apply a low-pass filter to a pulse, or whether we feed a step function through a series of high and low pass filters, because in the latter case, the first operation turns a step into a pulse function.

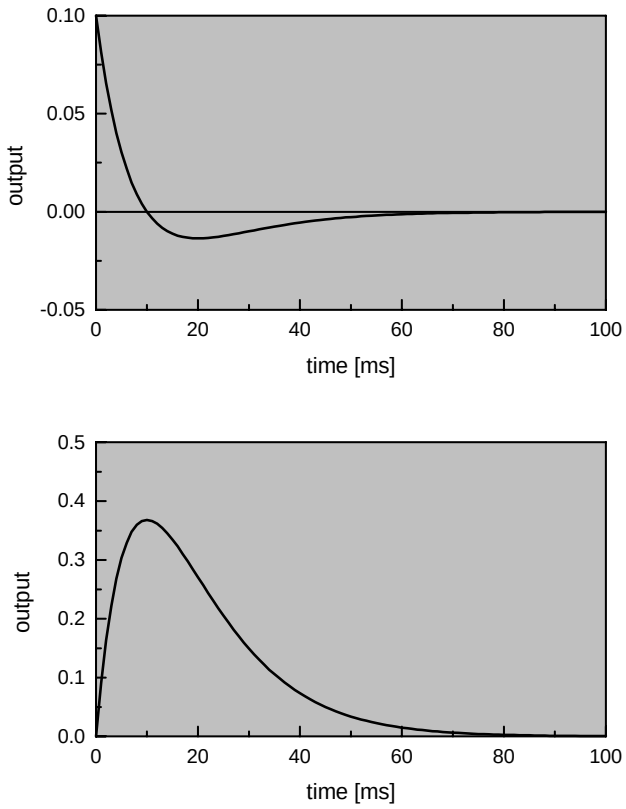


Fig.1.19 *Characteristics of 1st order band-pass filters having a time constant τ of 10 ms. The top graph illustrates the impulse response, the middle one the step response, and the bottom graph shows the amplitude response as a function of frequency.*

Without deriving the following formulas at this point, we simply state that amplitude responses, when written as a function of frequency, of filters in series simply multiply. Phase responses add, respectively. Thus, knowing the amplitude and phase responses of low- and high-pass filters, we can write the amplitude and phase response of a 1st order band-pass filter:

$$A(\omega) = \frac{\tau\omega}{1+\tau^2\omega^2} \text{ and } \varphi(\omega) = arctg(\frac{1}{\tau\omega}) - arctg(\tau\omega); \qquad (1.32)$$

The latter can be simplified using the following formula:

$$arctg(a) - arctg(b) = arctg\left(\frac{a-b}{1+a\cdot b}\right);$$

The phase response of the 1st order band-pass filter then turns into:

$$\varphi(\omega) = arctg(\frac{1}{\tau\omega}) - arctg(\tau\omega) = arctg\left(\frac{\frac{1}{\tau\omega}-\tau\omega}{1+\frac{1}{\tau\omega}\tau\omega}\right) = arctg\left(\frac{1}{2}\left(\frac{1}{\tau\omega}-\tau\omega\right)\right);$$

