

The Fourier transform in optics: from continuous to discrete or from analogous experiment to digital calculus

P. C. LOGOFATU*, D. APOSTOL

Lasers Dept., National Institute for Laser, Plasma and Radiation Physics, PO Box MG-36, 077125, Magurele, Romania

Although extensively used in digital computation, the discrete Fourier transform (DFT) has a problematic relation with the physical or continuous Fourier transform (CFT), as we encounter it in practice, for instance in Fourier optics. This article first establishes the physical meaning of DFT by expressing it in terms of CFT using the Fourier series as an intermediary link between the two types of transforms. Then it studies some concrete examples of binary masks Fourier transformed using an optical plane wave and a lens and shows how one can make DFT a good approximation of CFT.

(Received June 29, 2007; accepted September 1, 2007)

Keywords: Fourier optics, Discrete Fourier transform, Signal sampling

1. Introduction

There are many textbooks dedicated to the study of the Fourier transform [1-8]. One issue that is not sufficiently dealt with in these references is the relation between the continuous (or analogous) and the discrete Fourier transforms (CFT and DFT). For instance in Fourier optics a common problem is the discrepancy between the Fourier transform of an input function performed physically using a lens and the digitally calculated Fourier transform of the sampled input. We do not refer here to the perturbations introduced by physical limitations such as the aperture or the aberrations of the lens. On the contrary, we assume infinite aperture and no aberrations. We refer to the inconsistency between the continuous input and output of CFT on the one hand and the discrete, sampled input and the resulting discrete output of DFT on the other hand. One may ask then what is physical meaning of the digitally calculated transform and how and if one can make this transform to be a good approximation of the physical transform. To this problem is this article dedicated.

In mathematics there are three types of Fourier transforms: (I) CFT, (II) the Fourier series and (III) DFT. Only the first type has full physical meaning, and can be accomplished for instance in optics by Fresnel diffraction using a lens or by Fraunhofer diffraction. The third type is a pure mathematical concept, although it is much more used in computation than the previous two for practical reasons. These three types of Fourier transforms are independent formalisms, they can stand alone without reference to one another and they are often treated as such, regardless of how logic and necessity generated one from another. Because only CFT has physical meaning, the other two types of transforms are mathematical constructs that have meaning only by expressing them in terms of the first. DFT is in use for multiple reasons among which the high computation speed due to the discovery of the Fast Fourier Transform (FFT) (or the Cooley-Tookey algorithm) [9] is not the least. Another reason is the fact

that there are hybrid cases in which DFT takes a physical appearance and not merely mathematical, as is the case of digital holography where either the input function to be Fourier transformed is, in a sense, discrete, as the computer generated holograms, or the recording devices have a discrete character, as the CCD cameras (see for instance references [10] and [11]). Therefore the need for a unifying view of the three types of Fourier transform arises, not only for reasons of conceptual economy, but from practical reasons as well. Fourier analysis may be experimental as well as digital. It is important to make the connection between the analogous Fourier transform that we perform experimentally and the discrete numeric calculus performed by computers. Not only in the above mentioned hybrid situations but in any situation in which one wants to use DFT for computation convenience, one has to switch back and forth from one formalism to another and this is difficult and confusing if the meaning of DFT is not well known in terms of the genuine Fourier transform which is the continuous one. The insight brought by such a unifying view of the three types of Fourier transforms makes one see them interconnected and deriving from one another rather than separate domains.

To our knowledge no mathematics or physics textbooks present such a unifying view of the three types of Fourier transforms, although all the necessary knowledge lies in pieces in the literature [1-8], not even in those dedicated specifically to DFT. In references [10] and [11] a correspondence is worked out between DFT and CFT, but only for band-limited functions and it is not rigorous. Also, it is just a correspondence, a link between them not an equivalence, not an expression of one in terms of the other. An integrated unifying presentation of the three types of Fourier transforms has then the character of a creative review, so to speak. In this paper such a unifying view is presented. The Fourier transform of type II, the Fourier series, besides its own independent worth, is shown to be an intermediary link between the Fourier transforms of types I and III, a step in the transition between them. Also some concrete cases are analyzed to illustrate how the discrete representation of the Fourier

transform should be interpreted in terms of the physical Fourier transform and how one can make DFT a good approximation of CFT.

2. Fourier transforms

The three types of Fourier transform are defined as following: (I) continuous, i.e. the calculation of the transform $G(f)$ is done for functions $g(t)$ ^a defined over the real continuum, that is the interval $(-\infty \dots +\infty)$, and the transformation is made by integration over the same interval

$$G(f) = \mathcal{F}\{g\}(f) = \int_{-\infty}^{+\infty} g(t) \exp(-i 2 \pi f t) dt, \quad (1)$$

$\mathcal{F}$ being the symbol for Fourier transform, with G being also generally defined over the real continuum, (II) Fourier series, where the function to be transformed is defined over a limited range $(0 \dots \Delta t)$ of the continuum, and instead of a Fourier transformed function defined over $(-\infty \dots +\infty)$ we have a series which represents the discrete coefficients of the Fourier expansion

$$G_n = \frac{1}{\Delta t} \int_0^{\Delta t} g(t) \exp(-i 2 \pi n t / \Delta t) dt, \quad (2)$$

with n an integer ranging from $-\infty$ to $+\infty$, and (III) the purely discrete Fourier transform where a list of numbers is transformed into another list of numbers by a summation procedure and not by integration

$$G_q = \frac{1}{N^{1/2}} \sum_{p=0}^{N-1} g_p \exp(-i 2 \pi q p / N), \quad (3)$$

where the list G_q has N elements just like g_p . Obviously, q runs from 0 to $N-1$. (For clarity purposes, in this paper we use consistently m and n to indicate the periodicity, and p

^a For simplicity we will refer to 1D functions as long as possible while having in mind 2D functions, which are of particular interest in optics. Also we did not take into consideration the sine and cosine Fourier transforms partly because they are particular cases of the complex Fourier transform, partly for simplifying the calculations but mostly because we are interested in the Fourier transform in optics, where the complex representation of the optical field is predominant. Moreover, although Fourier optics is about 2D optical fields, if the field is constant along one direction, which is not an uncommon situation, a 1D treatment is possible and desirable because it is simpler.

