REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)* 07-11-2005	**2. REPORT TYPE** Final Report	**3. DATES COVERED** *(From – To)* 1 June 2003 - 24 October 05

4. TITLE AND SUBTITLE Robust Optimal Filtering and Argument Transforming for Feedback Controlling Optical Systems	**5a. CONTRACT NUMBER** FA8655-03-D-0001, Delivery Order 0008
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) Dr. Alexander Vitalievich Razgulin	**5d. PROJECT NUMBER**
	5d. TASK NUMBER
	5e. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Moscow State University Moscow 119899 Russia	**8. PERFORMING ORGANIZATION REPORT NUMBER** N/A
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) EOARD PSC 802 BOX 14 FPO 09499-0014	**10. SPONSOR/MONITOR'S ACRONYM(S)**
	11. SPONSOR/MONITOR'S REPORT NUMBER(S) EOARD Task 02-9007

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

A profound mathematical study has been accomplished for the problem of optimal discrete Fourier filtering: existence and uniqueness theorems both for direct and conjugate initial-boundary value problems for a functional differential parabolic equation have been proved; the solvability of the optimal filtering problem has been demonstrated; a functional gradient formula has been obtained. A projection finite-element scheme has been developed both for the direct and conjugate problems. A conditional gradient iterative procedure has been developed and applied for finding the optimal Fourier filter to control the feedback optical system. The optimization performance has been analyzed depending on the parameters of the optical system and on the number of the control Fourier channels.

A mathematical statement of the optimal control problem has been developed for a distributed spatial argument transform. It is applied to studying controlling optical systems with nonlocal transforms of the light wave in the feedback loop. The approach developed utilizes a generalized way of the determination of the argument transform. As an advantage, one can use wide range of nonsmooth and irreversible argument transforms, and consequently achieve better results with the optimal control. The approach developed forms the basis for constructing a projection finite-element method of approximation of both direct and conjugate problems in the same manner. As a result, computational versions of projection gradient and conditional gradient methods have been worked out for target functional minimization.

15. SUBJECT TERMS

EOARD, Optical Theory, Optical communications, texture coding, Hermite series, edge detection, optimal discrete Fourier filtering, feedback optical system, conditional gradient, irreversible argument transforms, projection finite-element method

16. SECURITY CLASSIFICATION OF:			**17. LIMITATION OF ABSTRACT**	**18, NUMBER OF PAGES**	**19a. NAME OF RESPONSIBLE PERSON**
a. REPORT UNCLAS	**b. ABSTRACT** UNCLAS	**c. THIS PAGE** UNCLAS	UL		PAUL LOSIEWICZ, Ph. D.
					19b. TELEPHONE NUMBER *(Include area code)* +44 20 7514 4474

Robust Optimal Filtering and Argument Transforming for Feedback Controlling Optical Systems

AFOSR grant through CRDF Project # RP0-1391-MO-03

Final Report

Project Director: Alexander Vitalievich Razgulin
Principal Organization: Moscow State University, Computational Mathematics and Cybernetics Department, room 691, Leninskie Gory, 119992, GSP-2, Moscow, Russia.
Effective Period: June 1, 2003 – May 30, 2005

1. Objectives.

The main goal of the Project is developing a solid theoretical and numerical background for the analysis of control systems with spatially distributed feedback. The subjects of particular interest are the study of new approaches to classical Fourier filtering and irreversible spatial transforms as novel types of distributed feedback control. For such a type of analysis, new effective numerical methods for processing information in the frequency domain should be developed. These numerical techniques can be useful in multi-frame alignment for low-light applications (optical coherence tomography), stereo-image analysis for robotics, automated medical image processing, texture matching, etc. Therefore, the objectives split into three closely connected tasks, accomplishment of which could provide substantial improvement of the performance and robustness of distributed feedback-controlling systems subjected to micro-control devices.

Task 1: Distributed Fourier-Hermite Projection Filtering.

Nowadays, “texture” is an important term in the world of computer vision, graphical applications, computer modeling, and optical system controlling. The task of texture processing first came up with the task of making automatic image recognition systems. Textures in the context of image recognition can be used in different ways: image segmentation; detecting the physical properties and materials of the surfaces represented in the image; detecting the shape and position of the objects and surfaces in the image. These problems need also an effective method of edge detection that is stable to noise.

Our main goal was to develop a new effective method for the tasks of image filtering, texture identification and edge detection.

The first approach to implement the above goals was based on Hermite series expansion time-frequency filtering techniques. Non-hierarchical and hierarchical texture coding schemes using this approach have been developed. This approach allowed us also to create a new stable method of edge detection.

The second approach is based on Tikhonov’s regularization method. New image filtering algorithms using analytically found Green’s functions have been created and implemented. A numerical method for solving the corresponding Euler-Lagrange equation has been also designed.

Task 2: Controlling Discrete Fourier Filtering.

The development of a new class of optical filters based on Fourier series with a discrete filtering kernel (discrete filters) instead of traditional integral convolution attracts a substantial practical interest. Discrete spatial filters are much easier to fabricate using photolithography technology. Contemporary technology allows producing filters with a certain number of “steps” for a micron of height. On the other hand, there are technological capabilities for producing high-speed dynamic discrete Fourier filters, using MEMS devices. Consequently, the possibility to use filters with a discrete number of steps in thickness and a finite number of elements in the transverse plane is a favorable option, and such filters should be modeled numerically. In the framework of our Project, spatial filters are expected to enhance the characteristics of 2D nonlinear optical systems with a controlling feedback loop. Thus, specialized software is required to model the dynamics of such a system containing a spatial filter in its feedback loop as well as the theoretical background of the methods implemented.

Task 3: Controlling Irreversible Transforms of Spatial Arguments.

Large-scale spatial transformations in the feedback loop create novel possibilities for synthesizing light fields with desired properties, as compared with conventional spatial filtering techniques. Contraction, reflection, shift, and rotation are examples of reversible large-scale transforms capable of dramatically changing the spatio-temporal dynamics of the feedback system and actually appear as a novel and convenient controlling tool. More advantages could be accomplished by applying irreversible transforms resulting in the formation of localized and periodic 1D optical patterns. A new mathematical approach to give an adequate description of irreversible transforms was recently suggested by participants of the Project. The approach is based on the theory of distributions and takes into account a wide range of Lebesgue-measurable 2D irreversible transforms including focusing onto a point or a planar curve. It is thus necessary to accomplish comprehensive investigation of optimal control problems for optical systems governed by PDEs with controlling reversible and irreversible 2-D argument transforms in the feedback loop. Consequently, a solid mathematical background and specialized numerical methods are required to study the corresponding optimization algorithms and to apply them to optical information processing.

2. Technical Description of Work Accomplished

2.1. Distributed Fourier-Hermite Projection Filtering

In the scope of the Project, non-hierarchical and hierarchical texture coding schemes using Hermite series expansion time-frequency approach have been developed. This approach allowed us to create a new stable method for edge detection. A new method of image filtering and edge detection based on the Tikhonov regularization method has been also designed.

1.0.0. Non-hierarchical texture coding scheme

A filtering method for texture analysis has been elaborated. It is based on applying a set of band-pass filters to extract the so-called *feature vectors* from the image. Each filter responds most strongly to a selected spatial-frequency and orientation bands. In our work, Hermite expansion series are taken to define the filters. The Hermite functions are defined as:

$\psi_n(x) = \dfrac{(-1)^n e^{x^2/2}}{\sqrt{2^n n! \sqrt{\pi}}} \cdot \dfrac{d^n(e^{-x^2})}{dx^n}$. They can be also calculated recursively:

$$\psi_0 = \frac{1}{\sqrt[4]{\pi}} \cdot e^{-x^2/2}$$

$$\psi_1 = \frac{\sqrt{2}x}{\sqrt[4]{\pi}} \cdot e^{-x^2/2}$$

$$\psi_n = x\sqrt{\frac{2}{n}} \cdot \psi_{n-1} - \sqrt{\frac{n-1}{n}} \cdot \psi_{n-2}, \forall n \geq 2$$

1-D Hermite functions, $\psi_n(x,y) = \psi_n(x) \cdot 1$, have a strongly pronounced orientation. To perform an adequate analysis, the functions should be rotated by certain angles. During the first stage of investigations, it was found that 8 orientation angles, 0, 22.5, 45, 67.5, 90, 112.5, 135, and 157.5 degrees, are enough to obtain stable representation of texture features.

In this approach, to get the feature vectors, we considered functions $\psi_n(x,y)$, where n=0..64, and introduced 6 energy coefficients:

$$E_1 = (\alpha_0)^2 + (\alpha_1)^2,$$

$$E_2 = (\alpha_2)^2+(\alpha_3)^2 + (\alpha_4)^2,$$

$$E_3 = (\alpha_5)^2 + (\alpha_6)^2 + (\alpha_7)^2 + (\alpha_8)^2,$$

$$\ldots$$

$$E_6 = (\alpha_{33})^2 + (\alpha_{34})^2 + \ldots + (\alpha_{63})^2 + (\alpha_{64})^2 ,$$

where (*x*,*y*) is the source image, α_i are the Hermite series coefficients. Hence, with the 8 orientations we get a 48-dimensional feature vector for each texture example. An analysis shows (see Fig.1.1) that the feature vectors obtained by this filtering algorithm are suitable for texture parameterization.

Fig. 1.1 Two examples of the same texture and corresponding feature vectors. In the 2-D representations of the feature vectors, spatial frequency levels increase bottom-up, and the orientation changes from left to right, being wrapped around.

2.0.0. Hierarchical texture coding scheme

When decomposing a function by the standard scheme and using the coefficients of 1-D Hermite series expansion without rescaling, some problems arise with texture analysis. The most significant of these is that images taken from the real world are unstructured, i.e. the luminosity, scale, size, and/or perspective of objects may vary, causing inaccuracy during computer processing. To avoid the above-mentioned problem, we used the idea of hierarchical coding.

This approach allows us to use several sets of patterns, which correspond to different features the image may contain. These sets differ in spatial scale. 1-D Hermite series expansion with a changeable spatial scale enables solving these problems.

To compare texture parameterization methods, a special program module has been developed. It allows loading grayscale texture samples and applying to them three different texture parameterization methods. As a result, the program returns a table of comparison of the methods applied to the loaded texture patterns. An example of such a comparison table is shown in Fig.1.2

a	*B*	Ordinary Coding	Hierarchical Coding	Hierarchical Coding without subtractions
A1	A2	0.004505	0.005349	0.003739
B1	B2	0.009905	0.008160	0.007742
C1	C2	0.017778	0.011848	0.009946
A3	A4	0.008807	0.006047	0.002422
B3	B4	0.020075	0.010667	0.006275
C3	C4	0.128322	0.101591	0.080176

A1	B1	0.488420	0.492238	0.491653
B1	C1	0.315644	0.304597	0.305442

Fig. 1.2 Differences in processing some texture samples with different coding schemes. Brodatz textures of the same type are denoted by the same letters.

The comparison table contains the absolute values of the differences between the feature vectors of the loaded texture patterns. The difference is calculated using the following formula: $\rho(a,b) = \dfrac{\sum_{i=1}^{48}(\alpha_i - \beta_i)^2}{\max\{\sum_{i=1}^{48}\alpha_i^2, \sum_{i=1}^{48}\beta_i^2\}}$, where α_i and β_i are the Hermite series coefficients for the corresponding coding scheme.

The analysis of the comparison table shows that the best method to be applied to texture parameterization and segmentation tasks is the method called “Hierarchical Coding without subtractions”. Assuming this, we have chosen this method to solve the texture segmentation problem.

To perform the texture segmentation task using the “Hierarchical Coding without subtractions” method, another program module has been developed. It allows one to load a grayscale picture, fracture it into blocks of a selected size, tune the method parameters, and perform the segmentation. Some results of the segmentation program operation are shown in Fig.1.3 and Fig.1.4

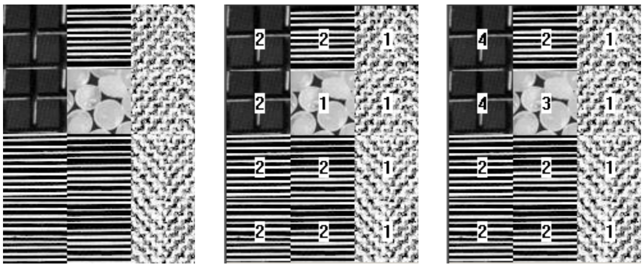


Fig. 1.3. The results of applying the texture segmentation algorithm with different block texture differences thresholds to a Brodatz texture patterns-based image.

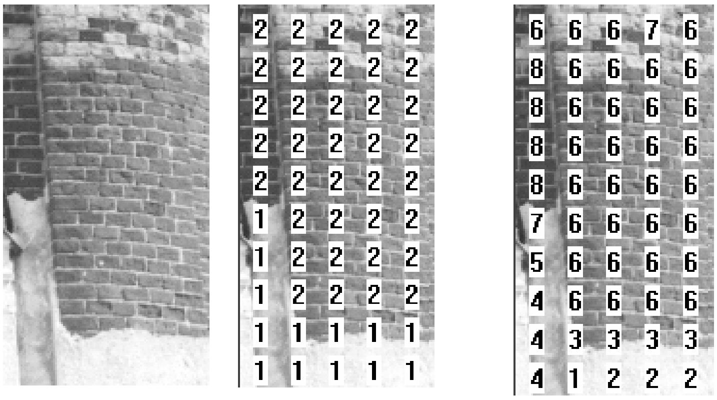


Fig. 1.4. The results of applying the texture segmentation algorithm with different blocks texture differences thresholds to a natural-scene image.

Hermite function-based parameterization implies some calculating problems. The main problem is that the calculation of the Hermite function $H_n(x)$ value at a given point, x, is very machine-time expensive for a large n. So, we were forced to calculate these recursively. To solve the problem and to increase the calculation performance of the program, we developed a caching calculation scheme, which is based on the fact that many values of $H_n(x)$ are retrieved repeatedly at different stages of calculation. Such an approach requires more memory, but, on the other hand, it allows us to increase the calculation speed up to 10 times.

3.0.0. Stable edge detection methods

The problem of texture analysis, parameterization, and discrimination using 1-D and 2-D Hermite series, which we are developing, needs a stable method to perform image edge detection. In (J.F. Canny. A computational approach to edge detection // *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 8(6):679--698, November 1986) three criteria for edge detection are given. These have been used in a similar form in a good deal of vision research. They are: 1. Good detection. There should be the minimum number of false negatives and false positives. 2. Good localization. The edge location must be reported as close as possible to the correct position. 3. Only one response to a single edge. Basically, these criteria are connected with the mathematical problem of stable second derivative zero crossing detection for a perturbed function. Our research was focused on developing new algorithms for solving this problem on the basis of the Tikhonov regularization method and 1-D Hermite series approach.

A. Edge detection method based on Tikhonov's regularization

The method based on Tikhonov's regularization looks as follows: for a function $\bar{y} \in H^1$, we have an approximation $\|\bar{y} - y_\delta\|_{L^2} \le \delta$. We want to find $\tilde{y} \in H^1$, such that $\|\tilde{y} - \bar{y}\|_{H^1} \to 0$ for $\delta \to 0$. This problem is being solved by minimizing the Tikhonov functional:

$$\|\tilde{y} - y_\delta\|_{L^2}^2 + \alpha \left\| \frac{d}{dx} \tilde{y} \right\|_{L^2}^2 \to \min .$$

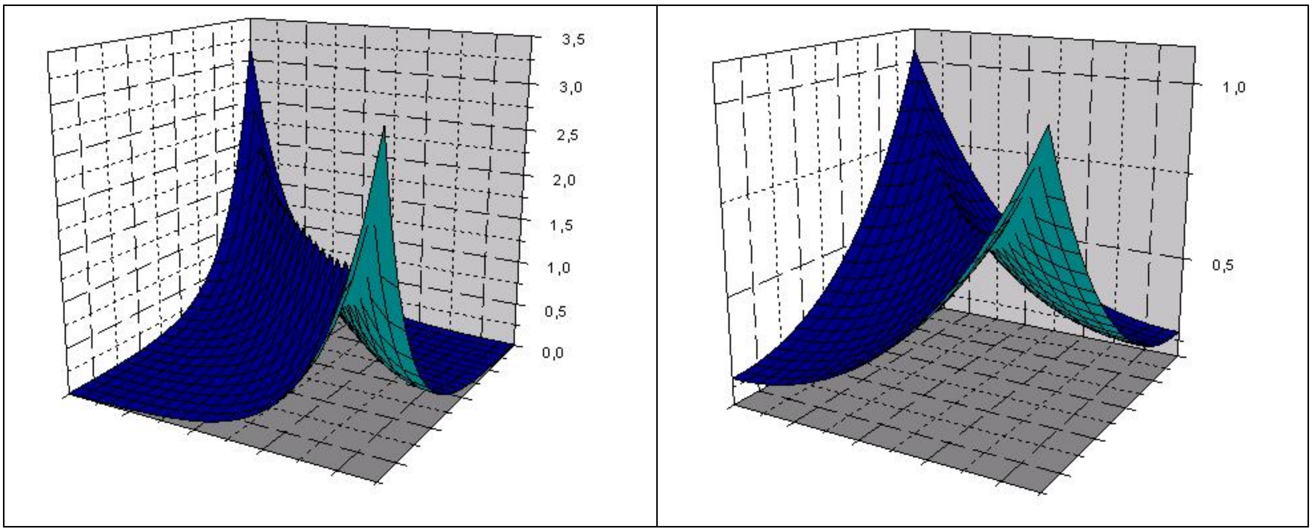
The problem of the Tikhonov functional minimization is reduced to solving the Euler equation for $\tilde{y}$ (below we use notation 'y' instead):

$$\begin{cases} -\alpha y'' + y = y_\delta \\ y'(1) = y'(-1) = 0 \end{cases} \Rightarrow \begin{cases} y'' - \dfrac{1}{\alpha} y = -\dfrac{1}{\alpha} y_\delta \\ y'(1) = y'(-1) = 0 \end{cases}$$

Its solution is $y_\alpha(x) = \int_{-1}^{1} G_\alpha(x,s) y_\delta(s) ds,$ with the kernel

$$G_\alpha(x,s) = \begin{cases} \dfrac{1}{\sqrt{\alpha}\, sh \dfrac{2}{\sqrt{\alpha}}} ch \dfrac{1}{\sqrt{\alpha}}(x-1) ch \dfrac{1}{\sqrt{\alpha}}(s+1) & s \le x \\ \dfrac{1}{\sqrt{\alpha}\, sh \dfrac{2}{\sqrt{\alpha}}} ch \dfrac{1}{\sqrt{\alpha}}(s-1) ch \dfrac{1}{\sqrt{\alpha}}(x+1) & s \ge x \end{cases}$$

The graphs of the kernel look as follows:



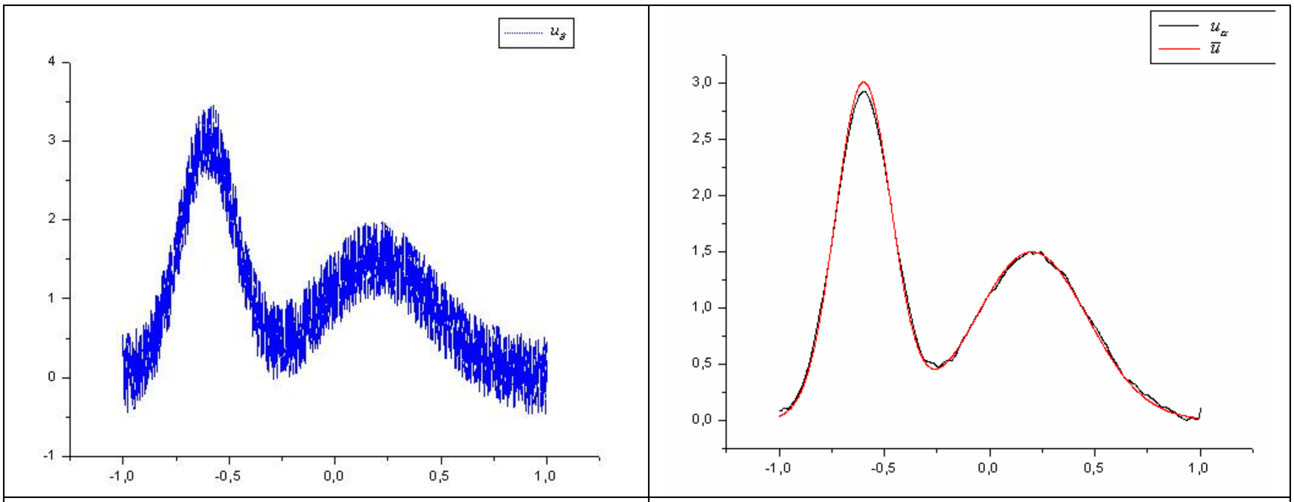
$\alpha = 0.1$ $\alpha = 10$

Fig. 1.5 Graphs of the kernel $G_\alpha(x,s)$, $-1 < x < 1, -1 < s < 1$.