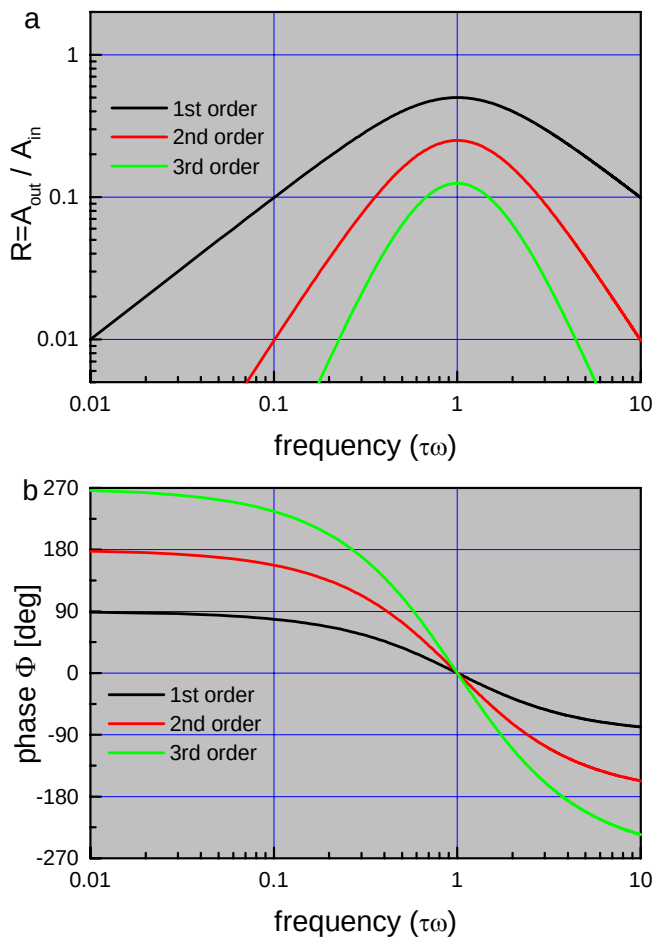


Fig.1.20 **a** *Double-logarithmic ('Bode') diagram of the amplitude spectra of band-pass filters of various orders. Using this logarithimc scaling, the slope of the filter curve turns out to be a straight line for higher frequencies. The higher the order of the filter, the steeper the slope. The x-axis has been normalized by* $\tau\omega$*. So when this value equals one, the circular frequency is just* $1/\tau$*.* **b** *Phase spectra of the same filters. The y-axis is scaled in linear units of phase angle in degree.*

Let us examine the amplitude and phase spectra in some detail. Where is the amplitude response maximum? This is when dA(ω)/dω equals zero.

$$dA(\omega)/d\omega = \frac{d}{d\omega}\left(\frac{\tau\omega}{1+\tau^2\omega^2}\right) = \frac{\tau\cdot\left(1+\tau^2\omega^2\right)-\tau\omega\cdot\left(\tau^2 2\omega\right)}{\left(1+\tau^2\omega^2\right)^2} = \frac{\tau+\tau^3\omega^2-2\tau^3\omega^2}{\left(1+\tau^2\omega^2\right)^2} = \frac{\tau-\tau^3\omega^2}{\left(1+\tau^2\omega^2\right)^2};$$

This expression becomes zero, when the nominator is zero, and the denominator is not. We thus obtain:

$$\tau-\tau^3\omega^2=0; \qquad \tau^3\omega^2=\tau; \qquad \omega^2=\frac{1}{\tau^2}; \qquad \omega=\frac{1}{\tau}; \qquad (1.33)$$

We make sure that the denominator is different from zero for this value of ω. Thus, when the circular frequency is 1 over the time constant, or the product of both equals 1, the amplitude response of the band pass is maximum. Inserting this value into the phase response reveals, that at this so-called center frequency of the filter, the phase shift is zero.

When a series of band-pass filters with different center frequencies are operated in parallel, one obtains a filter bank. The output of the different channels then represents the amplitude by which the different frequencies are present in an incoming signal. This is essentially an on-line Fourier Analyzer, and we will later see where this name comes from.

It might also be interesting to know that also complementary filters to band-pass filters can be constructed. These so-called ‘band reject’ filters allow all frequencies to go through, but one. Their amplitude response can be thought of as 1 minus the one of a band-pass filter. Such filters are often used if a signal is corrupted by some noise source which has only a limited frequency. Adjusting the center frequency of the band-reject filter to the noise frequency cuts off the noise efficiently, but one also has to keep in mind, that signals which occupy the same frequency range are cut out as well.

1.3.3. Dynamic Response Properties of Neurons

The dynamic response properties of many neurons can be well approximated by linear filters. In the following we will review some examples from the fly visual system, taken from two stages: the first example deals with the properties of units of the first optical lobe, the so-called lamina. We then turn to movement sensitivity, a property which can be found in many neurons in the visual system and which has been also investigated at the behavioral level.