^b There are two conventions in optics for representing the evolution in time of the phase, $\exp(-i \omega t)$ and $\exp(i \omega t)$. In this article we use the former, because it is the convention used in reference [0], a classic in the field of Fourier optics. This convention requires the minus sign of the argument of the exponential that is the kernel of the Fourier transform.

^c Eqs. (1), (2) and (3) are not the only possible definitions of the Fourier transforms, they are often slightly changed for convenience purposes. (We do not refer here to the sign convention discussed in the previous footnote but to some multiplicative constants.) We chose these definitions not necessarily for convenience but because they seem the most natural to us, especially because the inverse Fourier transform has the same form as the direct Fourier transform except for the sign of the exponential.

and q to indicate the sampling of g and G respectively.) Although these three types of Fourier transforms can be considered independent formalisms, the three types of transforms are strongly connected. Indeed, the second may be considered a particularization of the first, by restricting the class of input functions g to periodic functions only and then to take into consideration for calculation purposes only one period from the interval $(-\infty \dots +\infty)$ over which g is defined. The third can be considered a particularization of the second, by requesting not only that the functions g to be periodic, but also discrete, to have values only at even spaced intervals Δt . Therefore the third type of Fourier transform may be considered an even more drastic particularization of the most general Fourier transform I.

That is the top-bottom view. It is possible another view, a bottom-top one, in which the second type of Fourier transform is considered a generalization of the third type, or a construct made starting from the third type, and the same thing can be said about the relation between the first and the second type. In order to pass from the Fourier transform type III to type II, in the list g_m which is a discrete sampling made at equal intervals Δt we make $\Delta t \rightarrow 0$ and $N \rightarrow \infty$, which results in the list g_m becoming a function of continuous argument and the numbers G_n are not anymore obtained by summation as in Eq. (3) but by integration as in Eq. (2). Now we are dealing with the Fourier transform type II. Making the function g periodic, by imposing

$$g(t + m \Delta t) = g(t), \quad (4)$$

where m is an arbitrary integer, and making the period $\Delta t \rightarrow \infty$, the discrete coefficients G_n become a continuum and we are back to the Fourier transform of type I.

3. Unifying view

3.1. From continuous Fourier transform to Fourier series

Since the formalism of the Fourier transform of type I is the most general and the only one with full physical meaning, we will express the two other formalisms in its terms. As we said before, if the input function g is periodic, then the corresponding output function G degenerates into a series. Indeed, if g has the period Δt as in Eq. (4) then the corresponding G can be written as

$$\begin{aligned} G(f) &= \sum_{m=-\infty}^{+\infty} \int_{m\Delta t}^{(m+1)\Delta t} g(t) \exp(-i 2 \pi f t) dt = \sum_{m=-\infty}^{+\infty} \int_0^{\Delta t} g(t) \exp[-i 2 \pi f (t + m\Delta t)] dt = \\ &= \int_0^{\Delta t} g(t) \exp(-i 2 \pi f t) dt \sum_{m=-\infty}^{+\infty} \exp(-i 2 \pi m f \Delta t) \end{aligned} \quad (5)$$

Without rigorous demonstration we will state here that the infinite sum of exponentials in the uttermost right hand side of Eq. (5) is an infinite sum of delta functions also called the comb function,

$$\sum_{m=-\infty}^{+\infty} \exp(-i 2 \pi m f \Delta t) = \frac{1}{\Delta t} \sum_{n=-\infty}^{\infty} \delta\left(f - \frac{n}{\Delta t}\right) \equiv \text{comb}(\Delta t f) \quad (6)$$

Indeed, one may check that the sum of exponentials is 0 everywhere except for f of the form $n/\Delta t$ when it becomes infinite. When f is of the form $n/\Delta t$ all the members of the infinite sum are equal to 1. When f differs however slightly from $n/\Delta t$, the phaser runs uniformly from 0 to 2π when the index m of the sum grows incrementally with the result that the contributions of the terms to the sum cancels each other out, although not necessarily term by term. (One should not confuse the delta function $\delta(t)$ with the sampling interval δt .) Introducing (6) into (5) one obtains

$$G(f) = \frac{1}{\Delta t} \sum_{n=-\infty}^{\infty} \delta\left(f - \frac{n}{\Delta t}\right) \int_0^{\Delta t} g(t) \exp(-i 2\pi n t / \Delta t) dt = \sum_{n=-\infty}^{\infty} G_n \delta\left(f - \frac{n}{\Delta t}\right) \quad (7)$$

where G_n were defined in Eq. (2). One can see from Eq. (7) that the Fourier series are a particular case of CFT, namely the Fourier transform G of a periodic function g of period Δt is a sum of delta functions of arguments shifted with $1/\Delta t$ intervals and with coefficients G_n that are the same as the coefficients defined in Eq. (2). Actually it is the coefficients from Eq. (2) that are the Fourier series, and not the function defined in Eq. (7), but the correspondence is obvious.

In optical experiments one can see that happening when a double periodic mask, with perpendicular directions of periodicity, modulates a plane monochromatic light wave and the resulting optical field distribution is Fourier transformed with the help of a lens. In the focal plane of the lens, where the Fourier spectrum is formed, we have a distribution of intensely luminous points along the directions of periodicity. The luminous points do not have, of course, a rigorous delta distribution, they are not infinitely intense and they have areas. This departure from ideal is due to the fact that neither the mask nor the plane wave are ideal. The image that is Fourier transformed is neither infinite nor rigorously periodic, since the intensity of the plane wave decreases from a maximum in the centre to zero towards the periphery. But such an experiment is a good physical illustration of the mathematics involved in Eq. (7).