To find $\alpha = \alpha(\delta)$, the solution $y_\alpha(x)$ must satisfy the following equation:

$$\varphi(\alpha) = \int_{-1}^{1} \left(y_\alpha(x) - y_\alpha(x)\right)^2 dx = \delta^2 \Rightarrow \alpha(\delta)$$

The proposed method finds function y with the minimum norm of the derivative in L^2 among all the functions $y(x)$ from the set $\|y - y_\delta\|_{L^2} \le \delta$.



Noisy source data — Signal restoration

Fig. 1.6 Parameter α selection by discrepancy method (δ^2 =0.165, α =0.000986).

A very interesting consequence of the above approach is the fact that the solution satisfies the equation $-\alpha y'' + y = y_\delta$. So, zero crossing of the second derivative is equivalent to the change of the sign of $y - y_\delta$. This allows us to construct a method for stable second derivative zero crossing detection. Its application to image processing gives us a new method of stable edge detection. For edge detection we use Laplacian zero-crossing criterion. From the equation

$$-\alpha u'' + u = u_\delta$$

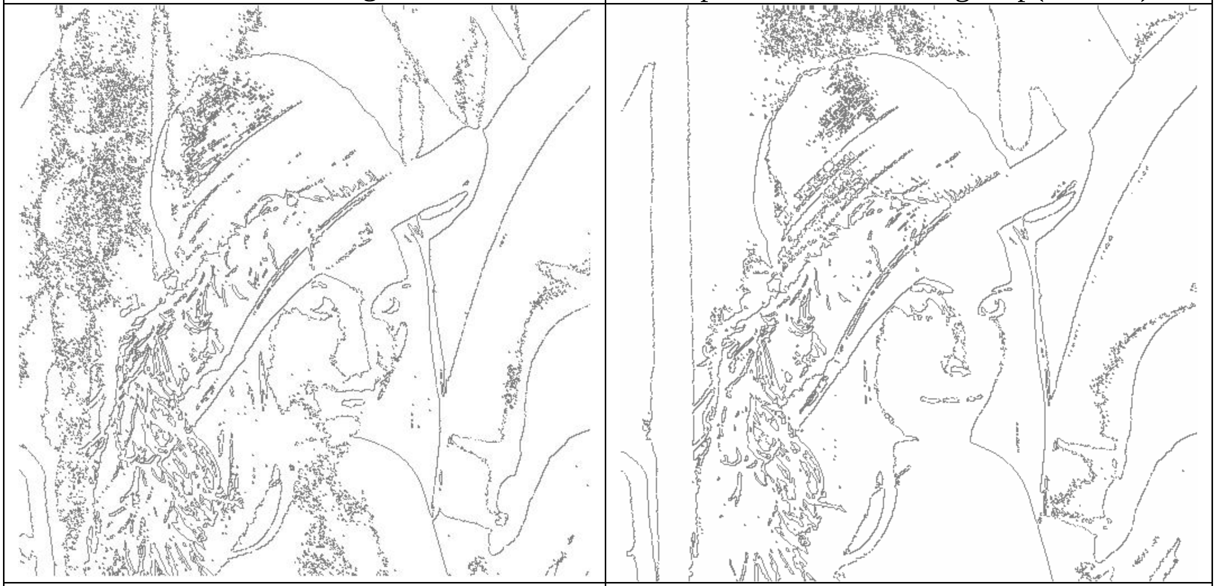
it follows that

$$u'' = 0 \Leftrightarrow u - u_\delta = 0.$$

So, we smooth every row of the noisy image u_δ, and find the second derivative in the y direction, $\frac{\partial^2 u}{\partial y^2}$. Then, we smooth every column of the initial image u_δ, and find the second derivative in the x direction, $\frac{\partial^2 u}{\partial x^2}$. Summing these results, we obtain a non-directional second derivative (the Laplacian operator: $\Delta u = \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}$). Figure 1.7 shows applying this algorithm to edge detection.

Source image

Laplacian zero crossing map($\Delta u = 0$)



Map of $\frac{\partial^2 u}{\partial y^2} = 0$

Map of $\frac{\partial^2 u}{\partial x^2} = 0$

Fig. 1.7 Edge detection results.

The derivatives used for edge thresholding can be calculated by the following formula:

$$u'(x) = \frac{1}{\sqrt{\alpha}\, sh\frac{2}{\sqrt{\alpha}}} \int_{-1}^{x} sh\frac{1}{\sqrt{\alpha}}(x-1)\, ch\frac{1}{\sqrt{\alpha}}(s+1)\, y_\delta(s)\, ds - \frac{1}{\sqrt{\alpha}\, sh\frac{2}{\sqrt{\alpha}}} \int_{x}^{1} ch\frac{1}{\sqrt{\alpha}}(s-1)\, sh\frac{1}{\sqrt{\alpha}}(x+1)\, y_\delta(s)\, ds.$$

The proposed method can be used for the considered task of edge detection, but the design of 2D methods can make the algorithms more effective. To implement this approach, variational regularization methods for 2D image filtering were considered. Typical variational methods of image restoration obtain a filtered version of some degraded image u_δ as the minimizer, u_α, of:

$$E_\alpha(u) = \int_\Omega [(u - u_\delta)^2 + \alpha\Psi(|\nabla u|^2)]dx.$$

The first summand encourages the similarity between the restored image u_α and the original one u_δ, while the second summand rewards smoothness. The smoothness weight α > 0 is called the regularization parameter. In our case, the regularizer Ψ is supposed to satisfy the following conditions:

1. $\Psi(.)$ is continuous for any compact $K \subseteq [0,\infty)$.
2. $\Psi(|.|^2) : \Re^n \to \Re$ — is a convex function .
3. $\Psi(.)$ is increasing in $[0,\infty)$.
4. There exists a constant $\varepsilon > 0 : \Psi(s^2) \geq \varepsilon s^2$.

For this class of regularization methods, one can establish a well-posed and scale-space framework, as for nonlinear diffusion filtering. The regularization parameter α should be considered as a scale.

Theorem 1.1. (Properties of regularization methods)

(a) (Well-posedness and regularity) Let $u_\delta \in L^\infty(\Omega)$. Then the functional $E_\alpha(u)$ has a unique minimizer u_α in the Sobolev space $H^1(\Omega)$. Moreover, $u_\alpha \in H^2(\Omega)$ and $\|u_\alpha\|_{L^2(\Omega)}$ continuously depends on α.

(b) (Extremum principle) Let $a = \inf_\Omega u_\delta$ and $b = \sup_\Omega u_\delta$. Then $a \leq u_\alpha \leq b$ on Ω.

(c) (Average grey level invariance) The average grey level $\mu = \frac{1}{|\Omega|}\int_\Omega u_\delta(x)dx$ remains constant under regularization: $\frac{1}{|\Omega|}\int_\Omega u_\alpha(x)dx = \mu.$

(d) (Convergence to a constant image for $\alpha \to \infty$) $\lim_{\alpha\to\infty}\|u_\alpha - \mu\|_{L^p(\Omega)} = 0$ for any $p \in [1,\infty)$.

If Ψ is differentiable, then the minimizer of $E_\alpha(u)$ satisfies the Euler-Lagrange equation:

$$\frac{u - u_\delta}{\alpha} = div(\Psi'(|\nabla u|^2)\nabla u).$$

Summary of some Ψ functions:

Name of the function Ψ	Ψ(t)	$\Psi'(t)$	Convexity
Tikhonov	T	1	Yes
Total Variation	$\sqrt{t}$	$\frac{1}{2\sqrt{t}}$	Yes
Bouman and Sauer	t^α 0.5<α<1	$\alpha t^{\alpha-1}$ 0.5<α<1	Yes
Perona and Malik	1-exp(-t)	exp(-t)	No
Geman and McClure	$\frac{t^2}{1+t^2}$	$\frac{1}{(1+t)^2}$	No

Hebert and Leahy	Log(1+ t)	$\frac{2}{1+t}$	No
Green	Log[ch($\sqrt{t}$)]	$\frac{th(\sqrt{t})}{2\sqrt{t}}$	Yes
Hyper surfaces	$\sqrt{1+t}-1$	$\frac{1}{2\sqrt{1+t}}$	Yes

The above methods were numerically analyzed. The Tikhonov and Perona-Malik approaches were found as the most promising. The first of these approaches (edge detection by Tikhonov's regularization) is a 2D generalization of our previous 1D method. In this case, if we use $\Psi(t)=t$, the Euler-Lagrange equation is reduced to: $u-u_\delta=\alpha\Delta u$. Therefore, $\Delta u=0 \Leftrightarrow u-u_\delta=0$. The proposed method allows us to obtain edges in an image on the basis of Laplacian zero crossing analysis.

B. Edge detection based on 1-D Hermite series approach

The Hermite expansion of function *f(x)* looks as follows:

$$f(x)=\sum_{i=0}^{\infty}\alpha_i\cdot\Psi_i(x)\,,\quad \alpha_i=\int_{-\infty}^{\infty}\Psi_i(x)\cdot f(x)dx\,,$$

where α_i is the *i*-th Fourier coefficient, and $\Psi_i(x)$ is the *i*-th Hermite function. Subsequently,

$$f''(x)=\left(\sum_{k=0}^{\infty}\alpha_k\cdot\Psi_k(x)\right)''=\sum_{k=0}^{\infty}\alpha_k\cdot\Psi_k''(x)\,.$$

Basing on the following formula that is valid for the Hermite functions

$$\Psi_n''+\left(2\cdot n+1-x^2\right)\cdot\Psi_n=0\,,$$

we can evaluate $f''(x)=\sum_{k=0}^{\infty}\alpha_k\cdot(x^2-2\cdot k-1)\cdot\Psi_k(x)$.

Let us now have a two-dimensional image $I(x,y)$ of size $m * n$. Basing on the above formula for one-dimensional functions decompositions, we define the following scheme for processing the image $I(x, y)$. The rows $\{L_i(x)\}_{i=1..n}$ and columns $\{C_j(y)\}_{j=1..m}$ are apparently analyzed for second derivate zeros. For each row $L_i(x)$ we have a set of second derivative values, $L_i''(x)$, x being a discrete argument: $x=1..m$. Having the set of second derivative values, we define the following criterion for a zero-crossing of $L_i''(x)$: x is considered a zero-crossing point if it matches the inequality

$L_i''(x)\cdot L_i''(x-1)<-\Delta$, Δ being a threshold parameter.

Second derivative zero points for $\{L_i(x)\}_{i=1..n}$ and $\{C_j(y)\}_{j=1..m}$ can be considered as edge points of the source image $I(x, y)$. Correspondingly these points are then plotted in the resulting image that depicts the edges found in the source image.

We have compared the described method of Hermite decomposition-based edge detection with the method based on Fourier decomposition. The method differs from the described one only in obtaining the second derivative values:

$$f(x)=\frac{a_0}{2}+\sum_{k=1}^{\infty}\left(a_k\cos kx+b_k\sin kx\right),$$

$$a_0=\frac{2\bar{f}_0}{\sqrt{2\pi}}=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)dx,\ a_k=\frac{\bar{f}_k}{\sqrt{\pi}}=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\cos kxdx,\ b_k=\frac{\bar{\bar{f}}_k}{\sqrt{\pi}}=\frac{1}{\pi}\int_{-\pi}^{\pi}f(x)\sin kxdx.$$

Then the second derivative can be calculated:

$$f''(x)=\left(\frac{a_0}{2}+\sum_{k=1}^{\infty}(a_k\cos kx+b_k\sin kx)\right)''=\sum_{k=1}^{\infty}k^2(-a_k\cos kx-b_k\sin kx).$$

The algorithm of finding edges using this formula is the same as that described above for Hermite decomposition-based edge detection.

Even before practical comparison of using the Hermite and Fourier series for edge detection in noisy images, we can see a great advantage of Hermite series approach. It lies in the fact that the coefficient a_k of the function f expansion is multiplied by k for the Hermite series, while the k^2 multiplier arises instead in the Fourier series approach. This means that the Hermite series approach is more stable when dealing with noisy data.

While having almost identical results when detecting edges in noise-free images, our method displays preferable results when detecting edges in noisy images. The followings Figures illustrate the results obtained.

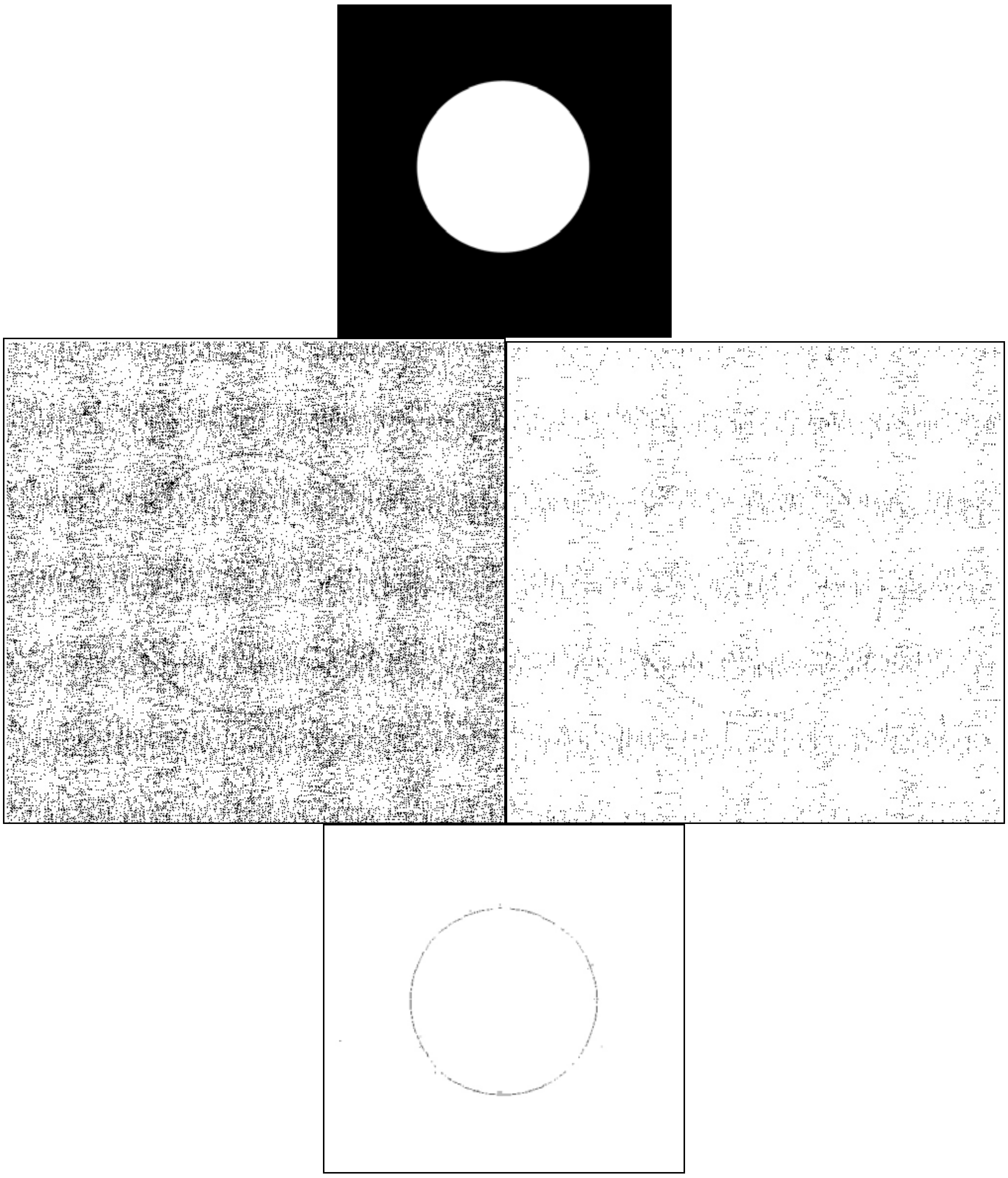


Fig. 1.8 Four images are plotted: source image of a white circle, with a 20% Gaussian noise; the results of 190-harmonic Fourier-based edge detection method application (two samples with different thresholds); and the result of 190 Hermite functions based edge detection method application. It can be seen that the Hermite functions based method gives an adequate result while the Fourier method does not yield adequate edge detection. It is also illustrated that changing the threshold does not help the Fourier method to produce good results, because either (with a small threshold) we obtain a too noisy resultant image that does not emphasize edges or (with a bigger threshold) we obtain an image with very little points that depict the edges but still many points that are just noise.

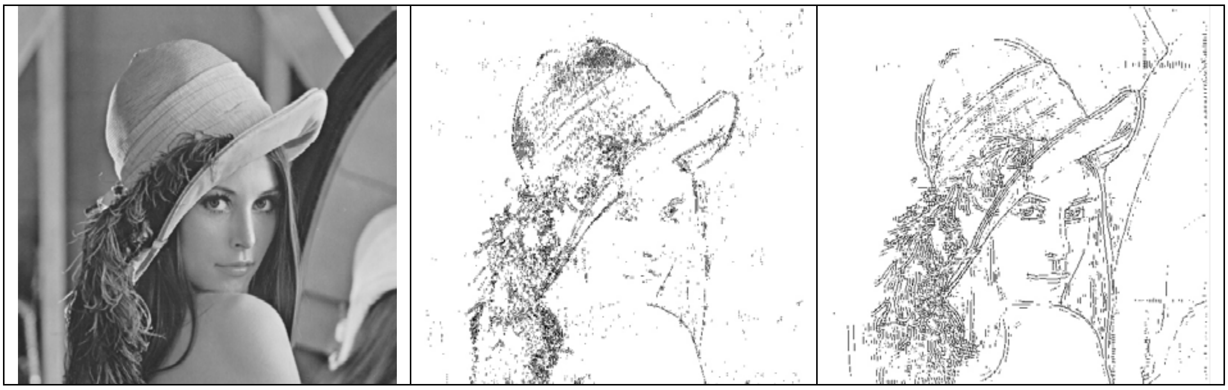


Fig. 1.9 Source image “Lena”; the result of 190-harmonic Fourier-based edge detection method application; and the result of 190 Hermite functions based edge detection method application.