1.3.3.1. Transient and Sustained Units

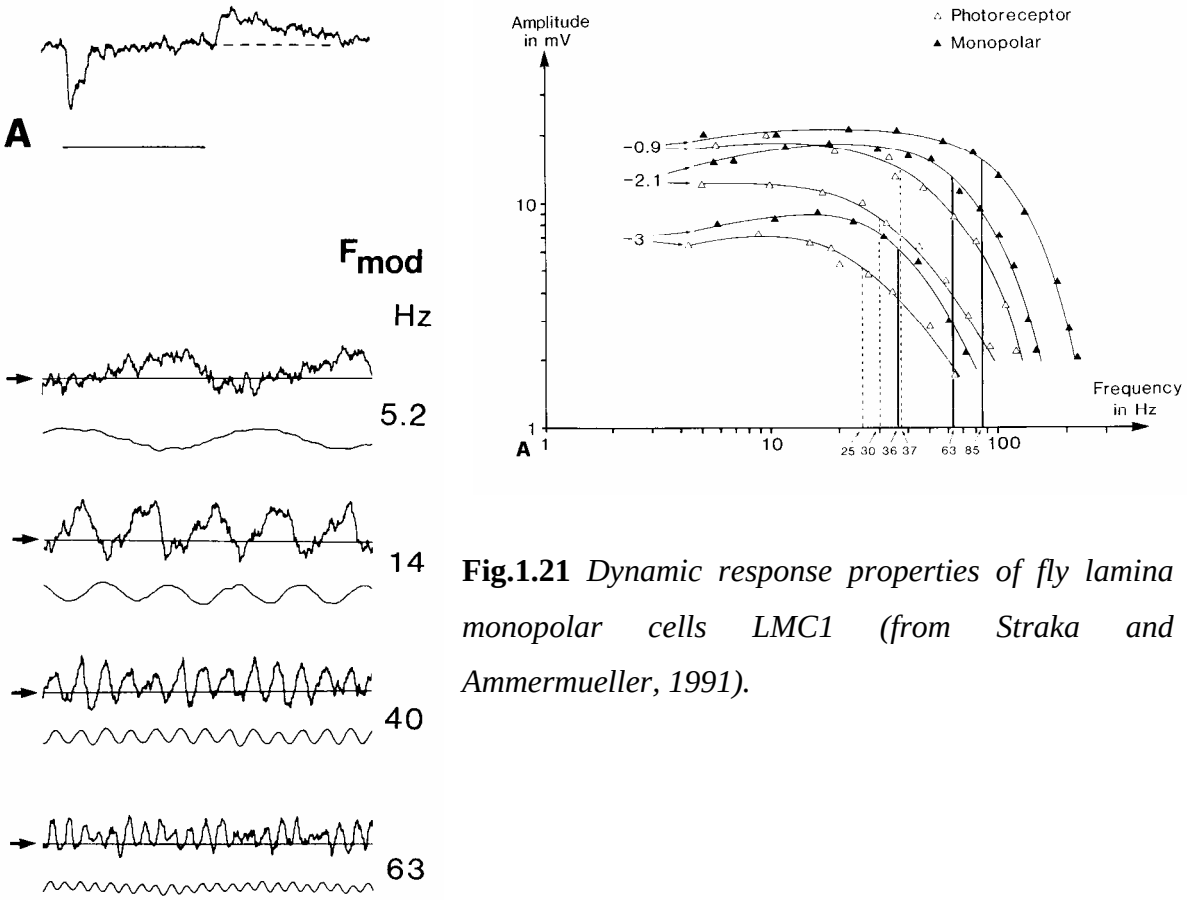


Fig.1.21 *Dynamic response properties of fly lamina monopolar cells LMC1 (from Straka and Ammermueller, 1991).*

In the fly lamina, each cartridge comprises 5 different cells receiving either direct or indirect input from the photoreceptors R1-6. While the first three of them can be examined by intracellular recordings, the latter two can, due to their small size, only be recorded extracellularly. Since their anatomical identity is not proven yet, they are referred to as ‘units’. Fig.1.21 Shows the response properties of LMC1 cells. They are graded response neurons and hyperpolarize transiently when the light level is increased, and depolarize slightly when the light is dimmed (Fig.1.21 A). Such a behavior is similar to the step-response of a high or band-pass filter (compare with Figs.1.17 and 1.19). The right panel shows the

amplitude spectrum of these cells at three different mean light intensities measured as peak-to-peak amplitudes in response to sinusoidal light intensity modulations of various frequency. Although the spectra were not measured up to low enough frequencies, a band-pass characteristic is revealed. The real filter properties of these neurons is uncovered when their response spectra are compared with the ones of their input elements, i.e. photoreceptors. LMCs can be seen to boost the responses of the photoreceptors in the frequency range from 10 to about 100 Hz significantly.