The difference between CFT and the Fourier series can be seen also from the inverse perspective of the Fourier expansion of g ,

$$g(t) = \mathcal{F}^{-1}\{G\}(t) = \int_{-\infty}^{+\infty} G(f) \exp(i 2\pi f t) dt = \int_{-\infty}^{+\infty} \sum_{n=-\infty}^{\infty} G_n \delta\left(f - \frac{n}{\Delta t}\right) \exp(i 2\pi f t) dt = \sum_{n=-\infty}^{\infty} G_n \exp(i 2\pi n t / \Delta t) \quad (8)$$

The difference is that in the case of a periodic function the Fourier expansion is a sum and not an integral anymore.

3.2. The discrete Fourier transform

For computation purposes we cannot use always a function g defined over the continuum but a sampled

version instead. There are many reasons that make the sampling of g necessary. It is possible that the function g , representing a physical signal, an optical field for instance, is not known a priori and only a detected sample of it can be known. It is possible that the function g cannot be integrated because it is too complicated or it cannot be expressed in closed form functions or it causes numerical instabilities in the calculation of CFT. A very important reason might be the fact that the function g may be the results of calculations too as in the case of computer-generated holograms. In this case, g is known only as a 2D matrix of numbers. (For simplicity, however, we will continue to work with 1D functions as long as possible. Hence, the input function g , the function to be Fourier transformed is a vector, a 1D matrix of numbers.) Then we have to sample the input function, and we can do that with the help of the comb function; this type of sampling will yield the value of g at evenly spaced intervals of chosen value δt and it can be written as

$$g^s(t) = g(t) \delta t \sum_{p=-\infty}^{\infty} \delta(t - p \delta t) = \delta t \sum_{m=-\infty}^{\infty} g(p \delta t) \delta(t - p \delta t) = \delta t \sum_{p=-\infty}^{\infty} g_p \delta(t - p \delta t), \quad (9)$$

where g_p are the sampled values $g(p \delta t)$ and p is integer. The subscript “s” comes from “sample”. One may notice that when $\delta t \rightarrow 0$ we have $g^s \rightarrow g$.

Now in order to obtain G we may apply CFT to g , but we may prefer to apply instead DFT defined in Eq. (3) for the reasons already mentioned in section 1. A sampled Fourier spectrum does not mean necessarily that the Fourier transform has to be performed discretely; we may perform it continuously and then sample the resulting continuous spectrum. But this would be a waste of effort. It is preferable to compute the Fourier spectrum discretely. But what physical correspondence has the discrete transform in reality, where the transform is continuous? To find out one needs to express DFT in terms of CFT, the physical transform. In other words using CFT of the sampled function we try to obtain the Fourier spectrum also under a sampled form. Here we can make use of the Fourier series, which now prove to be, as we stated before, an intermediary link between CFT and DFT. We want the Fourier spectrum to be discrete, but how can we accomplish that?

We know from subsection 2.1 that CFT of a periodic function is a discrete even spaced function. Since we want G to be discrete, then we have to make g periodic. The input function g is not generally periodic but in most practical cases it has values only over a finite domain of arguments. Suppose g is non-zero only for arguments in the interval $t \in (0 \dots \Delta t)$. We define a periodic function g^p as

$$g^p(t) = \sum_{m=-\infty}^{\infty} g(t + m \Delta t). \quad (10)$$

The subscript “p” means that g^p is the periodic version of g . A function that is non-zero only over the interval $(0 \dots \Delta t)$, does not have to be sampled to infinity, but only where has non-zero values. Sampling g over the interval Δt gives us the same quantity of information as sampling g_p

to infinity if, for simplicity, we choose the sampling interval δt so that

$$\Delta t = N \delta t, \quad (11)$$

where N is an integer. To make the input function both discrete and periodic, we have to combine the forms (9) and (10) of g together, taking into account (11),

$$g^{sp}(t) = \delta t \sum_{m=-\infty}^{\infty} \sum_{p=0}^{N-1} g_p \delta(t + m \Delta t - p \delta t). \quad (12)$$

Here the subscripts “sp” means that the function g^{sp} is both sampled (discrete) and periodic. We know from the general properties of CFT of its reciprocal character [0,0] and that the inverse CFT is very similar to CFT itself [see Eqs. (1) and (8); DFT has a similar property]. A double CFT reproduces the input function up to an inversion of coordinates. Therefore, if the CFT of a periodic function is an even-spaced discrete sampling, then CFT of an even-spaced discrete sampling has to be a periodic function. Let us check this assertion by calculating the CFT of (12)

$$\begin{aligned} G^{sp}(f) &= \delta t \int_{-\infty}^{\infty} \sum_{m=-\infty}^{\infty} \sum_{p=0}^{N-1} g_p \delta(t + m \Delta t - p \delta t) \exp(-i 2 \pi f t) dt = \\ &= \delta t \sum_{p=0}^{N-1} g_p \exp(-i 2 \pi f p \delta t) \sum_{m=-\infty}^{\infty} \exp(i 2 \pi f m \Delta t) \end{aligned} \quad (13)$$

We make use again of the property (6) and obtain

$$\begin{aligned} G^{sp}(f) &= \delta t \sum_{p=0}^{N-1} g_p \exp(-i 2 \pi f p \delta t) \frac{1}{\Delta t} \sum_{q=-\infty}^{\infty} \delta\left(f - \frac{q}{\Delta t}\right) = \\ &= \frac{1}{N} \sum_{q=-\infty}^{\infty} \delta\left(f - \frac{q}{\Delta t}\right) \sum_{p=0}^{N-1} g_p \exp(-i 2 \pi q p / N) \end{aligned} \quad (14)$$