4.0.0. Variational methods of image filtering

The proposed method based on Tikhonov’s regularization was also aimed at solving the edge detection problem. A more general but still very important problem of image filtering can be solved by minimizing the Tikhonov functional of a higher order:

$$E_\alpha(\tilde{u}) = \left\|\tilde{u} - u_\delta\right\|_{L^2}^2 + \beta\left\|\frac{d^2}{dx^2}\tilde{u}\right\|_{L^2}^2 .$$

The first term of the functional is a data fidelity term, while the second rewards smoothness of the solution. This problem can be reduced to solving the Euler equation for u:

$$\begin{cases} u^{(4)} + 4\alpha^4 u = u_\delta \\ u'(1) = u'(-1) = 0 \\ u'''(1) = u'''(-1) = 0 \end{cases} \text{, with the kernel } G_\alpha(x,s) = \begin{cases} \sum\limits_{i=1}^{4} a_i(s)u_i(x) & x \le s \\ \sum\limits_{i=1}^{4} b_i(s)u_i(x) & x \ge s \end{cases} .$$

The following procedure has been developed to find the analytical expression for the kernel. The fundamental solutions look as follows:

$$u_1(x) = e^{\alpha(x+1)}\sin\alpha(x+1),\ u_2(x) = e^{\alpha(x+1)}\cos\alpha(x+1),$$

$$u_3(x) = e^{-\alpha(x-1)}\sin\alpha(x-1),\ u_4(x) = e^{-\alpha(x-1)}\cos\alpha(x-1)$$

After substituting $\xi = \alpha(x+1) \quad \eta = \alpha(x-1)$, their derivatives are given by

$u_1'(x) = \alpha e^{\xi}[\ \ \sin\xi + \cos\xi]$	$u_1''(x) = 2\alpha^2 e^{\xi}[\ \ \ \cos\xi]$	$u_1'''(x) = 2\alpha^3 e^{\xi}[-\sin\xi + \cos\xi]$
$u_2'(x) = \alpha e^{\xi}[-\sin\xi + \cos\xi]$	$u_2''(x) = 2\alpha^2 e^{\xi}[-\sin\xi]$	$u_2'''(x) = 2\alpha^3 e^{\xi}[-\sin\xi - \cos\xi]$
$u_3'(x) = \alpha e^{-\eta}[-\sin\eta + \cos\eta]$	$u_3''(x) = 2\alpha^2 e^{-\eta}[-\cos\eta]$	$u_3'''(x) = 2\alpha^3 e^{-\eta}[\ \ \sin\eta + \cos\eta]]$
$u_4'(x) = \alpha e^{-\eta}[-\sin\eta - \cos\eta]$	$u_4''(x) = 2\alpha^2 e^{-\eta}[\ \ \ \sin\eta]$	$u_4'''(x) = 2\alpha^3 e^{-\eta}[-\sin\eta + \cos\eta]]$

Here, $c_i(s) = b_i(s) - a_i(s)$. From the conjunction conditions

$$\begin{cases} \sum_{i=1}^{4} c_i(s)u_i(s) = 0, \quad \sum_{i=1}^{4} c_i(s)u_i'(s) = 0, \\ \sum_{i=1}^{4} c_i(s)u_i''(s) = 0, \quad \sum_{i=1}^{4} c_i(s)u_i'''(s) = 1 \end{cases}$$

we get

$c_1(s) = \frac{e^{-\alpha(s+1)}}{8\alpha^3}[-\sin\alpha(s+1) + \cos\alpha(s+1)]$	$c_2(s) = \frac{e^{-\alpha(s+1)}}{8\alpha^3}[-\sin\alpha(s+1) - \cos\alpha(s+1)]$
$c_3(s) = \frac{e^{\alpha(s-1)}}{8\alpha^3}[\sin\alpha(s-1) + \cos\alpha(s-1)]$	$c_4(s) = \frac{e^{\alpha(s-1)}}{8\alpha^3}[-\sin\alpha(s-1) + \cos\alpha(s-1)]$

After substituting $G_\alpha(x,s)$ in the boundary conditions

$$\begin{cases} G_\alpha'(-1,s) = G_\alpha'(1,s) = 0, \\ G_\alpha'''(-1,s) = G_\alpha'''(1,s) = 0, \end{cases}$$

we finally obtain the coefficients $b_i(s), a_i(s)$:

$$p = c_1(s) + e^{2\alpha}[c_3(s)\cos 2\alpha + c_4(s)\sin 2\alpha], \quad q = c_2(s) + e^{2\alpha}[c_3(s)\sin 2\alpha - c_4(s)\cos 2\alpha]$$

$$b_1(s) = -\frac{e^{4\alpha}[p\cos 4\alpha + q\sin 4\alpha] - p}{1 - 2e^{4\alpha}\cos 4\alpha + e^{8\alpha}}, \quad b_2(s) = \frac{e^{4\alpha}[p\sin 4\alpha - q\cos 4\alpha] + q}{1 - 2e^{4\alpha}\cos 4\alpha + e^{8\alpha}},$$

$$b_3(s) = e^{2\alpha}[-b_1(s)\cos 2\alpha + b_2(s)\sin 2\alpha], \quad b_4(s) = e^{2\alpha}[\ b_1(s)\sin 2\alpha + b_2(s)\cos 2\alpha]$$

We are flexible to change the boundary conditions for the Euler equation. For the case of the following boundary conditions

$$\begin{cases} G_\alpha(-1,s) = G_\alpha(1,s) = 0 \\ G_\alpha''(-1,s) = G_\alpha''(1,s) = 0 \end{cases},$$

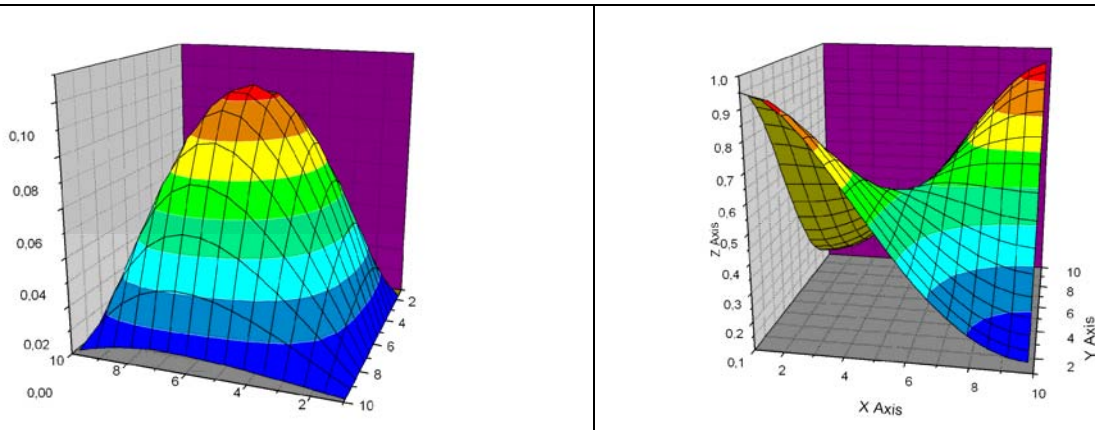
the coefficients $b_i(s), a_i(s)$ are given by expressions:

$$p = c_1(s) - e^{2\alpha}[c_3(s)\cos 2\alpha + c_4(s)\sin 2\alpha], \quad q = c_2(s) - e^{2\alpha}[c_3(s)\sin 2\alpha - c_4(s)\cos 2\alpha]$$

$$b_1(s) = -\frac{e^{4\alpha}[p\cos 4\alpha + q\sin 4\alpha] - p}{1 - 2e^{4\alpha}\cos 4\alpha + e^{8\alpha}}, \quad b_2(s) = \frac{e^{4\alpha}[p\sin 4\alpha - q\cos 4\alpha] + q}{1 - 2e^{4\alpha}\cos 4\alpha + e^{8\alpha}},$$

$$b_3(s) = e^{2\alpha}[b_1(s)\cos 2\alpha - b_2(s)\sin 2\alpha], \quad b_4(s) = -e^{2\alpha}[b_1(s)\sin 2\alpha + b_2(s)\cos 2\alpha]$$

The above versions of the kernel are illustrated in Fig.1.10.



$G_\alpha(x,s)$, $\alpha = 1$, boundary conditions with even derivatives

$G_\alpha(x,s)$, $\alpha = 1$, boundary conditions with odd derivatives

Fig. 1.10 Green's functions

The obtained analytical expressions of Green's functions give us a new tool for solving the problem of image filtering.

The general scheme of the methods based on Tikhonov's regularization using both first and second order terms, which we analyzed separately before, can be posed as a problem of minimizing the Tikhonov functional:

$$E_\alpha(\tilde{u}) = \left\|\tilde{u} - u_\delta\right\|_{L^2}^2 + \alpha\left\|\frac{d}{dx}\tilde{u}\right\|_{L^2}^2 + \beta\left\|\frac{d^2}{dx^2}\tilde{u}\right\|_{L^2}^2 .$$

The method is intended to be also used in the edge detection procedure based on the analysis of the second order zero crossings of the intensity function. It can be expected that the method based on minimizing the first derivative ($\beta = 0$) will suppress noise in the image and the number of the second order zero crossings will be reduced. The method with the second derivative ($\alpha = 0$) is expected to perform stronger smoothing, while preserving the second order zero crossings.

Here we can consider not only self-adjoint statements as it was done for analytical investigations. A special program unit has been designed to solve this problem. We will describe the algorithm for the case

$$\begin{cases} \beta u^{(4)} - \alpha u'' + u = u_\delta, \\ u'(1) = u'(-1) = 0, \quad u''(1) = u''(-1) = 0 \end{cases} .$$

The numerical implementation looks as follows. By approximating the derivatives contained in the equation with finite differences and proceeding to the vector notation $u_i = u(ih)$ (where h is a step size, $i = 0 \ldots n$), we can rewrite the equation as

$$\frac{\beta}{h^4}\left[\left(u_{i-2} - 2u_{i-1} + u_i\right) - 2\left(u_{i-1} - 2u_i + u_{i+1}\right) + \left(u_i - 2u_{i+1} + u_{i+2}\right)\right] + \frac{\alpha}{h^2}\left[u_{i-1} - 2u_i + u_{i+1}\right] + u_i = f_i .$$

The boundary conditions can be rewritten as

$$u_0 - u_1 = 0, \qquad u_0 - 2u_1 + u_2 = 0, \quad u_n - u_{n-1} = 0, \quad u_n - 2u_{n-1} + u_{n-2} = 0 .$$

These boundary conditions can be reduced to $u_0 = u_1 = u_2 \quad u_n = u_{n-1} = u_{n-2}$. By substituting these conditions into the equation, we obtain an n-2-order system of linear equations

$$\begin{cases} \frac{\beta}{h^4}\left[3u_2-4u_3+u_4\right]+\frac{\alpha}{h^2}\left[-u_2+u_3\right]+u_2=f_2 \\ \frac{\beta}{h^4}\left[-3u_2+6u_3-4u_4+u_5\right]+\frac{\alpha}{h^2}\left[u_2-2u_3+u_4\right]+u_3=f_3 \\ \dots \\ \frac{\beta}{h^4}\left[u_{i-2}-4u_{i-1}+6u_i-4u_{i+1}+u_{i+2}\right]+\frac{\alpha}{h^2}\left[u_{i-1}-2u_i+u_{i+1}\right]+u_i=f_i \quad i=4,\dots,n-4 \\ \dots \\ \frac{\beta}{h^4}\left[-3u_{n-2}+6u_{n-3}-4u_{n-4}+u_{n-5}\right]+\frac{\alpha}{h^2}\left[u_{n-2}-2u_{n-3}+u_{n-4}\right]+u_{n-3}=f_{n-3} \\ \frac{\beta}{h^4}\left[3u_{n-2}-4u_{n-3}+u_{n-4}\right]+\frac{\alpha}{h^2}\left[-u_{n-2}+u_{n-3}\right]+u_{n-2}=f_{n-2} \end{cases}$$

This system can be rewritten in a compact matrix form as $Au=f$.

Further analysis has been performed to compare the influence of the first and the second order term in the Tikhonov functional. The scheme of the analysis was as follows. For a set of test images, we added a uniformly distributed noise (an example of noisy data is given below):

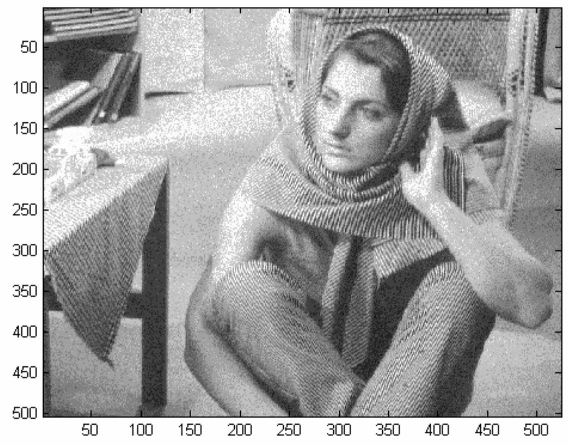


Fig. 1.11 Noisy test image.

Then, we analyzed the value of the PSNR metric between the initial image and the smoothed image obtained as a function of the regularization parameter α (for the case of the pure first order Tikhonov functional) or the regularization parameter β (for the case of the pure second order Tikhonov functional). It was found that the obtained maximal values of PSNR are very close in all the cases.

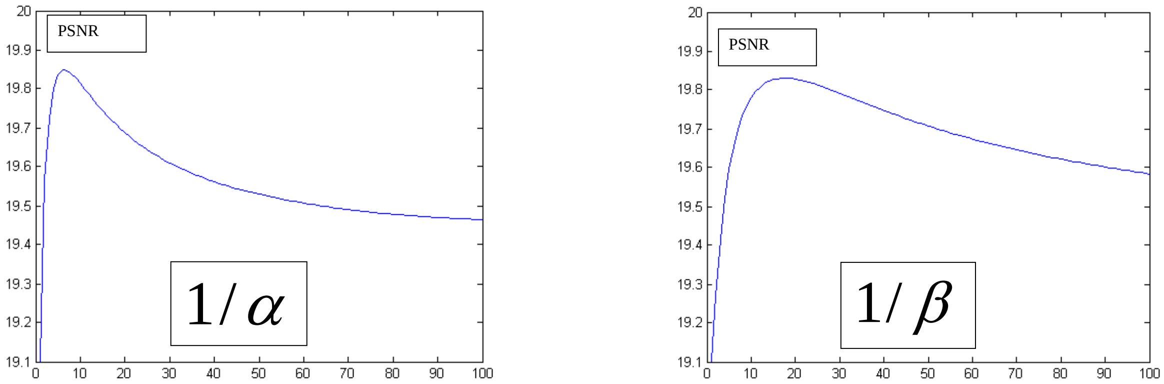


Fig. 1.12 PSNR values as functions of regularization parameters.

Visual analysis of the best-PSNR results (Fig. 1.13) was performed at the final stage to compare features of the first and second order Tikhonov regularization schemes.

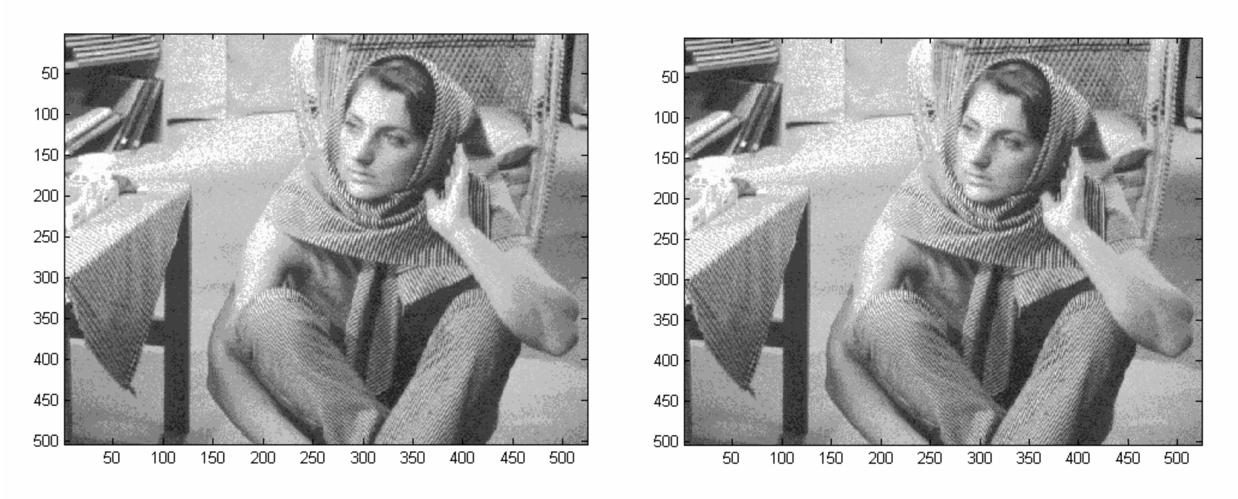


Fig. 1.13 First order (left) and second order (right) Tikhonov regularization filtering results.

The analysis showed that the first order Tikhonov regularization filtering works better for lower frequencies, while the second order Tikhonov regularization filtering is preferable to treat high-frequency details.

1.0. Controlling Discrete Fourier Filtering.

2.2.1. Model of discrete Fourier filtering in nonlinear optical systems

The theoretical background of optimal Fourier filtering has been developed and applied to the problem of producing phase-amplitude Fourier filters aimed at forming a desired phase distribution in a light wave by means of a nonlinear optical system with such a Fourier filter in the feedback loop (see. Fig.2.0).

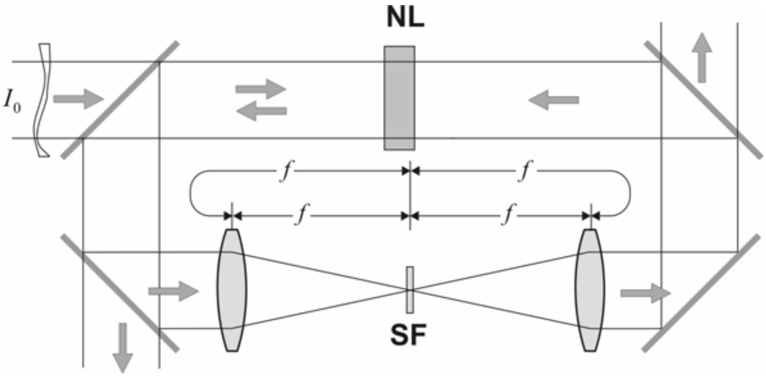


Fig. 2.0. Conceptual scheme of a nonlinear optical system with a spatial filter in the feedback loop.