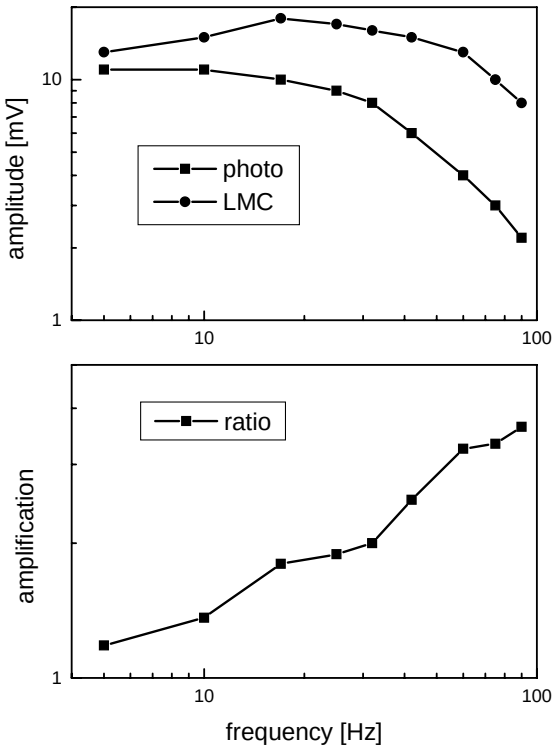


Fig.1.22 *Amplitude spectra of fly photoreceptors and LMCs (top) and the ratio of LMC and photoreceptor responses (bottom). The transfer characteristic of the signal processing from receptor to postsynaptic LMC reveals a strong high-pass component, i.e. high frequency signals are boosted as compared to low frequency ones. Data are redrawn from Straka and Ammermueller (1991) and further evaluated.*

The other response types of the fly lamina are called ‘on-off’ and ‘sustaining’ units, respectively. Their characteristics are shown in Fig.1.23. The top panel illustrates their spike frequency in response to an on-set and off-set of light. While the on-off units can be seen to give positive response to both events, the sustaining only respond transiently to the onset. Their response decays to resting level after the light is turned off again. As compared to LMCs 1-3, their high frequency cut-off is shifted substantially to lower frequencies. They, thus, operate much more in the low frequency range and are speculated to represent the input to the motion detection system (see next paragraph).