We notice that G^{sp} is periodic with the period $\Delta f = 1/\delta t$, because adding $n/\delta t$ to the argument f this causes only a reindexation with nN of the infinite sum of delta functions that does not cause any modification to G^{sp} precisely because the sum is infinite. We also notice that the coefficients of the delta functions are, up to a multiplication constant, DFT of Eq. (3). We may rewrite then G^{sp} as

$$G^{sp}(f) = \frac{1}{N^{1/2}} \sum_{n=-\infty}^{\infty} \sum_{q=0}^{N-1} G_q \delta\left(f - \frac{q}{\Delta t} - \frac{n}{\delta t}\right). \quad (15)$$

In the expression (15) the periodic character of G^{sp} , namely of period $1/\delta t$, is more clearly visible than in Eq. (14). Also G^{sp} is sampled at intervals of $\delta f = 1/\Delta t$. Both g^{sp} and G^{sp} are discrete and periodic. This is connected with the property of the function comb that it is invariant to CFT [0,0], in other words CFT of the comb function is also the comb function. The sampling of g is made with a comb function and it was to be expected to retrieve the comb function in the expression of G^{sp} . One may say that

the DFT of a sampled function g is the CFT of the comb function weighted with g and the result is a comb function weighted with G . It should be noted that G^{sp} has the same number of distinct elements as g^{sp} , N . This is to be expected since through Fourier transform no information is lost but only represented differently. We should be able to retrieve the same amount of information from G as from g , therefore the number of samples should be the same for both functions^d. It is also noteworthy the inverted correspondence between the sampling intervals and the periods of g^{sp} and G^{sp} . The sampling interval of G^{sp} is the inverse of the period of g^{sp} , and the period of G^{sp} is the inverse of the sampling interval of g^{sp} .

Eqs. (12) and (15) represent the physical equivalent of DFT. Only for a function like g^{sp} , periodic and consisting of evenly spaced samples, can one calculate the Fourier transform as in Eq. (3), i.e. discretely, and such functions do not exist in reality, and one may wonder what is the usefulness of it all. There is usefulness inasmuch as g^{sp} relates to g and G^{sp} to G , and they are related because we built g^{sp} starting from g . But in constructing g^{sp} we adapted g for the purpose of discrete computing and we departed from the original g and consequently from G . Now we have to find how close are CFT and the physical equivalent of DFT performed on g and how can one bring them closer.

One may see now that the correspondence between CFT and DFT given in references 0 and 0 is of a different kind than the one shown above. In those references the output function G is not discrete. Only the sampled values of G correspond to the DFT of only the sampled values of g . Also our relation between CFT and DFT has the advantage of illustrating better some properties of DFT such as the cyclic or toroidal character of DFT [0,0].

4. Correspondences to reality

4.1 Generalities

One step toward bringing G^{sp} closer to reality, i.e. closer to G , is choosing an appropriate sampling interval δt . We saw that the period of G^{sp} is $1/\delta t$ and the obvious choice for δt is one that makes the period of G^{sp} large enough to encompass the real Fourier spectrum G , if it is finite. This is the gist of the Whittaker-Shannon sampling theorem [0]. If the sampling interval is too large, then we are affected by aliasing. The Fourier spectrum is not

^d The value $1/\Delta t$ for the sampling interval δf was determined by our choice to obtain a discrete Fourier spectrum and to have a rigorous equivalent of DFT in terms of CFT. This choice enables also the use of the FFT algorithm, which is an important advantage both theoretical and practical not to be neglected. However, the sampling interval of the Fourier spectrum does not have to be $1/\Delta t$. It is a matter of choice. One may talk of DFT even in the case of a different δf . We will just loose the advantages mentioned above, plus the symmetry properties and the rigorousness and, more important, we would also loose information, because from the output function G thus obtained we will not be able to reproduce the input function g .

generally finite, but for most practical cases it decreases rapidly at high frequencies. An exception occurs for the inputs with sharp discontinuities, for instance binary masks in optics that either let all the light propagate or stop it completely. But even the Fourier spectrum of these masks decreases at high frequency, and one can always establish an ad-hoc criterion, by trial and error, that whatever passes below a certain level of intensity is negligible. This means that if we choose a sampling period small enough (or a sampling frequency high enough) we should be able to obtain a good approximation of G and consequently of g .

The unifying view presented in section 3, specifically the outline of the periodicity of the physical equivalent of DFT makes easier to understand another important piece of information for someone who manipulates a discretely computed Fourier spectrum, namely the fact that, due to computation convenience, the FFT algorithm provides the Fourier spectrum as a list in which the first element corresponds to the zeroth frequency [0,7], the nonmodulated part of the input function. In practice, in a genuine Fourier spectrum, the zeroth frequency is the middle one, with the negative frequencies to its left and the positive frequencies to its right. In DFT, where both g and G are infinite and periodic, it does not matter what is the starting point of a period, but it does in practice, where we deal with the continuous transform. The discrete Fourier spectrum shows large amplitude for the highest frequencies, which is contrary to experience, which shows that the central frequencies have the highest amplitudes and the further you are from the centre, the lower are the amplitudes. (Of course, the Fourier spectrum of highly regular, periodic structures has “spikes” after intervals of low amplitude, but, as we mentioned above, even these spikes decrease toward high frequencies.) To make use of this theoretical information in physical reality, one has to take into account the property of periodicity (or the cyclical character) of the discrete Fourier spectrum, that is

$$G_q = G_{q-N} \cdot \quad (16)$$

Therefore we can make $G_{-1} \equiv G_{N-1}$, $G_{-2} \equiv G_{N-2}$, and so on. We are closer to the physical reality if we cut the right half of a 1D Fourier spectrum and paste it to the left of the remaining half. The problem gets more complicated for a 2D spectrum. In such a case the 2D spectrum should be divided in 4 quadrants and their order should be interchanged as follows: the lower-left with the upper-right quadrant and the lower-right with the upper-left quadrant.