Here, a spatial filter, SF, is placed in the Fourier plane of a «4-*f*» system that produces a feedback signal to drive a nonlinear medium, NL. In the framework of a single-pass approximation, the dynamics of the above system is governed by a Debye-type nonlinear parabolic PDE. This equation governs the additional phase modulation, $u = u(\mathbf{x}, t)$, of a light wave passed through a thin layer of a Kerr-type nonlinear medium being a part of an optical feedback system with a Fourier filter in its feedback loop:

$$\begin{gathered}
\tau \frac{\partial u}{\partial t} + u - l_d^2 \Delta_\perp u = \Phi(u,T), \quad t \in [0, t_1], \quad u(\mathbf{x}, t=0) = u_0(\mathbf{x}), \\
\Phi(u,T) = K\delta_1 \, | A_{in} |^2 + 2K\delta_2 \operatorname{Re}(\overline{A_{in}} \hat{F}_T \{ A_{in} \exp(iu) \}) + K\delta_3 \left| \hat{F}_T \{ A_{in} \exp(iu) \} \right|^2 .
\end{gathered} \tag{2.1}$$

Here, τ is the characteristic response time of the nonlinear medium; l_d is the diffusion length of the medium; $\Delta_\perp = \nabla_\perp^2$ is the transverse Laplace operator with respect to transverse spatial variable $\mathbf{x} = (x_1, x_2) \in \Omega$, where Ω is the optical system aperture, being a bounded convex 2D domain.

Equation (2.1) is supplemented with homogeneous boundary conditions of the Dirichlet, Neumann, or periodical type (the latter type is valid for rectangular domains only). Since the Laplace operator with the mentioned boundary conditions allows only a discrete spectrum, the traditional convolution integral of the Fourier transform should be replaced with summing over a discrete index. More precisely, for a complex-valued function $A = A(\mathbf{x})$ we denote its direct discrete Fourier transform

$$\hat{F}\{A\} = \{a\}_{n=1}^{+\infty}, \ \{a\} = (a_1, a_2, ..., a_n, ...), \qquad a_n = \| e_n \|_{L^2(\Omega)}^{-1} \int_\Omega A(\mathbf{x}) \cdot e_n(\mathbf{x}) d\mathbf{x},$$

as a sequence of Fourier coefficients regarding to eigenfunctions $\{ e_n(\mathbf{x}) \}_{n=1}^{+\infty}$ of the Laplace operator meeting the corresponding boundary conditions: $\Delta_\perp e_n(\mathbf{x}) + \lambda_n e_n(\mathbf{x}) = 0$. Similarly, $\hat{F}^{-1}$ means the inverse discrete Fourier transform, which calculates the sum $\hat{F}^{-1}\{a\} = \sum\limits_{n=1}^{+\infty} a_n e_n(\mathbf{x})$. In such a way, we denote the discrete transfer function $T = (\rho_1, \rho_2, ..., \rho_n, ...)$ as a sequence of complex-valued multipliers, which directly affect the corresponding components of the Fourier spectrum. In our terms, the procedure of the discrete Fourier filtering of the function $A(\mathbf{x})$ consists of calculating its discrete Fourier transform $\{a\} = \hat{F}\{A\}$, component-wise multiplying with the transfer function,

$T \cdot a = (\rho_1 a_1, \rho_2 a_2, ..., \rho_n a_n, ...)$, followed by taking the inverse discrete Fourier transform. As a result, we come to the definition of discrete Fourier filtering in the form of the superposition $\hat{F}_T\{A\} = \hat{F}^{-1}\left\{T \cdot \hat{F}\{A\}\right\}$ that is placed into the right-hand side of equation (2.1).

Real-valued parameters $\delta_1, \delta_2, \delta_3$ contained in the right-hand side of Eq. (2.1) are responsible for the specific arrangement of the nonlinear optical system. Namely, equation (2.1) models the dynamics of optical systems both with summation of the input and the feedback light field and with separation of these fields. In the former case, the equation allows either taking into account the interference of these fields or neglecting it. Moreover, the choice of these parameters' values defines the type of the considered nonlinear medium (focusing or defocusing). For example, the choice $\delta_1 = \delta_2 = \delta_3 = 1$ allows us to simulate an optical system with coherent interaction of the input and feedback fields (the interference "is on"), while the choice $\delta_1 = \delta_3 = 1$, $\delta_2 = 0$ allows us to neglect the interference-type interactions. Should we set $\delta_1 = \delta_2 = 0, \delta_3 = -1$, then (1) will simulate an optical system with separated input and feedback fields in the case of a defocusing nonlinear medium. Feedback factor K is proportional to the intensity of the input light wave.

The first step in the mathematical study of the optimization problem is to prove the existence of the solution of state equation (2.1) in the appropriate functional space for all the admissible filters. To formulate the existence and uniqueness theorem, we denote by ℓ_∞ a space of complex-valued bounded sequences $T = (\rho_1, \rho_2, ..., \rho_n, ...)$ with the finite norm $\|T\|_{\ell_\infty} = \sup\limits_{n=1,2,...} |\rho_n|$. Using ℓ_∞ as the admissible set allows one to consider a wide range of Fourier filters with no restrictions of smallness to the amplitudes of high spatial frequency multipliers to be supposed.

Theorem 2.1. Let the initial condition $u_0 \in L^2(\Omega)$ and the input light wave $A_{in} \in H^1(\Omega) \cap C(\overline{\Omega})$. Then:

1) For any discrete filter, $T \in \ell_\infty$, on any finite time interval, $0 < t < t_1$, the initial-boundary value problem for equation (2.1) with homogeneous Dirichlet or Neumann-type boundary conditions or periodic boundary conditions has the unique energy class solution $u(\mathbf{x},t) \in L^2(0, t_1; H^1(\Omega)) \cap C([0,t_1]; L^2(\Omega))$. The solution satisfies the initial condition in the sense of traces, the equation (2.1) is realized in the sense of distributions.
2) The solution obeys the estimation $\|u\|_{C(0,t_1;L^2(\Omega))} + \|u\|_{L^2(0,t_1;H^1(\Omega))} \le C(t_1)(1 + \|u_0\|_{L^2(\Omega)})$, with the norms $\|u\|_{C(0,t_1;L^2(\Omega))} = \max\limits_{t \in [0,t_1]} \|u(t)\|_{L^2(\Omega)}$, $\|u\|_{L^2(0,t_1;H^1(\Omega))} = (\int\limits_0^{t_1} (\|u(t)\|^2_{L^2(\Omega)} + \|\nabla_\perp u(t)\|^2_{L^2(\Omega)}) dt)^{1/2}$, where the constant $C(t_1) > 0$ is independent of u_0.
3) The solution Lipschitz-continuously depends on the initial function and the Fourier filter, and the estimation
$$\|u - \tilde{u}\|_{C(0,t_1;L^2(\Omega))} + \|u - \tilde{u}\|_{L^2(0,t_1;H^1(\Omega))} \le C_2(\|u_0 - \tilde{u}_0\|_{L^2(\Omega)} + \|T - \tilde{T}\|_{\ell_\infty})$$
holds for the corresponding solutions $u = u(\mathbf{x}, t; u_0, T)$, $\tilde{u} = u(\mathbf{x}, t; \tilde{u}_0, \tilde{T})$.
4) Every bounded set of initial data is attracted to some set M (global attractor), which is compact in $L^2(\Omega)$ and has a finite Hausdorff dimension.

2.2.2. Statement and solvability of the problem of optimal Fourier filtering.

The optimization problem has been formulated as follows: let the Fourier filter take the form $T_\infty + T$, where T_∞ is the constant part of the filter (for instance, $T_\infty = 0$), and T appears as a variable one. The question is how the choice of the specific discrete Fourier filter T in equation (2.1) allows one to obtain at a given point of time, $t = t_1$, a solution, $u(\mathbf{x}, t = t_1; u_0, T_\infty + T)$, close to the desired state, $u_1(\mathbf{x})$. In the notation $u(\mathbf{x}, t; u_0, T_\infty + T)$ for the solution of problem (2.1), we underline its implicit dependence on the initial phase modulation $u_0(\mathbf{x})$ and the Fourier filter $T_\infty + T$. In the studies of this optimization problem, some new developments have been achieved.

Let us consider a linear bounded operator $D : L^2(\Omega) \to L^2(\Omega)$ and formulate a new minimization problem for a family of terminal functionals

$$J(T) = \int\limits_{\Omega} \left| D\big(u(\mathbf{x}, t = t_1; u_0, T_\infty + T) - u_1(\mathbf{x})\big) \right|^2 d\mathbf{x} \to \inf, \quad T \in \Xi, \tag{2.2}$$

with the admissible set Ξ, which takes into account the technical restrictions for the filters considered. In compare with the “classical” terminal target functional $J_1(T) = \int\limits_{\Omega} | u(\mathbf{x}, t = t_1; u_0, T_\infty + T) - u_1(\mathbf{x}) |^2 \, d\mathbf{x}$, the operator D is responsible for additional features of approximation to the target state for each point of domain Ω. In particular, by choosing the following operator

$$D(v) = \chi_{\Omega'}(\mathbf{x})\Big(v(\mathbf{x}) - |\Omega'|^{-1} \int_{\Omega} \chi_{\Omega'}(\mathbf{y}) \cdot v(y) d\mathbf{y}\Big),$$

we can embrace a practically important case of the numerically studied problem of the proximity of the sum of the input wave phase, $\varphi(\mathbf{x})$, and the additional nonlinear phase modulation, $u(\mathbf{x}, t)$, to the given phase distribution, $u_1(\mathbf{x})$, in the bounds of the controlled subdomain Ω' at the time point $t = t_1$:

$$J\ (T) =$$

$$= \int\limits_{\Omega} \left| \chi_{\Omega'}(\mathbf{x})\Big(u(\mathbf{x}; t_1; T_\infty + T) + \varphi(\mathbf{x}) - u_1(\mathbf{x}) - |\Omega'|^{-1} \int_{\Omega} \chi_{\Omega'}(\mathbf{y})\big(u(\mathbf{y}; t_1; T_\infty + T) + \varphi(\mathbf{y}) - u_1(\mathbf{y})\big) d\mathbf{y}\Big) \right|^2 d\mathbf{x}$$

where $\chi_{\Omega'}(\mathbf{x}) = \begin{cases} 1, \text{ if } \ \mathbf{x} \in \Omega', \\ 0, \text{ if } \ \mathbf{x} \notin \Omega' \end{cases}$, $|\Omega'|$ being the area of the subdomain Ω'.

It is important to describe possible admissible sets Ξ for which the optimization problem (2.1)-(2.2) is solvable, i.e. at least one optimal filter T^* satisfying the equality $J(T^*) = \inf\limits_{T \in \Xi} J(T)$ does exist. The following assertions on the solvability of the optimization task have been proven.

Theorem 2.2. Let $A_{in} \in C^1(\overline{\Omega})$, the operator D be a linear and continuous operator in $L^2(\Omega)$, and the constant part of the Fourier filter $T_\infty \in \ell_\infty$ be fixed.

1) If $\Xi = H_R$, where $H_R = \{T = (\rho_1, \rho_2, ...) \in \ell_2 : | \rho_k | \le r_k k^{-1}, 0 \le r_k \le R, k = 1, 2, ..., R > 0\}$, then, the optimal filter set H_R^* for problem (2.1)-(2.2) is not empty, $H_R^* = \{T^* \in H_R : J(T^*) = \inf\limits_{T \in H_R} J(T)\} \ne \varnothing$. Moreover, the optimal filter set H_R^* is compact, and any sequence minimizing terminal functional (2.2) converges to this set.

2) If $\Xi = B_R$, where $B_R = \{T = (\rho_1, \rho_2, ...) \in \ell_2 : \| T \|_{\ell_2}^2 \equiv \sum\limits_{n=1}^{\infty} | \rho_n |^2 \le R^2\}$, then, the optimal filter set B_R^* for problem (2.1)-(2.2) is not empty, $B_R^* = \{T^* \in B_R : J(T^*) = \inf\limits_{T \in B_R} J(T)\} \ne \varnothing$. Moreover, the optimal filter set B_R^* is weakly compact in the space ℓ_2.

3) For the gradient $J'(T)$, the following formula is valid:

$$\langle J'(T),\Delta T\rangle = \tau^{-1}\int_0^{t_1}\int_{\Omega}\left\langle \Phi_T(u(\mathbf{x},t),T_\infty+T),\Delta T\right\rangle_{\ell_2}\psi(\mathbf{x},t)d\mathbf{x}dt\,, \qquad (2.3)$$

where $\left\langle \Phi_T,\Delta T\right\rangle = 2K\,\mathrm{Re}\Big((\delta_2\overline{A}_{in}+\delta_3\overline{\hat{F}}_T\{A_{in}\exp(iu)\})\cdot\hat{F}_{\Delta T}\{A_{in}\exp(iu)\}\Big)$, and $\psi(r,t)$ is the solution of the conjugate problem

$$\begin{aligned}
&-\tau\frac{\partial\psi}{\partial t}+\psi-l_d^2\Delta_\perp\psi=\langle\Phi_u^*,\psi\rangle,\quad t\in[0,t_1],\\
&\psi(\mathbf{x},t=t_1)=2D^*D\big(u(\mathbf{x},t=t_1;u_0,T_\infty+T)-u_1(\mathbf{x})\big),
\end{aligned} \qquad (2.4)$$

$$\begin{aligned}
&\left\langle\Phi_u^*(u,T),\psi\right\rangle=2K\delta_2\,\mathrm{Re}\Big(A_{in}\cdot i\cdot\exp(iu)\overline{\hat{F}}_{\overline{T}}\{A_{in}\overline{\psi}\}\Big)+\\
&+2K\delta_3\,\mathrm{Re}\Big(A_{in}\cdot i\cdot\exp(iu)\overline{\hat{F}}_{\overline{T}}(\hat{F}_T\{A_{in}\exp(iu)\}\overline{\psi})\Big).
\end{aligned} \qquad (2.5)$$

The boundary conditions coincide with those of the direct problem (2.1).

4) Within any finite time interval, $0<t<t_1$, initial-boundary value problem (2.4)-(2.5) with homogeneous boundary conditions of the Dirichlet, Neumann, or periodical type has the unique solution $\psi\in L^2(0,t_1;H^1(\Omega))\cap C([0,t_1];L^2(\Omega))$, and this solution obeys the estimation

$$\|\psi\|_{C(0,t_1;L^2(\Omega))}+\|\psi\|_{L^2(0,t_1;H^1(\Omega))}\le C\ (1+\|\psi(t=t_1)\|_{L^2(\Omega)}),$$

where the constant $C>0$ is independent of u_1.

2.2.4. Projection finite-element approximations

To solve direct problem (2.1) with periodic boundary conditions set in a rectangle, a rectangular Cartesian coordinate system is usually chosen. An implicit projection-difference scheme based on the finite element method has been proposed and developed. We used space $S_1=Lin\{\varphi_{i,j}(\mathbf{x})\}$ of piecewise continuous linear functions (first-order B-splines) $\varphi_{i,j}(\mathbf{x})=\varphi_{i,j}(x_1,x_2)=\varphi_i(x_1)\cdot\varphi_j(x_2)$ built in accordance with the formulas

$$\begin{aligned}
&\varphi_n(x_j)=\varphi_0(x_j-nh_j),\ \varphi_N(x_j)=\varphi_0(x_j)+\varphi_0(x_j-2\pi),\\
&\varphi_0(x_j)=\max\{0,1-|x_j|h_j^{-1}\},\ n=1,\ldots,N_j-1,\ j=1,2\,.
\end{aligned}$$

Then, an approximate solution was being sought for, as the element $w=(w^0,\ldots,w^M)\in S$ of the space $S=\underbrace{S_1\times S_1\times S_1\times\ldots\times S_1}_{M+1\text{ times}}$ satisfying the projection-difference scheme:

$$\left\langle\tau^{-1}(w^m-w^{m-1}),\varphi\right\rangle+\left\langle w^m,\varphi\right\rangle+l_d^2\left\langle\nabla_\perp w^m,\nabla_\perp\varphi\right\rangle=\left\langle\Phi(w^m),\varphi\right\rangle,\ m=1,\ldots,M\ ,\forall\varphi\in S_1, \qquad (2.6)$$

where, as before, the operator Φ stays for the right-hand side of (2.1). The angular brackets in (2.6) denote inner products in the Hilbert space $L^2(\Omega)$. In the software developed, we used a more convenient matrix form of Eq. (2.6):

$$((1+\tau)B+\tau D\Lambda)w^m=\tau\,\Phi(w^m)+Bw^{m-1},\ m=1,\ldots,M\ , \qquad (2.7)$$

where $\Lambda=B_2\Lambda_1+B_1\Lambda_2$, $B=B_1B_2$, $B_l=E-h^2\Lambda_l/6$, E is the unit matrix, and

$$(\Lambda_l f)_i=-\frac{1}{h_l^2}(f_{i+1}-2f_i+f_{i-1}),\ f_{N_l+1}=f_1,\ f_{N_l}=f_0,\ i=0,\ldots,N_l-1,\ l=1,2\,.$$

The above scheme (2.7) is implicit. Therefore, it requires iterations in order to obtain the solution at a new temporal layer. Usually, 3-4 iterations per each temporal layer are enough to achieve the required accuracy.

The same approach is applied while constructing approximations to conjugate problem (2.4). Then, an approximate solution of the problem (2.4) is being sought for, as the element

$w = (w^0, ..., w^M) \in S$ satisfying a projection-difference scheme analogous to (2.6). The corresponding matrix notation looks as follows:

$$((1+\tau)B + \tau D\Lambda)w^{m-1} = \tau G(w^{m-1}) + Bw^m, \; m = M, ..., 1. \quad (2.8)$$

Scheme (2.8) is being solved along a decreasing sequence of indices. Therefore, this scheme is implicit. The choice of an implicit scheme is due to its greater stability margin with respect to explicit schemes.

2.2.5. Tikhonov regularization method

A number of computer experiments were performed with the aim of minimizing target functionals of the considered class by means of the conditional gradient method. On this way, we noticed that after a certain number of steps further iterations did not considerably decrease the functional value, while the structure of the optimal filter became more and more complex (due to a larger number of high-frequency harmonics to transmit). Such an unjustified complication of the filter structure is not only inconvenient in practice but also indicates certain nonrobustness of the optimization process. To make the optimization robust, we used Tikhonov's regularization, which had been earlier applied by the Team members to abstract extremum problems (A.N. Tikhonov and F.P. Vasiliev // Banach Center Publications. Vol. 3. Mathematical models and numerical methods. Warszawa, 1978, pp. 297-342; F.P. Vasiliev and M. Yachimovich // Soviet Mathematics Doklady, Vol. 250, No 2, 1980).

We have implemented a version of the Tikhonov regularization method for the problem of minimization of the functional $J(T)$, which implies constructing a family of Tikhonov functionals

$$J_k(T) = J(T) + \alpha_k \|T\|^2_{\ell_2},$$

dependent on the regularization parameters $\{\alpha_k\}_{k=1}^{+\infty}$. Then, on the basis of the expressions for the gradients, $J_k'(T) = J'(T) + 2\alpha_k T$, for every $k = 1,2,\ldots$ one step of the conditional gradient method was made for the functional $J_k(T)$ in the set H_R. As a result, a sequence of optimized filters, $\{T_k\}$, was obtained. In the theory of regularization of the conditional gradient method, the convergence of the iterations to the set of optimal controls can be proven under rather restricting conditions, such as convexity of the target functional, presence of a saddle node for the Lagrange function, etc. For the Fourier filtering models we consider, these conditions are most likely not met. However, this fact does not prevent us from studying possibilities of practical application of iterative regularization methods and robust construction of a sequence of optimized filters. The properties of such a sequence strongly depend on the choice of the regularization parameters. These properties were numerically studied for Tikhonov functionals of the two following types.

In the first case, all the regularization parameters were equal: $\{\alpha_k\}_{k=1}^{+\infty} = \overline{\alpha} = const > 0$. As a result, we proceed to so-called static regularization that implies adding to the target functional a stabilizing increment, $\overline{\alpha}\|T\|^2_{\ell_2}$, independent of the iteration number, which prevents the filter norm from increasing during the optimization process.