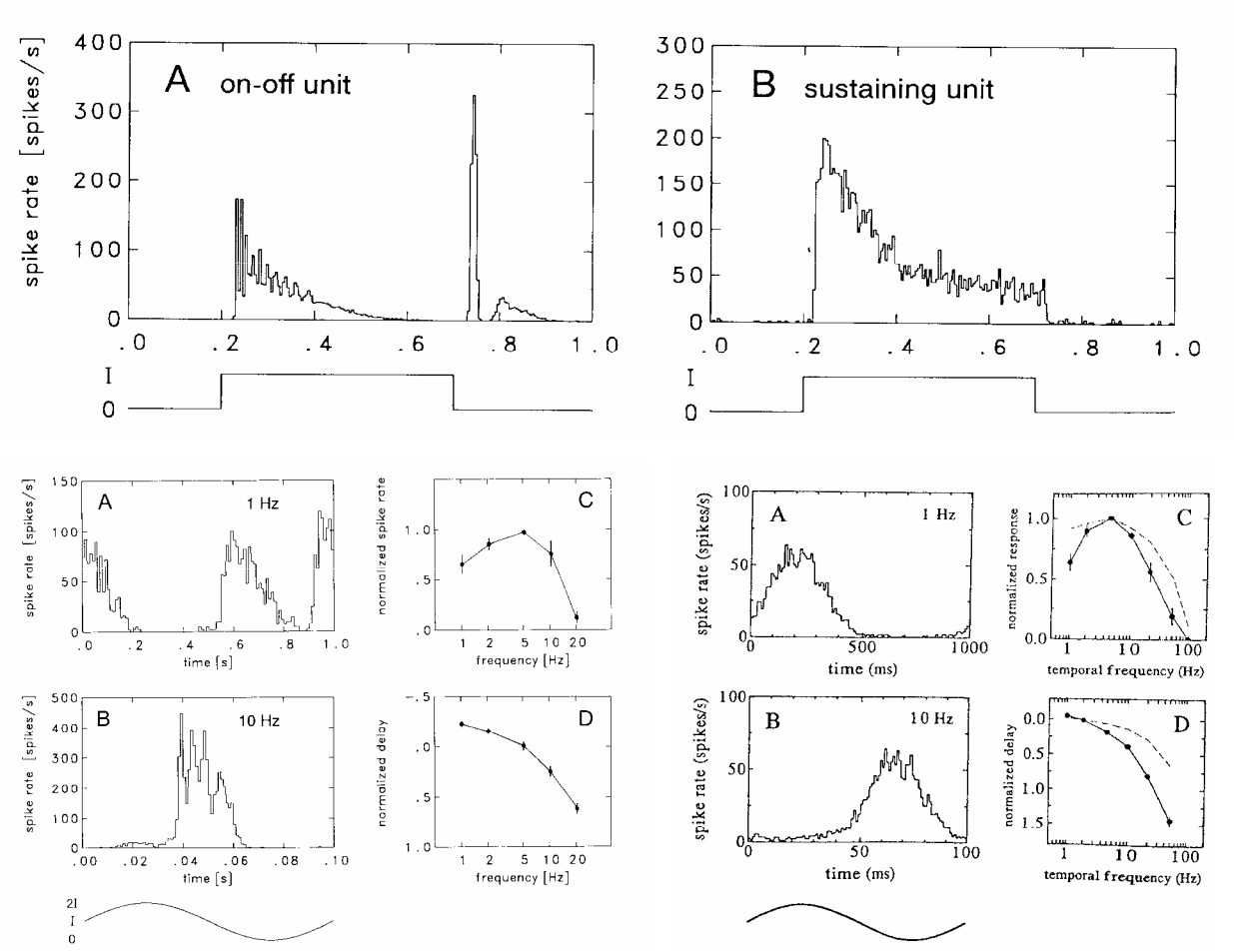


Fig.1.23 *Dynamic response properties of spiking units of the fly lamina (from van Hateren and Jansonius, 1991, 1993).*

1.3.3.2. Motion Detection

Many neurons in the visual system respond to visual motion. A particularly successful model for motion detection, shown in Fig.1.21, proposes that the light intensities derived form neighboring locations in the retina are multiplied with each other after one of them has been low-pass filtered (Reichardt, 1961, Borst and Bahde 1987, Borst and Egelhaaf, 1989, 1993). This operation is done in a mirror-symmetrical way, and the final output is the difference between the subunits. In such a motion detector, the spacing of the inputs and the time-constants in the detector determine the velocity tuning of the system.

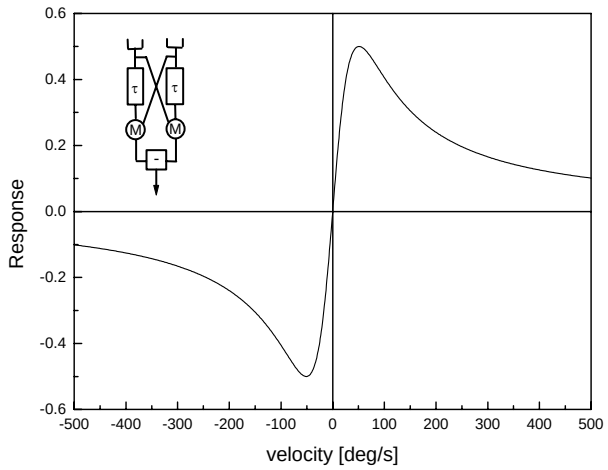


Fig.1.24 *Response of a Reichardt detector as a function of pattern velocity. The inset shows the structure of the model with two low-pass filters (1st order with time-constant τ), two multipliers (M) and a subtraction stage (-).*