4.2 Binary masks made out of identical squares

A particular case in which we may talk in a sense of “naturally” sampled input functions is the case of binary masks (transmittance 0 or 1) that are formed out of identical squares. A 1D grating as in Fig. 1.a (mask A), or a 2D grating as in Fig. 1.b (mask B) are such examples. We chose those masks because our purpose is to compare the discrete and the continuous Fourier transforms and for

those masks CFT can be calculated analytically. The masks are not sampled functions in the sense of Eq. (9), there are no delta functions in their expression. They are, however, sampled in the sense that for evenly spaced rectangular areas the transparency functions are constant, hence, the functions can be represented, as discrete matrices of samples. We chose masks with a low number of samples, only 32×32 , not because DFT is difficult or time-consuming (actually, due to the existence of the FFT algorithm DFT can be done very quickly for quite large number of samples), but to make easier the calculation of CFT, which is indeed considerably time-consuming. Also, at small number of samples the differences between the discrete and the continuous spectrum can be more easily seen. The period of the gratings, both horizontally and vertically, was chosen to be 7 squares, so that it does not divide 32; in this way the Fourier spectrum gets a little complicated and we avoid symmetry effects that may obscure the points we want to make.

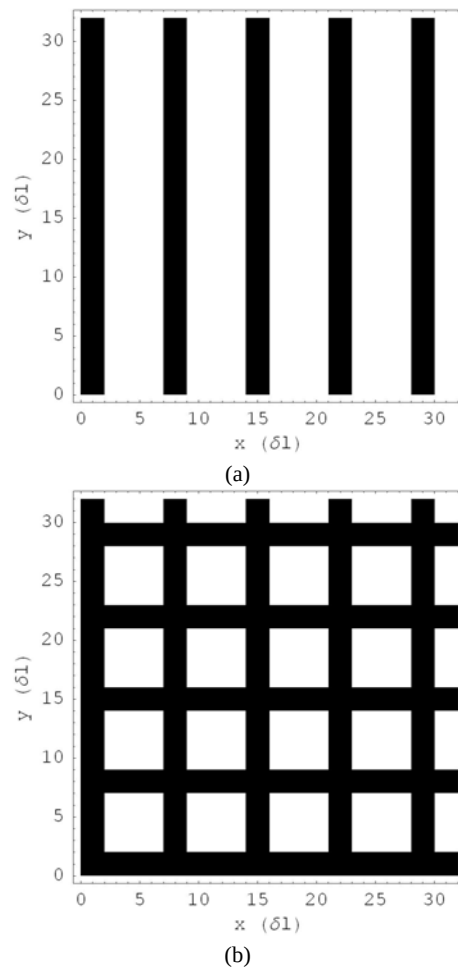


Fig. 1. Binary masks: a) 1D grating and b) 2D grating. Both gratings consist in 32×32 squares, black or white, of equal dimension δl . We labelled the two masks “A” and “B”. It is assumed that outside the represented area of the masks the input light is completely blocked, i.e. the transparency function is zero.

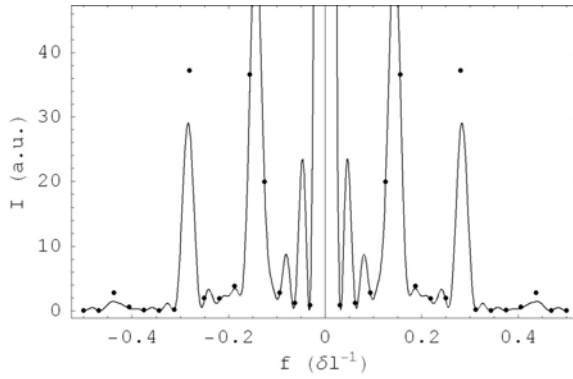


Fig. 2. Discrete (dots) and continuous (solid line) Fourier power spectra for mask A shown together vs the spatial frequency. Only the central part of the continuous spectrum for which DFT provides output values is represented. The spatial frequency is expressed in δl^{-1} units.

It seems natural to represent the masks from Fig. 1, for DFT calculations, as 32×32 matrices with values 1 or 0 corresponding to the transmission coefficient of the squares of the masks. In Fig. 2 are shown together the squared absolute values of the DFT of the previously discussed type of matrix for mask A and the CFT of mask A. The square absolute value is the power or the luminous intensity, the only directly measurable parameter of the light field. The CFT of mask A has the expression

$$G(f_x, f_y) = \delta \text{sinc}(f_x \delta) 32 \delta \text{sinc}(32 f_y \delta) \sum_{p=1}^{32} g_p \exp[-i 2 \pi f_x (p-1) \delta], \quad (17)$$

where δ is the dimension of the squares and

$$\text{sinc}(t) \equiv \frac{\sin(\pi t)}{\pi t}, \quad (18)$$

and obviously N from Eq. (3) is here 32. Since mask A is 1D [in a limited sense only; if it were truly 1D then the dependence on f_y of G would be of the form $\delta(f_y)$] in the calculation of $|G|^2$ in Fig. 2 and the following figures that represent $|G|^2$ for mask A, we discarded the dependence on f_y and the entire contribution of the integration over y and we used only the part corresponding to the integration over x . The DFT calculations were rescaled (multiplied with $N^{1/2}$) so that they could be compared to the CFT calculations. All the CFT spectra represented in this article are calculated, hence they are simulations. A correctly done experiment of Fourier optics would yield, of course, the same results.