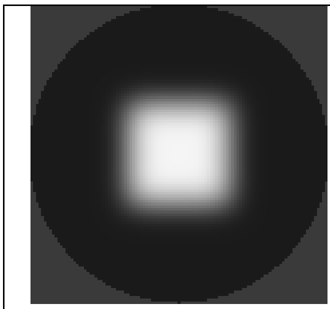
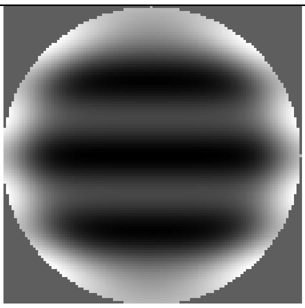
In the second case, we *a priori* constructed a power-function sequence of the regularization parameters, $\{\alpha_k\}_{k=1}^{+\infty} = C_0(k+1)^{-\beta}$, $C_0 > 0, \beta > 0$, which was often used in regularization methods. As a result, we obtained a particular case of the regularized method of conditional gradient.

In both cases, the key question was how many harmonics of the discrete Fourier filter should be controlled in order to achieve the desirable result with acceptable accuracy. Also, we compared our numerical data with the results obtained in the case of no regularization, that is, when $\{\alpha_k\}_{k=1}^{+\infty} \equiv 0$. Taking one group of such experiments:

$$T_\infty = (\underbrace{0,...,0}_{N-times}, -i, -i, ...), \; H_R = \{T = (\rho_1, \rho_2, ...) : |\rho_k| \le 1, k = 1,2,...,N; \rho_k \equiv 0, k \ge N+1\},$$

as an example, let us describe the general regularities revealed.

For the phase inhomogeneity of the input light beam having the form of $\varphi(\mathbf{x}) = 0.5\cos(9x_2)$, we took the target function $u_1(\mathbf{x})$ shown in **Fig 2.1a**. **Fig 2.1b** demonstrates the total phase profile, $u(\mathbf{x}, t_1; T) + \varphi(\mathbf{x})$, before the optimization is started.

 **Fig 2.1a.** Target phase profile $u_1(\mathbf{x})$.	 **Fig 2.1b.** Total phase profile $u(\mathbf{x}, t_1; T) + \varphi(\mathbf{x})$ before the optimization.

In the case of controlling a limited number of the filter components, which are inside a desired-radius circle on the Fourier plane, close results were obtained both for iterative regularization with parameters $C_0 = 1.52588\text{e-}5, \beta = 0.45$ (**Fig 2.2**) and in the experiment with no regularization, that is, for $\{\alpha_k\}_{k=1}^{+\infty} \equiv 0$ (**Fig 2.3**).

In the experiment with static regularization with parameters $\{\alpha_k\}_{k=1}^{+\infty} = \overline{\alpha} = 1.52588\text{e-}5$, the approximation (see **Fig 2.4a**) appeared to be less successful with respect to those shown in **Fig 2.3a** and **Fig 2.2a**. This can be explained by the fact that the resultant optimized control shown in **Fig 2.4b** has a smaller norm as compared with the controls obtained in the previous experiments (**Fig 2.2b** and **Fig 2.3b**).

Similar regularities were revealed during a series of computer experiments where all the available components of the discrete Fourier filter were controlled. In these experiments, good approximations to the target function $u_1(\mathbf{x})$ were obtained both with iterative regularization with parameters $C_0 = 1.52588\text{e-}6, \beta = 0.25$ (**Fig 2.5**) and with no regularization, that is, for $\{\alpha_k\}_{k=1}^{+\infty} \equiv 0$ (**Fig 2.6**). Nevertheless, these experiments qualitatively differ from the experiments with a limited number of controlled harmonics. This difference is in the fact that the iterative regularization performed for all the controlled filter components allowed us to clearly demonstrate the possibility of obtaining a good approximation with simultaneously decreasing the number of controlled harmonics (as compared with the experiment, in which no regularization was performed). In this case, it becomes possible not to control a certain number of harmonics belonging to peripheral portions of the filter (in **Fig 2.7**, these harmonics, which are ignored in the filter shown in **Fig 2.5b** with respect to the filter shown in **Fig 2.6b**, are red-highlighted). This fact emphasizes a positive effect of Tikhonov's iterative regularization for obtaining both efficient and non-complex Fourier filters.

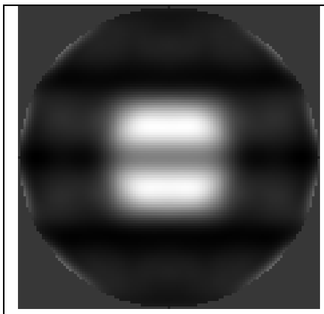


Fig 2.2a. Total phase profile $u(\mathbf{x}, t_1; T) + \varphi(\mathbf{x})$ after optimization with iterative regularization

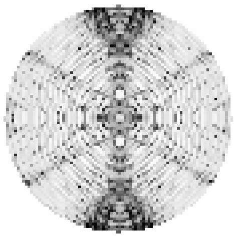


Fig 2.2b. Modulus of the discrete filter T .

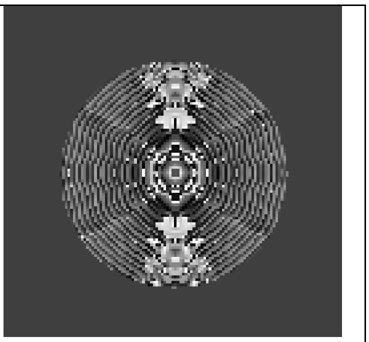


Fig 2.2c. Argument of the discrete filter T .

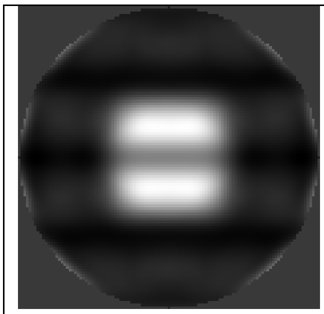


Fig 2.3a. Total phase profile $u(\mathbf{x}, t_1; T) + \varphi(\mathbf{x})$ after optimization without regularization

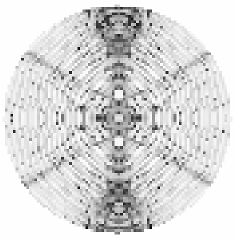


Fig 2.3b. Modulus of the discrete filter T .

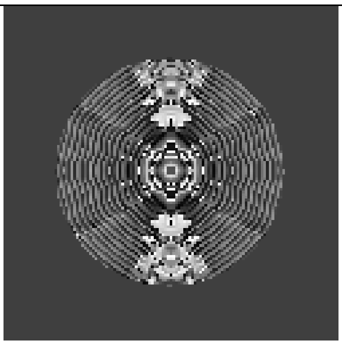


Fig 2.3c. Argument of the discrete filter T .

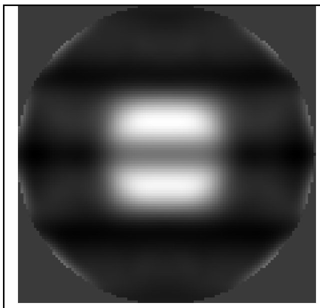


Fig 2.4a. Total phase profile $u(\mathbf{x}, t_1; T) + \varphi(\mathbf{x})$ after optimization with static regularization

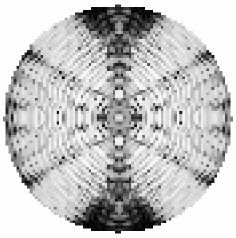


Fig 2.4b. Modulus of the discrete filter T .

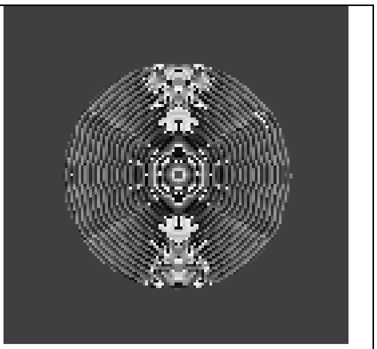
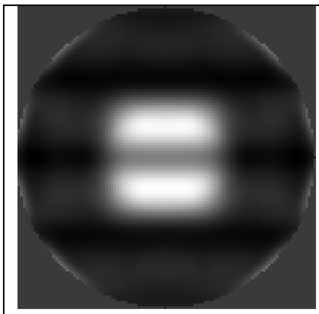
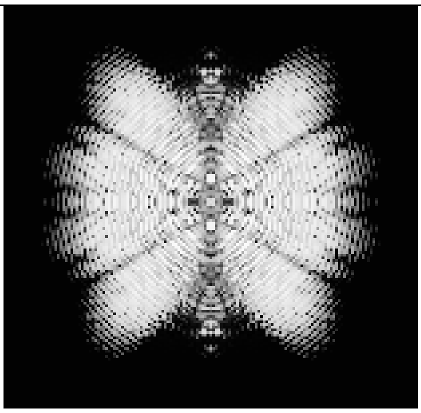
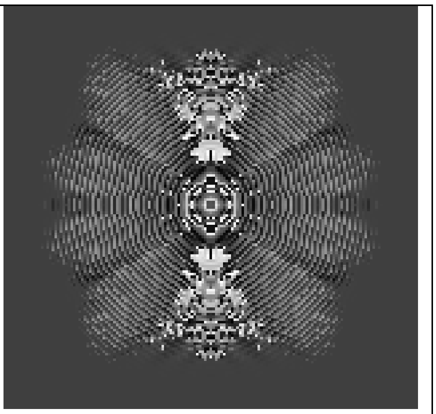
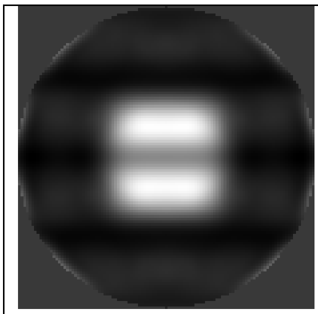
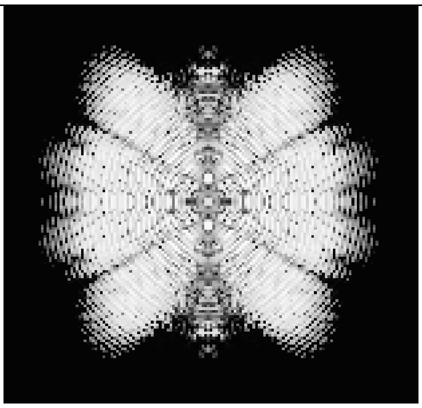
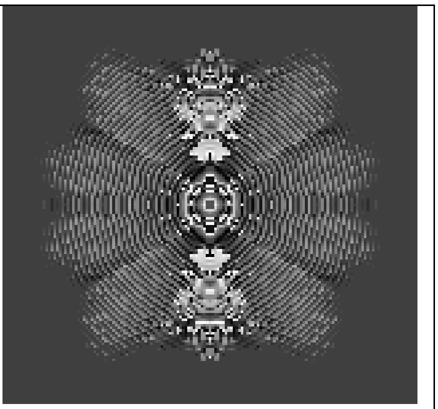


Fig 2.4c. Argument of the discrete filter T .

 **Fig 2.5a.** Total phase profile $u(\mathbf{x},t_1;T)+\varphi(\mathbf{x})$ after optimization with iterative regularization	 **Fig 2.5b.** Modulus of the discrete filter T .	 **Fig 2.5c.** Argument of the discrete filter T .

 **Fig 2.6a.** Total phase profile $u(\mathbf{x},t_1;T)+\varphi(\mathbf{x})$ after optimization without regularization	 **Fig 2.6b.** Modulus of the discrete filter T .	 **Fig 2.6c.** Argument of the discrete filter T .

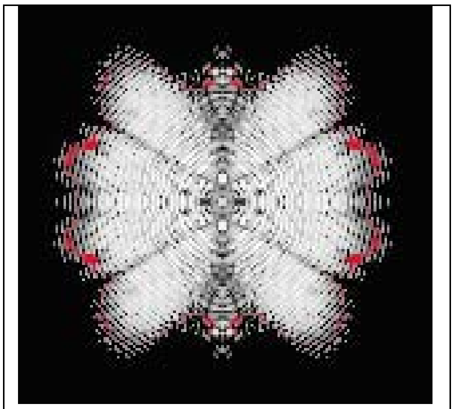


Fig 2.7. Red points mark those harmonics processed by the filter shown in **Fig 2.6b** and ignored in the filter shown in **Fig 2.5b**

2.2.6. Dependence of the optimization quality on the number of controlled harmonics

Some new statements were considered for the problem of forming a given phase distribution in a light wave by means of optical feedback systems with Fourier filters, which correspond to model (2.1) for the following values of right-hand side parameters: $\delta_1 = \delta_2 = 0$, $\delta_3 = -1$. The optimization quality was estimated by the following functional:

$$J^k(T) =$$

$$= \int_\Omega \left| \theta(\mathbf{x}) \left(u(\mathbf{x};t_1;T_\infty + T) + \varphi(\mathbf{x}) - u_1(\mathbf{x}) - \frac{\int_\Omega \theta(\mathbf{y})\big(u(\mathbf{y};t_1;T_\infty + T) + \varphi(\mathbf{y}) - u_1(\mathbf{y})\big)d\mathbf{y}}{\int_\Omega \theta(\mathbf{y})d\mathbf{y}} \right) \right|^2 d\mathbf{x} +$$

$$+ \alpha_k \|T\|^2_{\ell_2} .$$

In the numerical simulation series we performed, the following parameters of iterative regularization were chosen: $C_0 = 0.833N^{-1}, \beta = 0.45$, where N is the number of controlled harmonics, and the sequence $\{\alpha_k\}_{k=1}^{+\infty}$ was constructed according to the formula $\{\alpha_k\}_{k=1}^{+\infty} = C_0(k+1)^{-\beta}$. The goal of the numerical simulations was to figure out the number of controlled harmonics, enough to achieve a good quality of forming the desired phase distribution. The key parameters were the following: R, the radius of a circle centered with respect to the Fourier plane, which defined the number of "active" harmonics, $N(R)$ (all the harmonics being inside the circle). The two additional quantities, $\eta(\varphi)$ and $\eta(u_1)$, we used to explain the results obtained, were the fractions of the power spectra of (respectively) the input field $A_{in}(\mathbf{x}) = |A_{in}(\mathbf{x})| \exp\{i\varphi(\mathbf{x})\}$ and the target field $|A_{in}(\mathbf{x})| \exp\{iu_1(\mathbf{x})\}$, which lay within the controlled zone (the circle of the radius *R*). The control quality was estimated with the use of a smooth weight function, $\theta(\mathbf{x}) \in C^{\infty}(\Omega)$. Taking one group of such simulations as an example, let us describe the regularities revealed.

For the phase inhomogeneity of the input light beam having the form of $\varphi(\mathbf{x}) = 0.5\cos(16x_2)$, we took the target function $u_1(\mathbf{x}) = \exp\left(-5\left((x_1-\pi)^4 + (x_2-\pi)^4\right)\right)$ shown in **Fig 2.8a.**

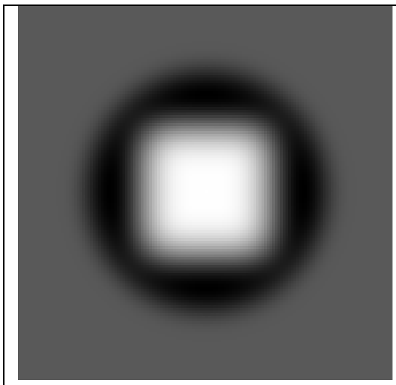


Fig 2.8a. Target phase profile $u_1(\mathbf{x})$.

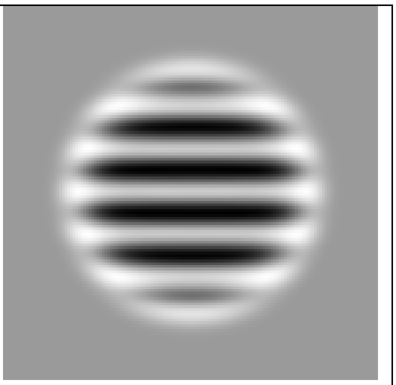


Fig 2.8b. Total phase profile $u(\mathbf{x}, t_1; T) + \varphi(\mathbf{x})$ before the optimization.

For those simulations where the controlled zone radius, R, was large enough to fulfill the conditions $\eta(\varphi) > 0.99$ and $\eta(u_1) > 0.99$, the optimization results were the best out of the series. In this case, the number of controlled harmonics were, of course, quite large, while the rest portion of the Fourier filter was just opaque (**Fig 2.9**). It can be seen from **Fig 2.9b** that only those harmonics belonging either to the spectrum of the input wave or to that of the target profile take part in the optimization process. This feature is important, because it may help with adaptively choosing the structure of the controlled set of Fourier filters for any specific input data.

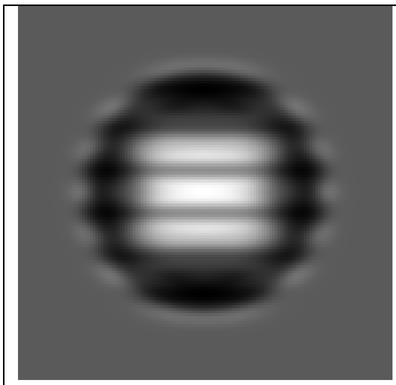


Fig 2.9a. Total phase profile $u(\mathbf{x}, t_1; T) + \varphi(\mathbf{x})$ after the optimization. $R = 50$.

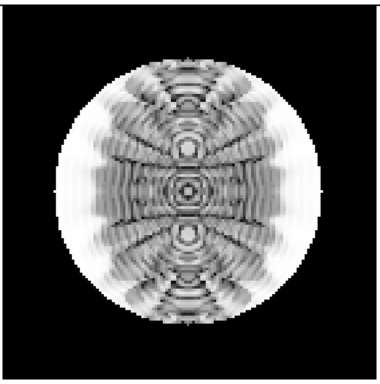


Fig 2.9b. Modulus of the discrete filter T. $N(R) = 7845$.

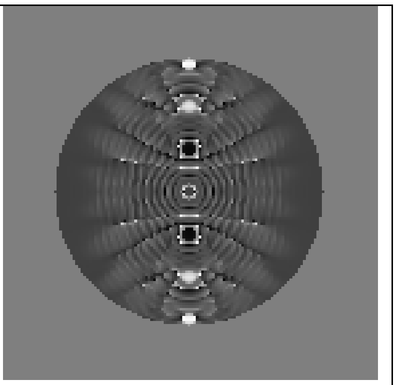


Fig 2.9c. Argument of the discrete filter T. $N(R) = 7845$.

When decreasing R, we revealed the following regularity. In those simulations where the controlled zone radius were still large enough to fulfill the conditions $\eta(\varphi) > 0.92 \div 0.93$ and $\eta(u_1) > 0.92 \div 0.93$, the optimization result remained rather good (**Fig 2.10**), though distinctly conceding to those simulations where $\eta(\varphi) > 0.99$ and $\eta(u_1) > 0.99$ (**Fig 2.9**). When we decreased R in such a way that to fulfill $\eta(\varphi) < 0.90$ or $\eta(u_1) < 0.90$, the optimization quality considerably degraded with respect to the previous simulation series (compare **Fig 2.11** and **Fig 2.9-Fig 2.10**).