How can we calculate this relation? We consider a sine-wave of wavelength λ traveling with a constant velocity v [°/s]. The inputs are spaced by Δφ. This results in the following input signals:

$$S_1 = I + \Delta I \cdot \sin(v / \lambda \cdot t)$$

$$S_2 = I + \Delta I \cdot \sin(2\pi \cdot v / \lambda \cdot (t - \Delta t)) = I + \Delta I \cdot \sin(2\pi \cdot v / \lambda \cdot t - 2\pi \cdot v / \lambda \cdot \Delta\varphi / v) =$$
$$I + \Delta I \cdot \sin(2\pi \cdot v / \lambda \cdot t - 2\pi \cdot \Delta\varphi / \lambda);$$

We introduce the following relations:

$$\omega = 2\pi v / \lambda; \text{ and } \Delta t = \Delta\varphi / v;$$

This results in:

$$S_1 = I + \Delta I \cdot \sin(\omega t);$$

$$S_2 = I + \Delta I \cdot \sin(\omega(t - \Delta t)) = I + \Delta I \cdot \sin(\omega - 2\pi \cdot v / \lambda \cdot \Delta\varphi / v) = I + \Delta I \cdot \sin(\omega t - 2\pi \cdot \Delta\varphi / \lambda);$$

Low-pass filtering these input signals results in:

$$L_1 = I + A(\omega)\Delta I \cdot \sin(\omega t + \Phi(\omega));$$

$$L_2 = I + A(\omega)\Delta I \cdot \sin(\omega t + \Phi(\omega) - 2\pi\Delta\varphi / \lambda);$$

The output of the detector then becomes:

$$R(t) = L_1 S_2 - L_2 S_1 =$$
$$= I^2 + I\Delta I \sin(\omega t) + IA(\omega)\Delta I \sin(\omega t + \Phi(\omega)) + \Delta I^2 \sin(\omega t + \Phi(\omega)) \sin(\omega t - 2\pi\Delta\varphi / \lambda) -$$
$$- I^2 - IA(\omega)\Delta I \sin(\omega t + \Phi(\omega) - 2\pi\Delta\varphi / \lambda) - I\Delta I \sin(\omega t) - \Delta I^2 \sin(\omega t + \Phi(\omega) - 2\pi\Delta\varphi / \lambda) \sin(\omega t);$$

The I^2-terms cancel. Being interested in the temporal average of the response only, all additive sine terms become zero, too. We are left with:

$$\bar{R} = \Delta I^2 A(\omega)\left[\sin(\omega t + \Phi(\omega)) \sin(\omega t - 2\pi\Delta\varphi / \lambda) - \sin(\omega t + \Phi(\omega) - 2\pi\Delta\varphi / \lambda) \sin(\omega t)\right];$$

Applying the formula

$$\sin(\alpha)\sin(\beta) = \frac{1}{2}(\cos(\alpha - \beta) - \cos(\alpha + \beta)); \text{ results in:}$$

$$\bar{R} = \Delta I^2 A(\omega)\frac{1}{2}[\cos(\omega t + \Phi(\omega) - \omega t + 2\pi\Delta\varphi / \lambda) - \cos(2\omega t + \Phi(\omega) - 2\pi\Delta\varphi / \lambda) -$$
$$- \cos(\omega t + \Phi(\omega) - 2\pi\Delta\varphi / \lambda - \omega t) + \cos(2\omega t + \Phi(\omega) - 2\pi\Delta\varphi / \lambda)];$$

$$\bar{R} = \Delta I^2 A(\omega)\frac{1}{2}[\cos(\Phi(\omega) + 2\pi\Delta\varphi / \lambda) - \cos(\Phi(\omega) - 2\pi\Delta\varphi / \lambda)];$$