Two types of discrepancies can be noticed in Fig. 2 between CFT and DFT. First the discrete and the continuous Fourier spectra have different values for the same spatial frequency, sometimes there is even a considerable difference. Second, the discrete spectrum does not offer information for the domain of spatial frequencies in between two discrete values, and the interpolation of the sampled values does not lead always to a good approximation. The structure of the continuous

spectrum is much richer than that of the discrete spectrum. There are of course many reasons why the discrete Fourier spectrum $|G^{sp}|^2$ is different than the genuine spectrum $|G|^2$. One reason is the periodicity. The g input functions are not generally periodic or at least they are not infinitely periodic. The input functions illustrated in Fig. 1 are periodic in a sense, meaning they have a limited number of periods, but rigorously periodic means an infinity of periods.

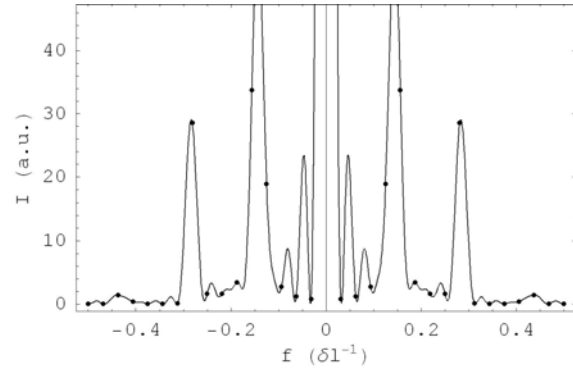


Fig. 3. Discrete (dots) and continuous (solid line) Fourier power spectra vs. the spatial frequency for mask A represented together, for a higher (10 times) sampling rate than that used for the computed data shown in Fig. 2. Only the portion of the both spectra corresponding to the spatial frequency range of Fig. 2 is shown.

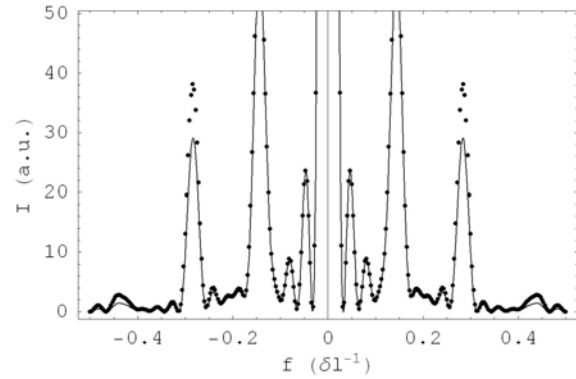


Fig. 4. Discrete (dots) and continuous (solid line) Fourier spectra vs. the spatial frequency for mask A extended so that to include part of the surrounding darkness shown together. Compared to the original mask used for the calculations of Fig. 2, the mask was now extended 10 times, which accounts for the higher density of dots of the DFT output, which now numbers 320 samples. Although in the figure 320 points are represented, the spatial frequency range is the same as that of Figs. 2 and 3.

Corresponding to the two types of discrepancies there are also two ways of improving the discrete spectrum. One way is to increase the sampling rate. We can do that by “swelling” the sampling array, inserting more than once the value corresponding to a square. The increased

sampling rate improves the agreement between DFT and CFT. We illustrated in Fig. 3 only the calculations for the same spatial frequencies as those represented in Fig. 2, not just to ease the comparison but also because the spectrum outside is very weak and, hence, negligible. One big source of the difference between CFT and DFT from Fig. 2 is the $\text{sinc}(f_x \delta l)$ function, that is the Fourier transform of the rectangular function. Because for Fig. 3 by allotting more samples for each square we had a higher sampling rate, the rectangular shape was felt in DFT calculations. We increased the sampling rate 10 times and, as one can see in Fig. 3, the continuous and the discrete spectrum have now closer values, they are almost on top of each other. Also CFT now provides values for higher spatial frequency but still only those corresponding to the spatial frequency range of Fig. 2 are shown in Fig. 3, not only for comparison purposes, but also because the power values corresponding to these new higher frequencies are very small. Because there are now 10 times more elements than in the “natural” sampling, and actually in this new sampling the same elements are just repeated 10 times, in order to be able to compare DFT and CFT we had to divide the DFT spectrum by 10^2 .

Another way to improve the conformity of the DFT to CFT is to increase the sampled area. Since outside the area of mask A there is no structure, only darkness, and g is just zero, when making the sampling of g we are generally tempted to discard this surrounding darkness. But when we express DFT in terms of CFT, as we saw in section 3.2, we consider that the structure of mask A is periodically repeated back to back in what we know to be just darkness. Therefore, in order to improve the likeness of DFT to reality it is a good idea to pad the original sampled function with zeros to the left and to the right to account for the surrounding darkness. We padded with zeros so that the original sampled array of values is increased 10 times. In Fig. 4 the squared absolute values of the DFT and CFT for the new sampled function are again compared, and, although their values are still different, now DFT offers more information, enough for interpolation. We did not need here another type of rescale of DFT than that done in Fig. 3 in order to have a meaningful comparison to CFT, because the new elements added to the input sampling were just zeros.

The two procedures for improving the likeness of DFT to CFT described above and illustrated in Figs. 3 and 4 may be combined and the result is shown in Fig. 5. The sampling array used for the calculation of DFT is now both “swollen” and extended, having 100 times more elements than the “natural” sampling. Now DFT is both closer to CFT and richer in information. Now DFT is both accurate and able to provide enough information for a correct interpolation.