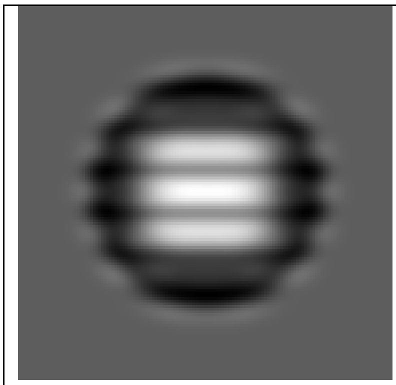


Fig 2.10a. Total phase profile $u(\mathbf{x}, t_1; T) + \varphi(\mathbf{x})$ after the optimization. $R = 19$.

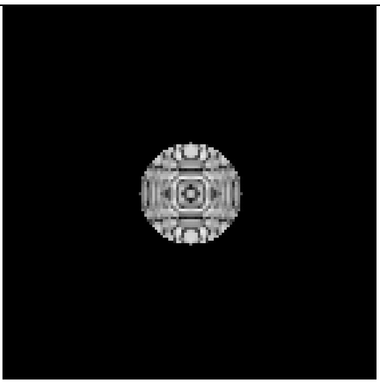


Fig 2.10b. Modulus of the discrete filter T. $N(R) = 1129$.

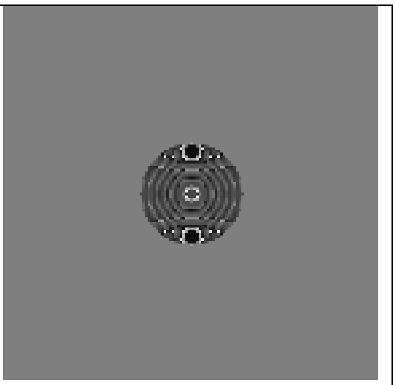


Fig 2.10c. Argument of the discrete filter T. $N(R) = 1129$.

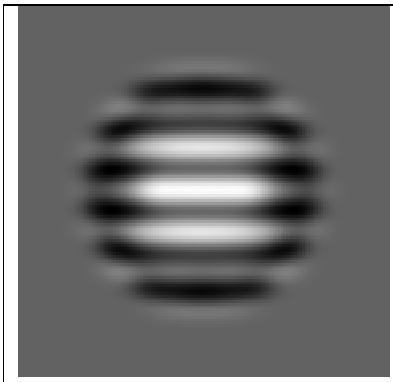


Fig 2.11a. Total phase profile $u(\mathbf{x}, t_1; T) + \varphi(\mathbf{x})$ after the optimization.

$R = 14$.

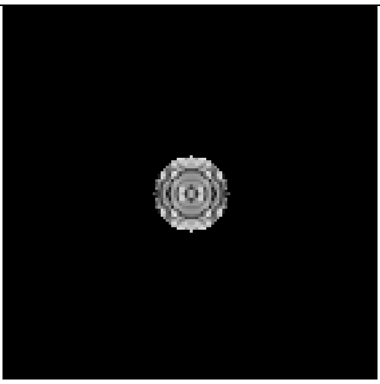


Fig 2.11b. Modulus of the discrete filter T. $N(R) = 613$.

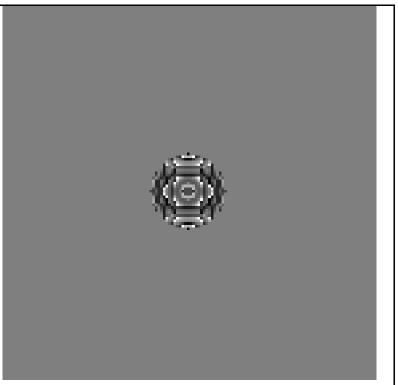


Fig 2.11c. Argument of the discrete filter T. $N(R) = 613$.

Further decreasing R, in such a way that $\eta(\varphi) < 0.88$ or $\eta(u_1) < 0.88$, we observed further deterioration of the optimization quality (**Fig 2.12**) with respect to the above series (**Fig 2.9-Fig 2.11**).

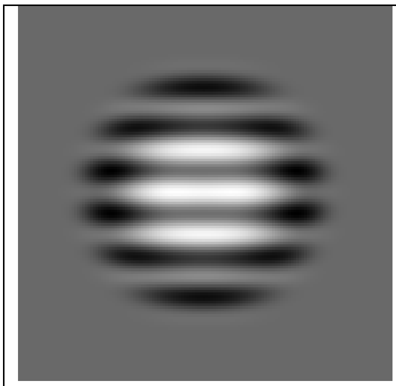
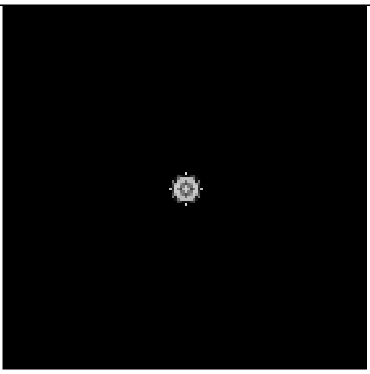
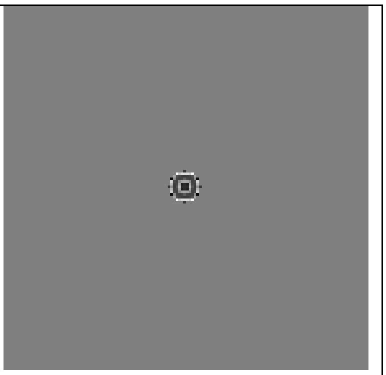
 **Fig 2.12a.** Total phase profile $u(\mathbf{x},t_1;T)+\varphi(\mathbf{x})$ after the optimization. $R=6$.	 **Fig 2.12b.** Modulus of the discrete filter T . $N(R)=113$.	 **Fig 2.12c.** Argument of the discrete filter T . $N(R)=113$.

Fig 2.13 presents the dependence of the optimized value of the target functional J on the radius R. The numerical data on which this plot is based are listed in the table below (**Fig 2.14**).

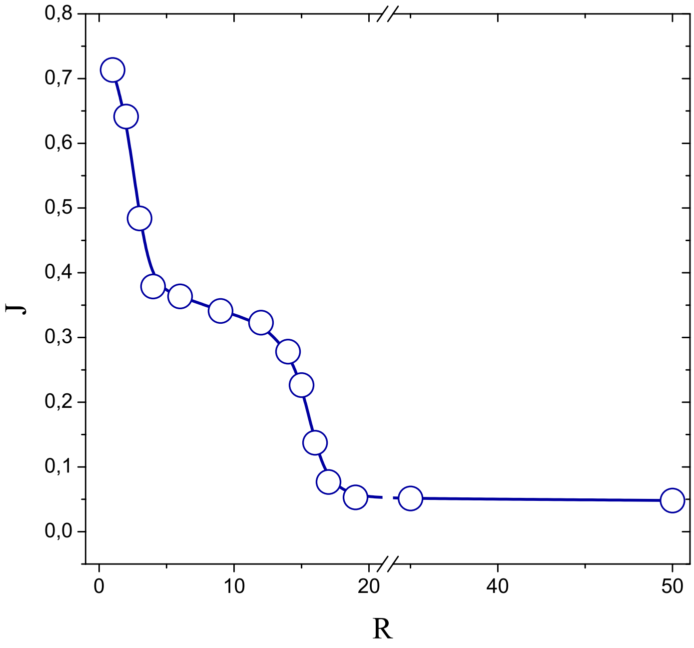


Fig 2.13. Target functional value after the optimization versus the radius of the controlled zone.

R	N(R)	J	$\eta(\varphi)$	$\eta(u_1)$
1	5	0,7134	47,71	47,58
2	13	0,6414	75,76	75,45
3	29	0,4839	82,94	85,43
4	49	0,3789	84,52	92,29
6	113	0,3636	87,25	98,58
9	253	0,3411	87,99	99,88
12	441	0,3231	88,12	99,98
14	613	0,2780	88,50	100,00
15	709	0,2264	89,95	100,00
16	797	0,1375	93,43	100,00
17	901	0,0767	97,37	100,00
19	1129	0,0530	99,60	100,00
35	3853	0,0513	100,00	100,00
50	7845	0,0482	100,00	100,00

Fig 2.14

The main conclusion we may derive from the above results is the following. To achieve the best optimization quality, it is necessary to choose such a number of harmonics to control that to fulfill the conditions $\eta(\varphi) > 0.99$ and $\eta(u_1) > 0.99$, meaning that the controlled zone of the Fourier plane should contain al least 99% of the power of both the input and the target field spectrum.

2.2.7. Dependence of the optimization quality on the parameter *K*

To figure out how the parameter *K* (proportional to the intensity of the input wave) affects the quality of forming the desired phase distribution in the output wave, we performed some numerical simulations, taking the following values of the iterative regularization parameters: $C_0 = 0.833N^{-1}, \beta = 0.45$, where $N = 1793$ is the number of controlled harmonics, and the sequence $\{\alpha_k\}_{k=1}^{+\infty}$ was constructed according to the formula $\{\alpha_k\}_{k=1}^{+\infty} = C_0(k+1)^{-\beta}$. The optimization problem was solved with the conditional gradient method, in which a combined stopping criterion was used to take into account both the discrepancy value and the number of executed iterations of the gradient method. The control quality was estimated using a smooth weight function, $\theta(\mathbf{x}) \in C^{\infty}(\Omega)$. Thus, all the numerical simulation results shown below were obtained at the same fixed parameters of the optimization method. Let us describe the regularities revealed, using one series of simulations as an example.

In each numerical experiment of this series, the diffusion coefficient was set to $l_d^2 = 0.0025$, while the parameters of the input field $A_{in}(\mathbf{x}) = |A_{in}(\mathbf{x})|\exp\{\ i\varphi(\mathbf{x})\}$ were the following:

$A_{in}(\mathbf{x}) \in C^{\infty}(\Omega)$, $A_{in}(\mathbf{x}) = \begin{cases} 1, & \text{if } \mathbf{x} \in \Omega', \\ 0, & \text{if } \mathbf{x} \notin \Omega'', \end{cases}$ where $\Omega' \subset \Omega''$, $\varphi(\mathbf{x}) = 2\cos(7x_2 - 3x_1)$. The target profile was defined by the formula $u_1(\mathbf{x}) = 2\cos(2x_1 + 3x_2)\cos(4x_2 - 5x_1)$ (see **Fig 2.15a.**):

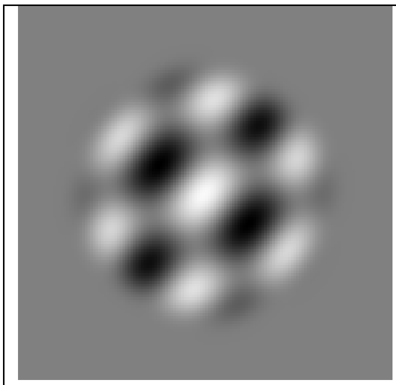


Fig 2.15a. Target phase profile $u_1(\mathbf{x})$.

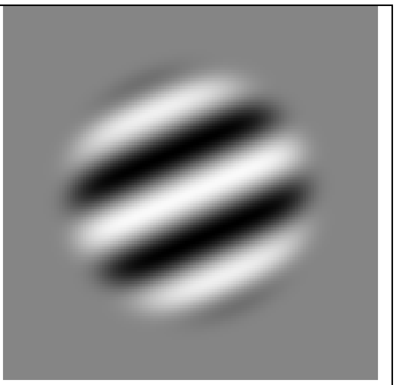


Fig 2.15b. Total phase profile $u(\mathbf{x}, t_1; T) + \varphi(\mathbf{x})$ before the optimization. $K = 0.5$.

Fig 2.15b shows the total phase profile $u(\mathbf{x}, t_1; T) + \varphi(\mathbf{x})$ before the optimization, for $K = 0.5$. We should stress that the 2-radian amplitude of both the input phase distortion and the target phase profile is quite large. **Fig 2.16** clearly demonstrates that $K = 0.5$ is absolutely not enough to achieve good optimization quality. Indeed, at this value of K no considerable approach to the target profile is obtained (compare **Fig 2.15a.** and **Fig 2.16a.**), and the final result of the optimization procedure weakly differs from the initial phase profile (**Fig 2.15b)**. **Fig 2.16b** shows that almost all the harmonics are equally (non-adaptively) transmitted by the filter. This fact clearly indicates that at $K = 0.5$ the depth of the additional phase modulation $u(\mathbf{x}, t_1; T)$ is not enough to efficiently suppress the distortion and simultaneously form the target profile shown in **Fig 2.15a.**

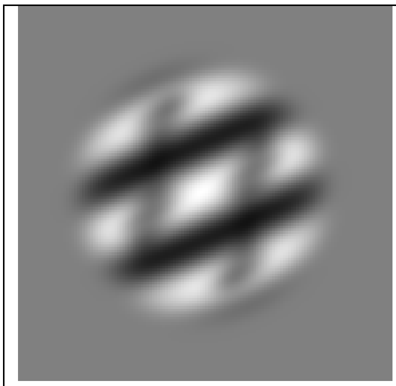


Fig 2.16a. Total phase profile $u(\mathbf{x}, t_1; T) + \varphi(\mathbf{x})$ after the optimization. $K = 0.5$

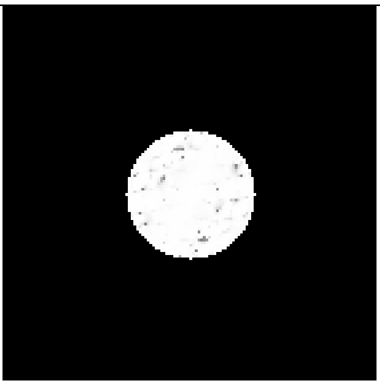


Fig 2.16b. Modulus of the discrete filter T.

$K = 0.5$.

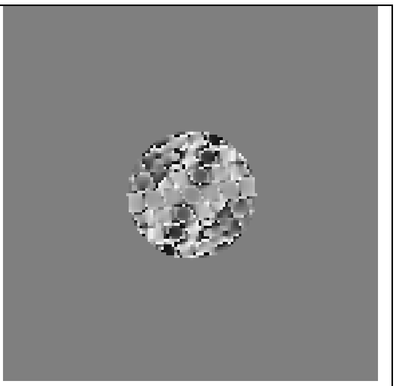
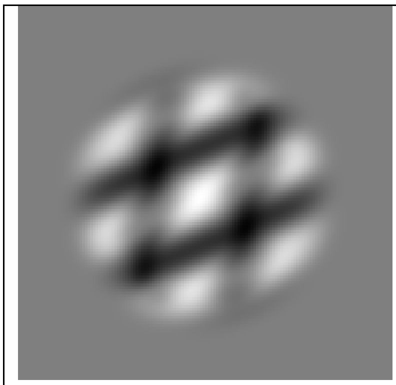
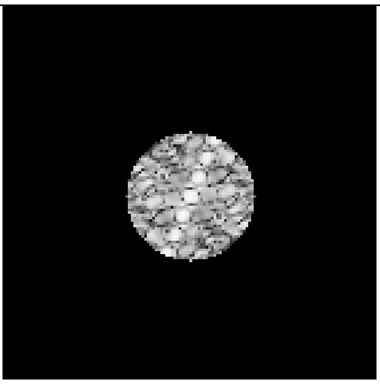
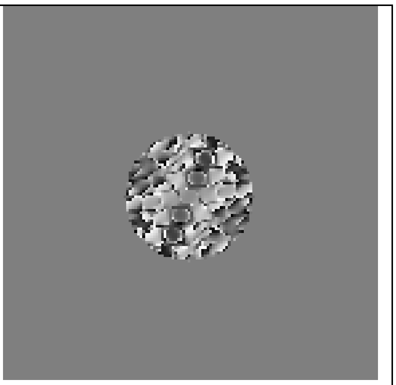


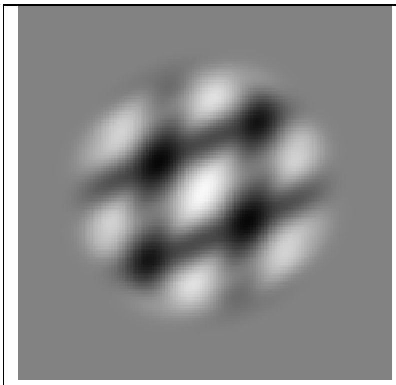
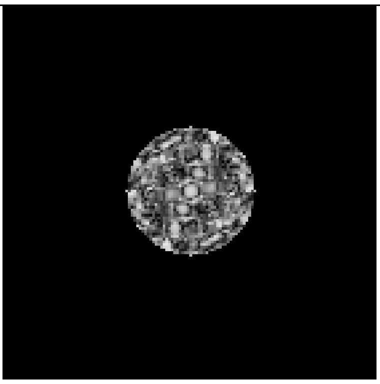
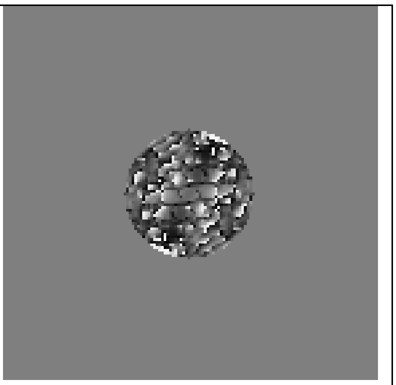
Fig 2.16c. Argument of the discrete filter T.

$K = 0.5$.

In the next numerical experiment, the input wave intensity was increased to $K = 1.1$. The results of this experiment, shown in **Fig 2.17**, demonstrate that the optimization quality considerably improved. It can be seen that the filter obtained for $K = 1.1$ (see **Fig 2.17b,c**) is less trivial and more adaptive than the previous one (**Fig 2.16b,c**). The latter feature is confirmed by the fact that the total phase profile after the optimization (**Fig 2.17a**) decently fits the target profile (**Fig 2.15a**).

 **Fig 2.17a.** Total phase profile $u(\mathbf{x},t_1;T)+\varphi(\mathbf{x})$ after the optimization. $K=1.1$.	 **Fig 2.17b.** Modulus of the discrete filter T . $K=1.1$.	 **Fig 2.17c.** Argument of the discrete filter T . $K=1.1$.

Similar regularities were revealed for $K=3.2$ (**Fig 2.18**) and $K=7.0$ (**Fig 2.19**).

 **Fig 2.18a.** Total phase profile $u(\mathbf{x},t_1;T)+\varphi(\mathbf{x})$ after the optimization. $K=3.2$.	 **Fig 2.18b.** Modulus of the discrete filter T . $K=3.2$.	 **Fig 2.18c.** Argument of the discrete filter T . $K=3.2$.

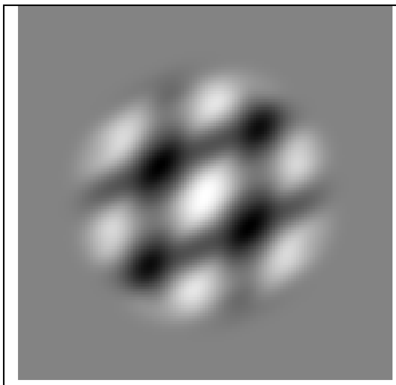
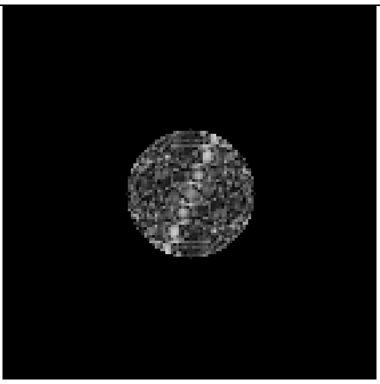
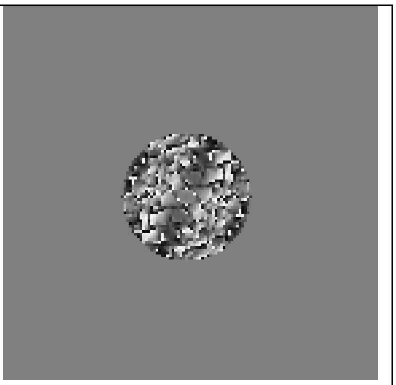
 **Fig 2.19a.** Total phase profile $u(\mathbf{x},t_1;T)+\varphi(\mathbf{x})$ after the optimization. $K=7.0$.	 **Fig 2.19b.** Modulus of the discrete filter T . $K=7.0$.	 **Fig 2.19c.** Argument of the discrete filter T . $K=7.0$.