Here the terms depending on t became zero because of the time average we calculate.
We now apply the above formula from right to left:

$$\bar{R} = \Delta I^2 A(\omega) \sin(\Phi(\omega)) \sin(-2\pi\Delta\varphi / \lambda);$$

This is identical to:

$$\bar{R} = \Delta I^2 A(\omega) \sin(-\Phi(\omega)) \sin(2\pi\Delta\varphi / \lambda); \qquad (1.34)$$

In case we assume the low-pass to be of 1st order, we obtain:

$$\bar{R} = \Delta I^2 \frac{1}{\sqrt{1 + \tau^2\omega^2}} \sin(a \tan(\tau\omega)) \sin(2\pi\Delta\varphi / \lambda);$$

The following relation exists between a sine and a tangent:

$$\sin(x) = \frac{\tan(x)}{\sqrt{1 + \tan^2(x)}};$$

Applying this yields:

$$\bar{R} = \Delta I^2 \frac{1}{\sqrt{1 + \tau^2\omega^2}} \cdot \frac{\tau\omega}{\sqrt{1 + \tau^2\omega^2}} \cdot \sin(2\pi\Delta\varphi / \lambda) = \Delta I^2 \frac{\tau\omega}{1 + \tau^2\omega^2} \cdot \sin(2\pi\Delta\varphi / \lambda); \qquad (1.35)$$

The term, which depends on the temporal frequency of the stimulus is identical to the amplitude spectrum of a band-pass (Formula 1.33). Thus, the response is optimum at the following frequency ω :

$$\omega = 1/\tau; \qquad (1.36)$$

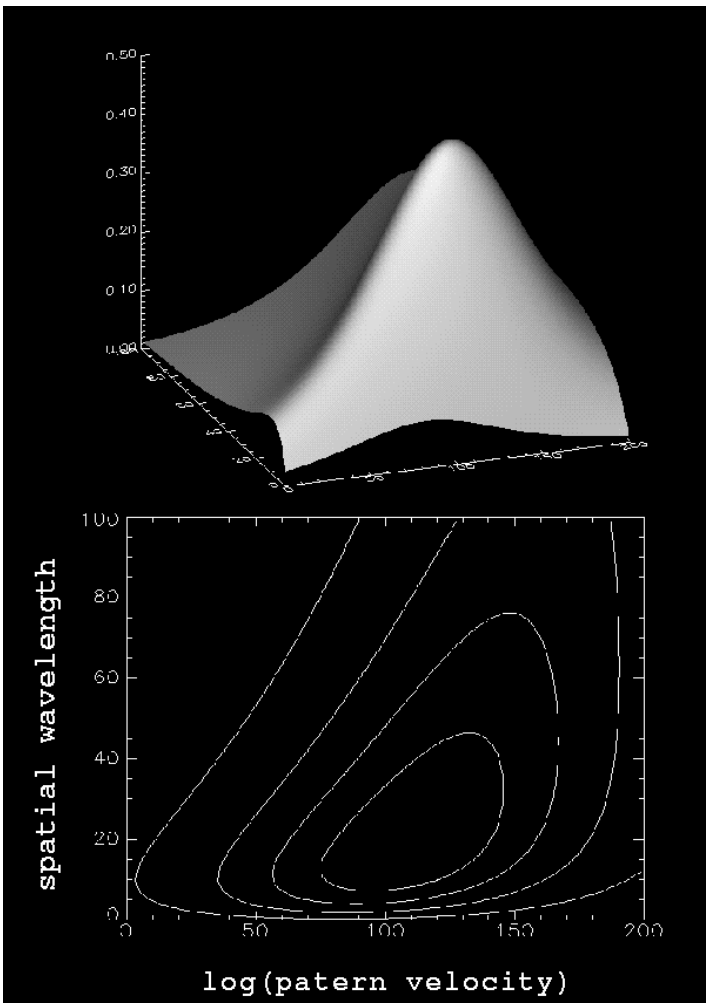


Fig.1.25 *Response of a motion detector of the correlation-type as a function of pattern velocity (logarithmic scaling) and spatial wavelength of the pattern. The pattern wavelength starts at values of twice the receptor spacing. Thus, no spatial aliasing occurs. For each wavelength, the response optima are shifted such that the ratio of velocity and wavelength remains constant.*