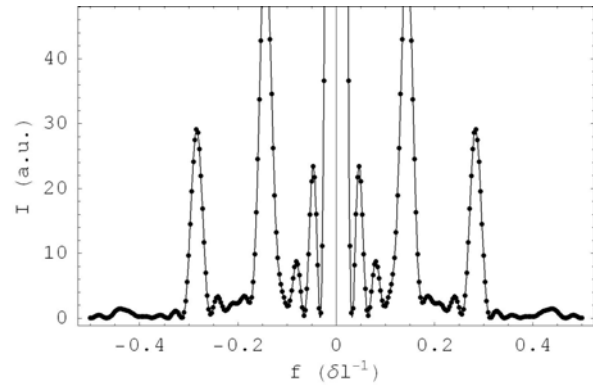


Fig. 5. Discrete (dots) and continuous (solid line) Fourier spectra vs. the spatial frequency for mask A both extended and sampled at higher rate shown together. For DFT an array of $32 \times 10 \times 10$ points was used, but only the 32×10 points corresponding to the spatial frequency range of Figs. 2-4 were shown.

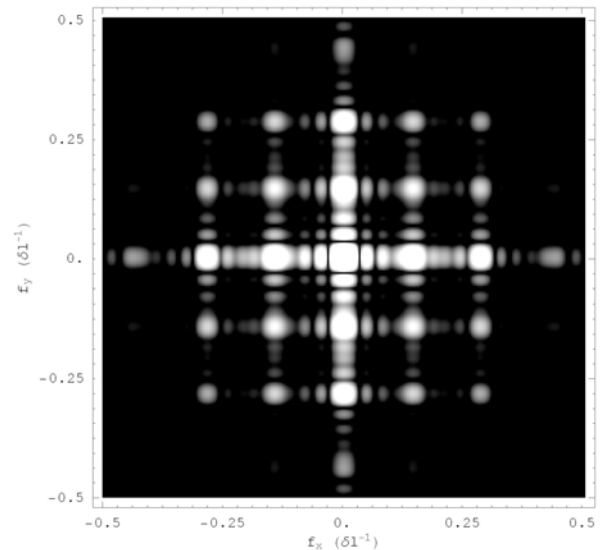


Fig. 6. The central high intensity portion of the continuous Fourier spectrum of mask B, chosen so that to match the spatial frequency ranges of the DFT of the “naturally” sampled input of mask B (see Fig. 7 below). The abscissa and the coordinate are the spatial frequencies and the light intensity of the spectrum is coded as levels of grey.

The similar procedure applied to mask A was also applied to mask B. We found appropriate to illustrate the procedure for mask B because optics is generally about images and these are 2D, not 1D, which is just a particular case, useful mostly for the easiness of the graphic representation than for practical purposes. The Fourier spectrum of a 2D mask is more difficult to represent. We chose to represent the spectrum as levels of grey, just as the eye would see it. Moreover, to simulate the vision of the eye we represented the logarithm of the luminous intensity. Another reason for using the logarithmic

representation is the fact that the Fourier spectrum decreases quite sharply with the spatial frequency and only the representation of the logarithm allows the fine shades to be visible.

In Fig. 6 the represented continuous Fourier spectrum is the logarithm of the squared absolute value of the function

$$G(f_x, f_y) = \delta^2 \text{sinc}(f_x \delta) \text{sinc}(f_y \delta) \sum_{p=1}^{32} \sum_{q=1}^{32} g_{pq} \exp[-i 2 \pi (f_x (p-1) + f_y (q-1)) \delta] \quad (19)$$

In Fig. 7 it is represented the discrete Fourier spectrum of the “natural” sampling of mask B, that is of a 32×32 matrix with each element having the value of the corresponding square element of mask B. The comparison of Figs. 6 and 7 shows marked differences. Just as in the case of mask A we tried next to compensate for the shortcomings of the “natural” sampling by extending the sampling and increasing the sampling rate. The result, shown in Fig. 8, is quasi identical to Fig. 6.

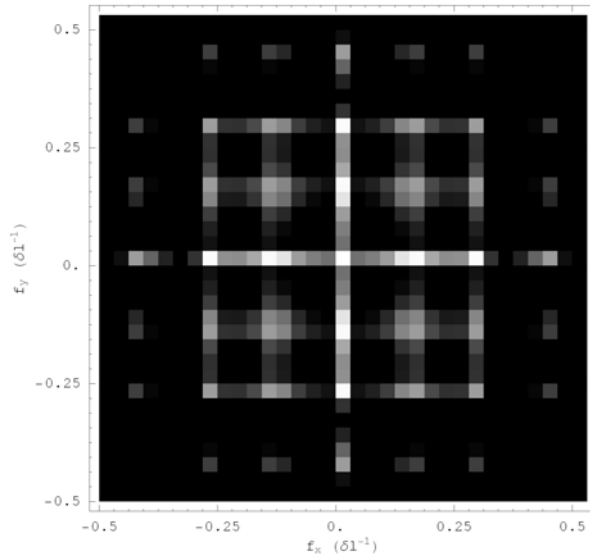


Fig. 7. The discrete Fourier spectrum of the “natural” sampling of mask B. The abscissa and the coordinate are the spatial frequencies, and the light intensity is coded as levels of grey, just as in Fig. 6.

As a side comment, one may notice that all the spectra shown in this article are symmetric. The 1D plots are symmetric with respect to the origin, and the 2D plots are symmetric with respect to the axes. There is a redundancy of information. The 1D plots contain in one vertical half all the information, and the 2D plots contain all the information in any of the 4 quadrants. This is due to the fact that we calculated the spectra of transmission masks that do not modify the phase of the input optical fields and we assumed the input wave to be a plane wave, which is actually a common situation in Fourier analysis. It is this property of the masks that causes the spectra to be symmetric. Rigorously speaking they are not symmetric, the phase differ in the 2 halves of the 1D plots and in the 2D plots the phase of 2 diagonal quadrants differs from

that of the other 2 diagonal quadrants. But the difference is just that they have conjugate complex values. The absolute values of 2 conjugate complex quantities is the same, hence the 4-fold symmetry of the 2D power spectra.