Fig 2.20 shows a plot of the target functional values obtained as a result of the optimization versus the parameter K. The plot illustrates the following feature. Choosing larger values of K allows one to synthesize a Fourier filter that is better adapted to the specific parameters of the input and the target field. For larger values of K, the optimization results are, as a rule, not worse than those for smaller values of the parameter. This feature is due to the fact that a larger value of K leads to a larger depth of the additional phase modulation $u(\mathbf{x},t_1;T)$ and, in turn, to wider possibilities of correcting the input phase profile $\varphi(\mathbf{x})$ with simultaneously forming the desired phase profile $u_1(\mathbf{x})$.

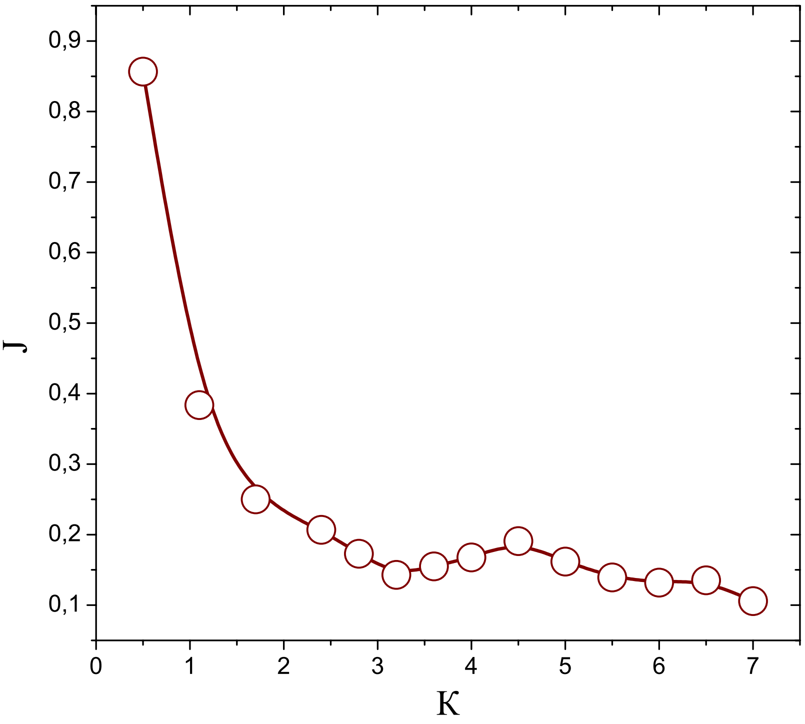


Fig 2.20. Target functional value versus the parameter *K*.

Thus, to achieve better optimization quality, it is advisable to choose a rather large value of the feedback factor K.

2.2.8. Dependence of the optimization quality on the diffusion coefficient of the nonlinear medium particles.

To figure out how the effective diffusion coefficient, $D = l_d^2$, of the nonlinear medium particles affects the quality of forming the desired phase distribution in the output wave, we performed some numerical simulations, taking the following values of the iterative regularization parameters: $C_0 = 0.833N^{-1}, \beta = 0.45$, where $N = 4293$ is the number of controlled harmonics, and the sequence $\{\alpha_k\}_{k=1}^{+\infty}$ was constructed according to the formula $\{\alpha_k\}_{k=1}^{+\infty} = C_0(k+1)^{-\beta}$. The optimization problem was solved with the conditional gradient method, in which a combined stopping criterion was used to take into account both the discrepancy value and the number of executed iterations of the gradient method. The control quality was estimated using a smooth weight function, $\theta(\mathbf{x}) \in C^\infty(\Omega)$. Thus, all the numerical simulation results shown below were obtained at the same fixed parameters of the optimization method. Let us describe the regularities revealed, using one series of simulations as an example.

In each numerical experiment of this series, the feedback factor was set to $K = 3{,}2$, while the parameters of the input field $A_{in}(\mathbf{x}) = |A_{in}(\mathbf{x})| \exp\{ i\varphi(\mathbf{x})\}$ were the following:

$A_{in}(\mathbf{x}) \in C^{\infty}(\Omega)$, $A_{in}(\mathbf{x}) = \begin{cases} 1, & \text{if } \mathbf{x} \in \Omega', \\ 0, & \text{if } \mathbf{x} \notin \Omega'', \end{cases}$ where $\Omega' \subset \Omega''$, $\varphi(\mathbf{x}) = 2\cos(8(x_1 + x_2))\cos(8(x_1 - x_2))$. The target profile is shown in **Fig 2.21a.**

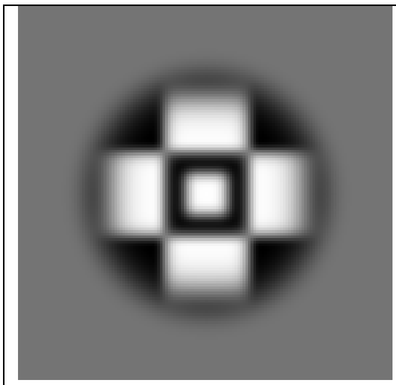


Fig 2.21a. Target phase profile $u_1(\mathbf{x})$.

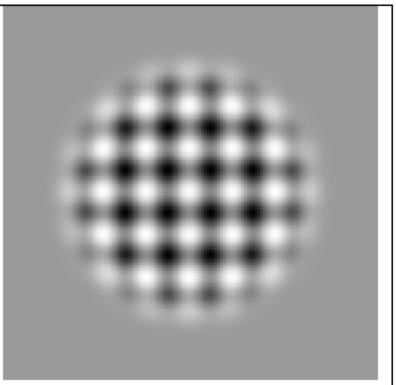


Fig 2.21b. Total phase profile $u(\mathbf{x}, t_1; T) + \varphi(\mathbf{x})$ before the optimization. $D = 0{,}0055$.

Fig 2.21b. shows the total phase profile $u(\mathbf{x}, t_1; T) + \varphi(\mathbf{x})$ before the optimization, for $K = 0.5$. We should stress that the 2-radian amplitude of both the input phase distortion and the target phase profile is quite large.

For the first numerical experiment of this series, we chose $D = 0{,}0025$. The results shown in **Fig 2.22** demonstrate rather good quality of forming the desired phase profile as well as a complex adaptive structure of the optimal filter (**Fig 2.22b,c**).

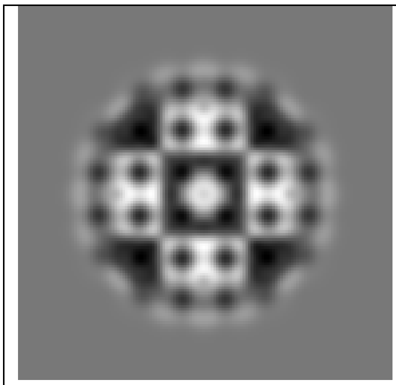


Fig 2.22a. Total phase profile $u(\mathbf{x}, t_1; T) + \varphi(\mathbf{x})$ after the optimization. $D = 0{,}0025$.

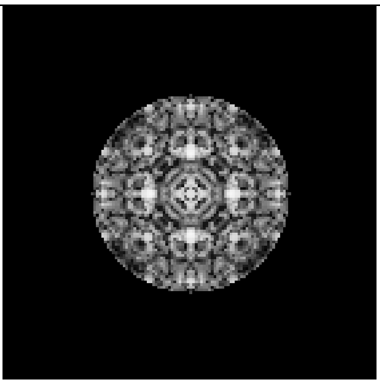


Fig 2.22b. Modulus of the discrete filter T $D = 0{,}0025$.

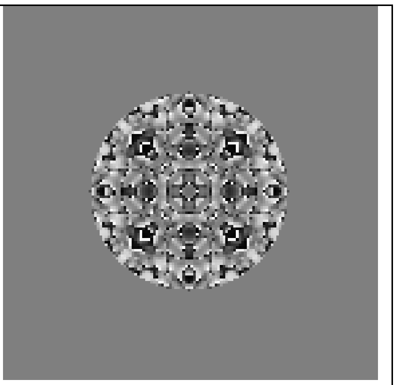


Fig 2.22c. Argument of the discrete filter T. $D = 0{,}0025$.

Increasing the diffusion coefficient D considerably decreases the quality of the numerical optimization. This is caused by strengthening of diffusive coupling between the nonlinear medium particles. At the same time, the larger the diffusive coefficient, the higher the average transmittance of the optimal filter (see **Fig 2.22b, 2.23b, 2.24b, 2.25b**).

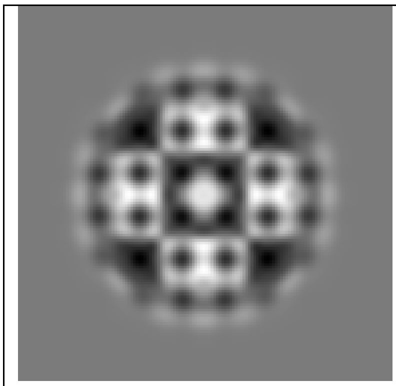


Fig 2.23a. Total phase profile $u(\mathbf{x},t_1;T)+\varphi(\mathbf{x})$ after the optimization. $D = 0{,}0040$.

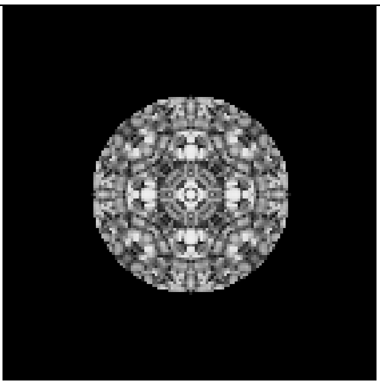


Fig 2.23b. Modulus of the discrete filter T. $D = 0{,}0040$.

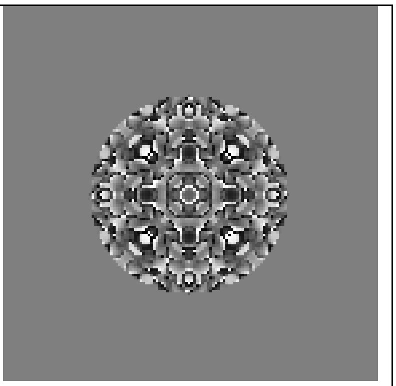


Fig 2.23c. Argument of the discrete filter T. $D = 0{,}0040$.

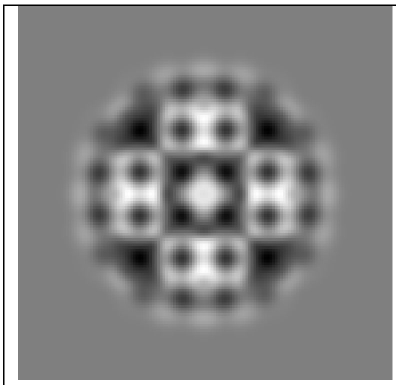


Fig 2.24a. Total phase profile $u(\mathbf{x},t_1;T)+\varphi(\mathbf{x})$ after the optimization. $D = 0{,}0055$.

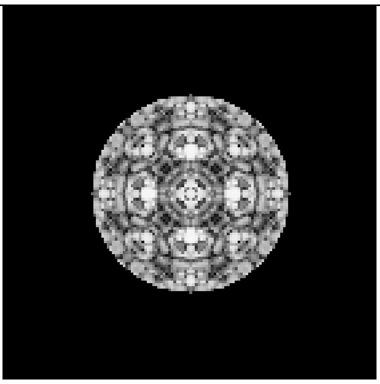


Fig 2.24b. Modulus of the discrete filter T. $D = 0{,}0055$.

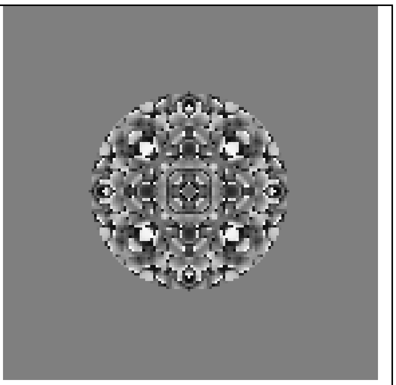


Fig 2.24c. Argument of the discrete filter T. $D = 0{,}0055$.

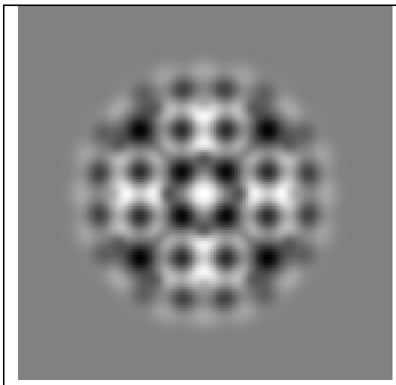


Fig 2.25a. Total phase profile $u(\mathbf{x},t_1;T)+\varphi(\mathbf{x})$ after the optimization. $D = 0{,}01$.

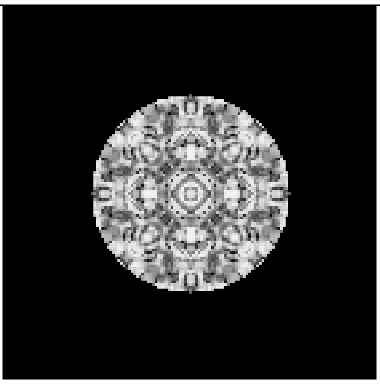


Fig 2.25b. Modulus of the discrete filter T.

$D = 0{,}01$.

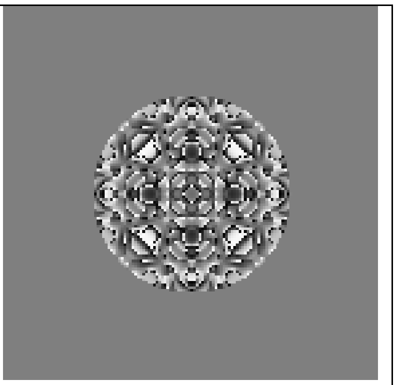


Fig 2.25c. Argument of the discrete filter T.

$D = 0{,}01$.

Fig 2.26 shows a plot of the target functional values obtained as a result of the optimization versus the diffusion coefficient D. The plot demonstrates the following regularity: increasing the diffusion coefficient sufficiently impairs the optimization quality. Since the diffusion coefficient defines the transverse resolution of the nonlinear medium, it can be concluded that in practice it is advisable to use media with a high spatial resolution. For example, it can be a high-resolution optically addressed liquid crystal light valve.

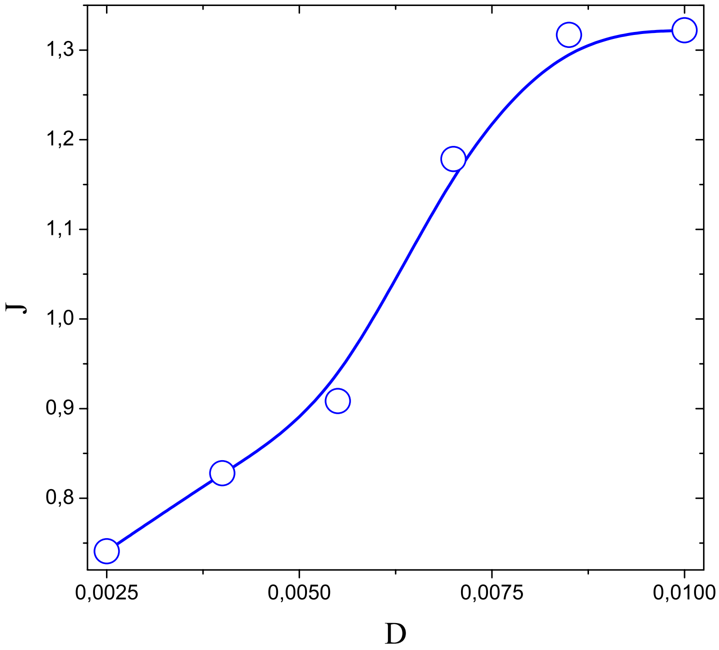


Fig 2.26. Target functional value versus the diffusion coefficient.

2.2.9. Study of the optimization quality achieved in optical systems with two feedback loops.

In this Section, we consider a theoretical model of an optical feedback system, in which the additional phase modulation $u \equiv u(x, y, t)$ of the light wave passed through the nonlinear medium is governed by the equation

$$\tau \frac{\partial u}{\partial t} + u - D\Delta_{\perp} u = I_{fb} + c\int_{0}^{t} I_{fb} dt', \quad I_{fb} \equiv K[1 + \gamma \cos(u + \varphi)], \tag{2.9}$$

Here, τ is the response time of the nonlinear medium, D is the diffusion coefficient of its particles, K is the feedback factor, dependent on the nonlinear properties of the medium and the intensity of the input wave, γ is a coefficient dependent on the average intensity of the feedback wave, φ is the phase profile of the input wave, and c is the factor of integral feedback. Equation (2.9) is to be supplemented with appropriate initial and boundary conditions.

At $c = 0$, the above model describes the dynamics of a number of well-known nonlinear optical systems, such as a nonlinear interferometer with optical feedback or a nonlinear single-pass ring resonator. It is known that such systems are capable of automatically suppressing the input phase distortions φ. In addition, active control of the light wave phase by means of a digital processor connected to the feedback loop can considerably improve the suppression quality.

At the same time, it is more practical and effective to use purely optical control provided by nonlinear media. It is supposed that it is possible to improve the quality of phase distortion suppression (with respect to the case $c = 0$) by supplementing the optical system with an additional feedback loop, whose functionality corresponds to the integral term in the right-hand size of Eq.(2.9). Our numerical results confirm this supposition. For instance, **Fig.2.27** shows the total phase profile, $u + \varphi$, at t=0 for the case when φ is a 2D phase grating with an amplitude of about 2 rad. **Fig.2.28** demonstrates the effect of suppression of these phase distortions for $c = 0$. It can be seen from **Fig.2.28** that, as time goes on ($t = 2\tau$), the phase profile $u + \varphi$ gets somewhat flattened. However, **Fig.2.28** clearly indicates that, under the same conditions, the integral feedback ($c = 1$) allows us to achieve much better results. Moreover, this feature is observed for a wide range of system's parameters. Solving Eq.(2.9) over a longer time interval, one could see that the total phase profile $u + \varphi$ gradually tends to a stationary one and simultaneously continues to flatten. Such a dynamics demonstrates that the system can be applied to phase distortion suppression tasks. From this point of view, it is necessary to further investigate in detail how the suppression quality depends on system's parameters. It would be also useful to consider the case of non-stationary phase distortions .