We see the following:

1. The average response should be, unlike a speedometer, not proportional to the pattern velocity but depend on it in a more complex way with an optimum (Fig. 1.24).
2. The average response of such a motion detector to a moving sine grating should be proportional to the square of its contrast (Formula 1.35).
3. It should be proportional to the ratio of input spacing and spatial wavelength (Formula 1.35).
4. Its response maximum should be always at the same temporal frequency, i.e. for larger pattern wavelength higher optimal velocities are expected (Formula 1.35, Fig.1.25).

The steady-state response of such a detector is shown in Fig.1.22 as a function of pattern wavelength and velocity. Such a particular behavior has indeed been observed in the so-called optomotor responses of insects and vertebrates, too. In all these experiments, the experimental subject was surrounded by a periodic grating rotating at a constant velocity. Usually, the

intended turning response of the subject was monitored while the subject itself was tethered and, thus, not allowed to move freely. For *Drosophila*, the optomotor response was systematically studied for various pattern wavelengths and velocities. The results are shown in Fig.1.23. The optimal temporal frequency was found to be about 1 Hz. From Formula (1.36), one can calculate the time constant to be around $1/(2\pi f) \cong 160$ ms. This is under the assumption of a 1st order low-pass filter inherent in the motion detection system of *Drosophila*.

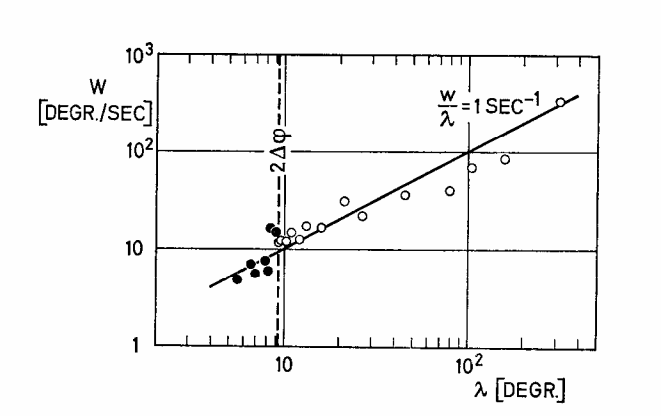


Fig.1.26 *Response optima of the turning behavior of the fruitfly Drosophila, measured for various spatial wavelengths as a function of pattern velocity. Note that the optima are shifted towards larger wavelengths with increasing velocity (from Goetz, 1978).*

What about a 2nd order low-pass? In this case, Formula (1.34) becomes:

$$\bar{R} = \Delta I^2 \frac{1}{1+\tau^2\omega^2} \sin(2a\tan(\tau\omega)) \sin(2\pi\Delta\varphi/\lambda);$$

Using the relation:

$$\sin 2\alpha = \frac{2\tan\alpha}{1+\tan^2\alpha};$$

We obtain:

$$\bar{R} = \Delta I^2 \frac{1}{1+\tau^2\omega^2} \cdot \frac{2\tau\omega}{1+\tau^2\omega^2} \sin(2\pi\Delta\varphi/\lambda) = \Delta I^2 \frac{2\tau\omega}{(1+\tau^2\omega^2)^2} \cdot \sin(2\pi\Delta\varphi/\lambda);$$

Calculating the frequency ω at which R is maximum yields:

$$\frac{dR}{d\omega} = \Delta I^2 \frac{2\tau(1+\tau^2\omega^2)^2 - 2\tau\omega\cdot 2(1+\tau^2\omega^2)\cdot 2\tau^2\omega}{(1+\tau^2\omega^2)^4} \cdot \sin(2\pi\Delta\varphi/\lambda) = \Delta I^2 \frac{2\tau - 6\tau^3\omega^2}{(1+\tau^2\omega^2)^3} \cdot \sin(2\pi\Delta\varphi/\lambda);$$

Setting this to zero, we obtain $\omega = 1/3\tau;$ This results in an estimate of the time-constant of around 50 ms.