To give some dimensional perspective to the considerations presented in this subsection, it might be instructive to give a concrete value to δ and to specify the experimental conditions in which the Fourier transform is performed. We need very small masks in order to make the Fourier spectrum macroscopic, but also large enough so that sufficient light passes through and the Fourier spectrum is visible. $100 \mu\text{m}$ is such a value for δ . Then the masks A and B would be squares of 3.2 mm dimension. The Fourier spectra (continuous or discrete) represented in Figs. 2-8 are segments for the 1D case and squares for the 2D case having the dimension $\Delta f = 1/\delta = 10 \text{ mm}^{-1}$ in spatial frequency units. If the Fourier transform is performed by a lens of focal length $F = 1 \text{ m}$ and the light source is a He-Ne laser of wavelength $\lambda = 632.8 \text{ nm}$, then the length of either dimension of the represented spectra in the Fourier plane (the posterior focal plane of the lens) is $\lambda F \Delta f = 6.328 \text{ mm}$.

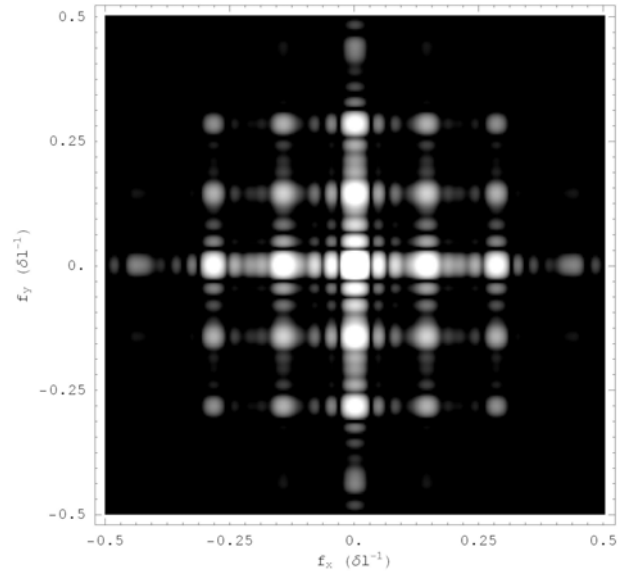


Fig. 8. The central high intensity portion of the discrete Fourier spectrum of a modified mask B, extended 9 times and sampled at 9 times a higher rate than the “natural” sampling rate. The portion corresponds to the spatial frequencies ranges of Figs. 6 and 7. The likeness of the image to Fig. 6 is obvious. As in Figs. 6 and 7 the abscissa and the coordinate are spatial frequencies and the light intensity is coded as levels of grey.

5. Conclusions

The problem of the relation between DFT and CFT is investigated in this article. In order to understand the physical meaning of DFT, we expressed it in terms of CFT. The Fourier series was a useful tool in this

endeavour, because it is an intermediary link between CFT and DFT. Namely, the two properties of both the input function and the Fourier spectrum of DFT, periodicity and discrete character, are present in the Fourier series, except that the input function is just periodic and the Fourier spectrum is just discrete. The connection between periodicity and the discrete character is stressed, namely it is shown that the periodicity of the input/output implies a discrete character of the output/input and vice-versa. For convenience the derivations were made for 1D input functions but they can be easily and straightforwardly extended to 2D input functions, if the sampling is done over 2 mutually perpendicular directions and the sampled area is rectangular. It is shown that DFT is the CFT of a periodic input of delta functions in which case the output is also periodic and composed of delta functions. The incongruence between DFT and CFT indicates that DFT may not be a good approximation of CFT, and some concrete numerical examples prove it. Two masks, first 1D and the second 2D were studied with respect to the concordance of their discrete with their continuous Fourier spectra. It has been shown that if the sampling rate and the extension of the masks are properly chosen (large enough) DFT is a good approximation of CFT. No generally valid criterion for the concordance between DFT and CFT is given, only the ways of improving the concordance are indicated and shown to be sufficient for the particular cases studied in this article.

References

- [1] J. W. Goodman, *Introduction to Fourier Optics*, McGraw Hill, New York (1996).
- [2] R. Bracewell, *The Fourier Transform and Its Applications*, McGraw-Hill, New York (1965).
- [3] E. O. Brigham, *The Fast Fourier Transform: An Introduction to Its Theory and Application*, Prentice-Hall, Englewood Cliffs NJ (1973).
- [4] L. Yaroslavsky, M. Eden, *Fundamentals of digital optics*, Birkhäuser, Boston (1996) ch 4.
- [5] R. J. Collier, C. B. Burckhardt, L. H. Lin, *Optical holography*, Academic Press, New York (1971) ch 19.
- [6] J. R. Fienup,
www.courses.rochester.edu/fienup/Optics_261/DFT&FresnelVoice.ppt
- [7] W. H. Press, S. A. Teukolsky, W. T. Vetterling, B. P. Flannery, *Numerical Recipes in C++*, Cambridge University Press, Cambridge (2002) chs. 12, 13.
- [8] G. B. Arfken, H. J. Weber, *Mathematical methods for physicists*, Harcourt / Academic Press, San Diego (2001) chs. 14, 15.
- [9] J. W. Cooley, J. W. Tookey, *Mathematics of Computation* **19**, 297 (1965).
- [10] A. W. Lohmann, D. P. Paris, *Appl. Opt.* **6**, 1739 (1967).
- [11] M. C. King, A. M. Noll, D. H. Berry, *Appl. Opt.* **9**, 471 (1970).

* Corresponding author: petre.logofatu@inflpr.ro