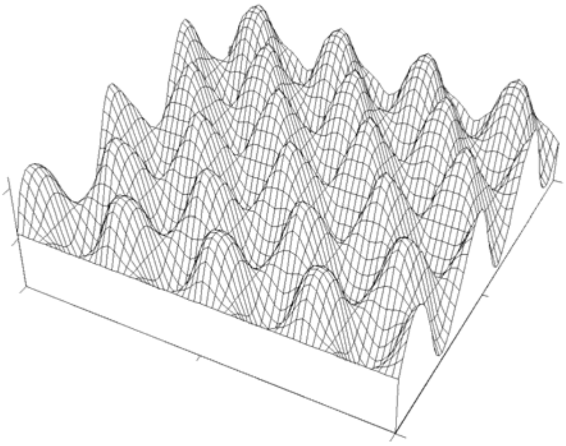


Fig. 2.27. Total phase profile $u + \varphi$ at the moment t=0.

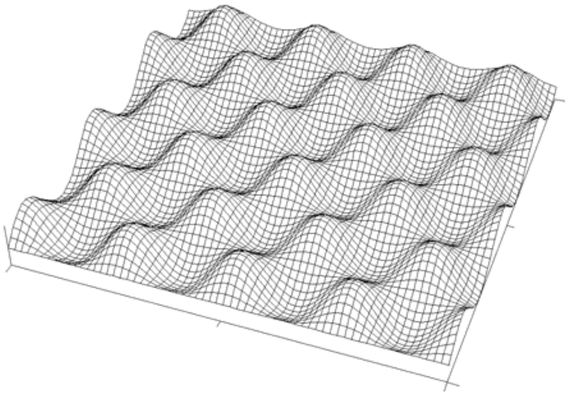


Fig. 2.28. Total phase profile $u + \varphi$ at the moment $t = 2\tau$ ($c = 0$).

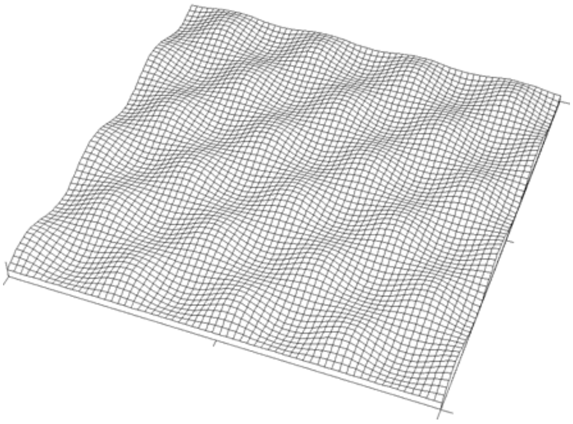


Fig. 2.28. Total phase profile $u + \varphi$ at the moment $t = 2\tau$ ($c = 1$).

2.3. Controlling Irreversible Transforms of Spatial Arguments

The statement of optimal argument transformation $\mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x})) : \overline{\Omega} \to R^2$ has been proposed for the model of 2D feedback optical systems, governed by the functional-differential diffusion equation

$$\frac{\partial u}{\partial t} + u - D\Delta_{\perp} u = F_{\mathbf{g}}(u), \quad t \in [0,T], \quad u(\mathbf{x}, t=0) = u_0(\mathbf{x}), \quad u\,|_{\partial\Omega\times(0,T)} = 0\,. \tag{3.1}$$

Here, $\mathbf{x} = (x_1, x_2)$ is the transverse spatial argument varying in spatial domain Ω with smooth boundary $\partial\Omega$ or in a rectangle; $u = u(\mathbf{x},t) = u(\mathbf{x},t;\mathbf{g})$ is the additional phase modulation in the nonlinear medium layer; $\Delta_{\perp}$ is the transverse Laplace operator; D is the diffusion coefficient; $F_{\mathbf{g}}(u)$ is the superposition of the nonlinear term $F(u)$ and a controllable transform, $\mathbf{g}$, of the spatial argument. For instance, the formula $F(u) = K(1+\gamma\cos u)$ is used for modeling field interactions in an interferometer or passive ring resonator with a thin Kerr-type medium, factor K is proportional to the intensity of the input light wave, γ is the visibility of the interference pattern.

2.3.1. Distributed statement of irreversible transform of spatial argument

In the case of a reversible transform, $\mathbf{g}(\mathbf{x}) = (g_1(\mathbf{x}), g_2(\mathbf{x})) : \overline{\Omega} \xrightarrow{\text{onto}} \overline{\Omega} \subset R^2$, the classical formula

$$F_{\mathbf{g}}(u)(\mathbf{x},t) = F\,(u(\mathbf{g}^{-1}(\mathbf{x}),t)) \tag{3.2}$$

is used. It describes a point-to-point superposition of $F(u)$ with the inverse transform $\mathbf{g}^{-1}(\mathbf{x})$. For instance, reflection and/or rotations in a circular domain are valid transforms that are easy to realize by means of reflecting mirrors and rotating prisms.

A complicated mathematical problem is how to give an adequate interpretation of such a superposition for an arbitrary Lebesgue-measurable transform. We propose to consider such a superposition in the generalized sense, i.e. as a distribution defined by the formula

$$\left\langle F_{\mathbf{g}}(u), \varphi \right\rangle = \int\limits_{\overline{\Omega}} F(u(\mathbf{x},t))\,\varphi(\mathbf{g}(\mathbf{x}))\,d\mathbf{x}\ ,\ \forall\varphi \in C(\overline{\Omega}). \tag{3.3}$$

The advantage of this approach is that there is no inverse transform implemented in the integral calculation and only a direct transform acts in the superposition with the test function $\varphi(\mathbf{x})$. It is important that for the class of reversible transforms, the distribution in (3.3) is regular and coincides with that of (3.2). The following new result on solvability and smoothness has been proved.

Theorem 3.1. Let the initial condition $u_0 \in L^2(\Omega)$ and the function F be bounded and continuously differentiable with a bounded derivative. Then, for any Lebesgue-measurable transform $\mathbf{g} \in G$, the conjugate problem (3.1)-(3.3) has the unique generalized solution $u \in L^2(0,T;H_0^{2-s}(\Omega)), \partial_t u \in L^2(0,T;H^{-s}(\Omega))$ with arbitrary $s \in (1, 3/2)$. The solution Holder-continuously depends on the argument transforms in the sense of the estimate

$$\|u - \tilde{u}\|_{L^2(0,T;H^{2-s}(\Omega))} + \|u - \tilde{u}\|_{L^p(Q_T)} \le C(p)\,\|g - \tilde{g}\|^{\beta}_{L^2(\Omega)}$$

with $\beta = 4p^{-1} - 1$ and arbitrary $p \in (2,4)$.

2.3.2. Statement of optimal control of argument transforms

The goal of the optimization is to approximate to a given function, $u_1(\mathbf{x},t)$, the solution of (3.1). The target functional estimates the quality of the approximation being the optimization result. This functional is determined by the requirements of the problem formulated and usually coincides with the norm of deviation from the assigned target, measured in a certain functional space. It is proposed to consider a weighed integral functional of the following kind:

$$J(\mathbf{g}) = \int\limits_{Q_T} \rho(\mathbf{x},t)\,|\,u(\mathbf{x},t;\mathbf{g}) - u_1(\mathbf{x},t)\,|^2 \,d\mathbf{x}dt\,, \tag{3.4}$$

which allows embracing different statements of optimization problems. In such a way, varying the form of the function $\rho(\mathbf{x},t)$ makes it possible to consider the solution approximation within the whole time interval and spatial domain, as well as on its certain portion and subdomain. Parameter $\rho(\mathbf{x},t)$ determines the relative contribution of each temporal layer in the above functional. The minimization problem for the functional $J(\mathbf{g})$ is formulated on the admissible set G of Lebesgue-measurable transforms satisfying pointwise restrictions:

$$J(\mathbf{g}) \to \inf_G\,, \quad G = \{\mathbf{g} \in L_2(\Omega) : \mathbf{g}(\overline{\Omega}) \subseteq \overline{\Omega}\}\,. \tag{3.5}$$

Theorem 3.2. Let the conditions of theorem 3.1 hold, the target function $u_1 \in L^2(\Omega)$, and $0 \le \rho \in L^\infty(\Omega)$. Then the target functional $J(\mathbf{g})$ is correctly defined on the admissible set G and Holder-continuously depends on the argument transforms.

2.3.3. Conjugate problem and target functional gradient

To compute the transform sequence that minimizes the functional, we have used a variation approach based on extracting the main linear part of the increment of functional $J(\mathbf{g})$. Let $\psi = \psi(\mathbf{x},t)$ denote the solution of so-called conjugated problem:

$$-\partial_t\psi + \psi - D\Delta_\perp\psi = 2\rho(u - u_1) + F'(u)\psi(\mathbf{g}(\mathbf{x}),t)\,,\;\; \psi\,|_{\partial\Omega\times(0,T)} = 0,\;\; \psi\,|_{t=T} = 0. \tag{3.6}$$

Note that, unlike that in Eq. (3.1), in the right-hand side of Eq. (3.6) the sought function is superposed to transform $\mathbf{g}(\mathbf{x})$ in the classical sense. Furthermore, transform $\mathbf{g}(\mathbf{x})$ is implicitly contained in (3.6) through the solution $u = u(\mathbf{x},t;\mathbf{g})$ of the diffusion equation (3.1).

Theorem 3.3. Let the conditions of theorems 3.1-3.2 hold and the target function have additional property, $u_1 \in L^4(Q_T)$. Then, for any Lebesgue-measurable transform $\mathbf{g} \in G$ the conjugate problem (3.6) has the unique solution $\psi \in H_p^{2,1}(Q_T)$ with arbitrary $p \in (2,4)$, and this solution obeys the estimation $\|\psi\|_{H_p^{2,1}(Q_T)} \le C(p)$, where the constant is independent of $\mathbf{g}$. The solution Holder-continuously depends on the argument transforms in the sense of the estimate

$$\|\psi - \tilde{\psi}\|_{H_p^{2,1}(Q_T)} + \|\psi - \tilde{\psi}\|_{C([0,T];H_q^1(\Omega))} \le C(p)\,\|g - \tilde{g}\|_{L^2(\Omega)}^{\beta}$$

with $\beta = 4p^{-1} - 1$ and arbitrary $q \in (2,+\infty)$.

Theorem 3.4. Let the conditions of theorem 3.3 hold. Then, for any Lebesgue-measurable transform $\mathbf{g} \in G$, the target functional $J(\mathbf{g})$ is uniformly quasi-differentiable, i.e. the following increment formula is valid:

$$J(\mathbf{g}+\mathbf{h}) - J(\mathbf{g}) \;\; = \langle \mathbf{J}'(\mathbf{g}),\mathbf{h}\rangle_{L^2(\Omega)} + R(\mathbf{g},\mathbf{h}), \quad \forall \mathbf{g},\mathbf{g}+\mathbf{h} \in G\,,$$

with the reminder term $R(\mathbf{g},\mathbf{h})$, which is uniformly estimated on G with the inequality

$$|\,R(\mathbf{g},\mathbf{h})\,| \;\; \le \;\; C\,\|\mathbf{h}\|_{L^2(\Omega)}^{1+\mu},\; \mu > 0.$$

The gradient $\mathbf{J}'(\mathbf{g}) \in L^\infty(\Omega)$ is calculated according to the formula

$$\mathbf{J}'(\mathbf{g}) = \int\limits_0^T F(u(\mathbf{x},t;\mathbf{g}))\nabla\psi(\mathbf{g}(\mathbf{x}),t)dt \tag{3.7}$$

and meets the Holder-continuity property, $\| \mathbf{J}'(\mathbf{g}) - \mathbf{J}'(\tilde{\mathbf{g}}) \|_{L^2(\Omega)} \quad \le \quad L \| \mathbf{g} - \tilde{\mathbf{g}} \|^{1/3}_{L^2(\Omega)}$.

2.3.4. Description of gradient-type methods

To numerically implement the method based on the previously derived gradient formula and on the additional information on its Holder-continuity, three methods were elaborated. The first one, a **special gradient projection method**, was elaborated and theoretically analyzed in the framework of the Project. In this version, the sequence of optimizing transforms is derived from the following expressions:

$$\mathbf{g}^{k+1} = \mathbf{g}^k + \alpha_k \mathbf{p}^k,\ \mathbf{p}^k = \frac{\bar{\mathbf{g}}^k}{\left\|\bar{\mathbf{g}}^k\right\|_{L^2(\Omega)}},\quad \bar{\mathbf{g}}^k = \mathrm{P}_G\left(\mathbf{g}^k - \mathbf{J}'(\mathbf{g}^k)\right) - \mathbf{g}^k,\quad \mathbf{g}^0 \in G, \qquad (3.8)$$

where P_G is the operator of projection onto the admissible convex set *G*. For a rectangular domain $\overline{\Omega}$, the projection operator periodically cuts those values going beyond $\overline{\Omega}$. The step of the method is chosen in accordance with the rule

$$\alpha_k = \min\left\{ \left(\frac{4/3}{L+\varepsilon} \cdot \left\|\mathbf{g}^k\right\| \right)^3, \left\|\mathbf{g}^k\right\| \right\},$$ where $\varepsilon > 0$ is a parameter of the method.

If, for some *k*, $\bar{\mathbf{g}}^k = 0$, then the necessary condition of minimum, $\mathbf{g}^k = \mathrm{P}_G\left(\mathbf{g}^k - \mathbf{J}'(\mathbf{g}^k)\right)$, is achieved, and the iterations should be stopped. The following statement governs the convergence of the method.

Theorem 3.4. Let the conditions of theorem 3.3 hold and, additionally, *G* is a closed convex set and the consequence $\mathbf{g}^k$ from (3.3) is infinite. Then

$$\left\|\mathrm{P}_G\left(\mathbf{g}^k - \mathbf{J}'(\mathbf{g}^k)\right) - \mathbf{g}^k\right\|_{L^2(\Omega)} \to 0, \quad k \to \infty. \qquad (3.9)$$

Comment. Condition (3.10) means that the consequence $\mathbf{g}^k$ from (3.3) converges to the asymptotic version of the necessary condition of minimum mentioned above.

For practical calculations, we also used a **simplified gradient projection method,** in accordance with the following iteration law:

$$\mathbf{g}^{k+1} = \mathbf{P}_\Omega(\mathbf{g}^k - \alpha_k \mathbf{J}'(\mathbf{g}^k)). \qquad (3.10)$$

The sequence $\{\alpha_k\}$ of steps in (3.10) was determined by the splitting method on the basis of the condition of a strict decrease of the functional at each step: $J(\mathbf{g}^{k+1}) < J(\mathbf{g}^k)$.

The third gradient-like method elaborated is a **conditional gradient method**. For the case of the rectangular domain $\overline{\Omega} = [0, 2\pi] \times [0, 2\pi]$, the corresponding iterative sequence is constructed according to the formula

$$\mathbf{g}_{k+1}(\mathbf{x}) = \mathbf{g}_k(\mathbf{x}) + \alpha(\overline{\mathbf{g}_k}(\mathbf{x}) - \mathbf{g}_k(\mathbf{x})), \qquad (3.11)$$

where $\overline{\mathbf{g}_k}(\mathbf{x}) = (\bar{g}_{k_1}(\mathbf{x}), \bar{g}_{k_2}(\mathbf{x}))$ is an additional vector taken from the following cut-off rule:

$$\overline{g_{k\,m}}(\mathbf{x}) = \begin{cases} 0, & J'_m(\mathbf{g}) \ge 0, \\ 2\pi, & J'_m(\mathbf{g}) < 0, \end{cases} \quad m = 1,2,$$

$J'_m(\mathbf{g})$ being the *m*-th coordinate of the gradient $\mathbf{J}'(\mathbf{g}) = (J'_1(\mathbf{g}), J'_2(\mathbf{g}))$.

2.3.5. Projection finite-element approximations

In the scope of the Project, special finite-dimensional approximations have been elaborated. During the investigation, the following three sufficient obstacles in implementing traditional finite-difference approximations have been successfully overcome:

- Superposition with an argument transform in (3.3) acts as a distribution and requires an adequate approximation to be implemented;
- The value of the transformed argument $\mathbf{g}(\mathbf{x})$ in integral (3.3) may coincide with none of the mesh nodes;
- Even for an infinitely smooth but irreversible argument transform (such as $\mathbf{g}(\mathbf{x}) \equiv \mathbf{x}_0 \in \Omega$ < that is, focusing onto a point), the corresponding solution of (3.1), $u(\mathbf{x},t)$, need not to be a smooth enough to implement classical finite differences.

The idea of the approach we implemented is to utilize projection finite-element scheme, which keeps the structure of distributed kind of generalized superposition in (3.1) and (3.3). To describe it, we use the space $S_1 = Lin\{\varphi_{i,j}(\mathbf{x})\}$ of piece-wise linear functions $\varphi_{i,j}(\mathbf{x}) = \varphi_i(x_1)\cdot\varphi_j(x_2)$ constructed on the basis of first-order B-splines, as above. The approximate solution of (3.1) is sought as an element $w = (w^0,...,w^M) \in S$ of the space $S = \underbrace{S_1 \times S_1 \times S_1 \times ... \times S_1}_{M+1\ \text{times}}$, satisfying the following finite-element projection analog of (3.1), (3.3):

$$\left\langle \tau^{-1}(w^m - w^{m-1}),\varphi \right\rangle + \left\langle w^m,\varphi \right\rangle + D\left\langle \nabla_\perp w^m, \nabla_\perp \varphi \right\rangle = \left\langle F(w^m), \varphi(g) \right\rangle,\ m = 1,...,M\ ,\forall \varphi \in S_1,$$

where brackets mean the inner product in the Lebesgue space $L^2(\Omega)$. There is another matrix form of the previous formula, being more convenient for computations:

$$((1+\tau)B + \tau\, D\Lambda)w^m = \tau\, \Phi(w^m) + Bw^{m-1},\ m = 1,...,M\ . \tag{3.12}$$

To solve nonlinear matrix equation. (3.12), we applied iterations:

$$((1+\tau)B + \tau\, D\Lambda)w^{(s+1)} = \tau\, \Phi(w^{(s)}) + Bw^{m-1},\ s = 0,1,...,\ w^{(0)} \equiv w^{m-1}. \tag{3.13}$$

It is important to underline that calculating the matrix $\Phi(f)$ with elements

$$\Phi_{k,l}(f) = \frac{1}{h_1 h_2}\int\limits_{\Omega} F(f(\mathbf{x}))\,\varphi_{k,l}(\mathbf{g}(\mathbf{x}))\,d\mathbf{x}\,, \qquad f \in S_1,$$

leads to sums $\Phi_{k,l} = \sum\limits_{m,n=1}^{N_1,N_2} F(f(x_1^m, x_2^n))\varphi_{k,l}(\mathbf{g}(x_1^m, x_2^n))$, which are very expensive for direct computations. Basing on the finiteness of spline support, substantial economy of the operations number can be achieved. Indeed, since

$$\operatorname{supp}\varphi_{k,l} \equiv (x_1^{k-1}, x_1^{k+1}) \times (x_2^{l-1}, x_2^{l+1}) = \Omega_{k,l},\ \ k = 1,..., N_1 - 1,\ \ l = 1,..., N_2 - 1,$$

it is sufficient to sum over those indexes (m,n) for which $\mathbf{g}(x_1^m, x_2^n) \in \Omega_{k,l}$ and to keep the information of prototypes $\mathbf{g}^{-1}(\varphi_{k,l}(\overline{\Omega}))$ in the form of the following structure (see fig. 3.1):

`struct fiGArray{double value; struct fiGArray* next; int m,n;},`

where `value` means the value of $\varphi_{k,l}(\mathbf{g}(x_1^m, x_2^n))$; `next` is the pointer to the next element; `m,n` are the node numbers before the transform.

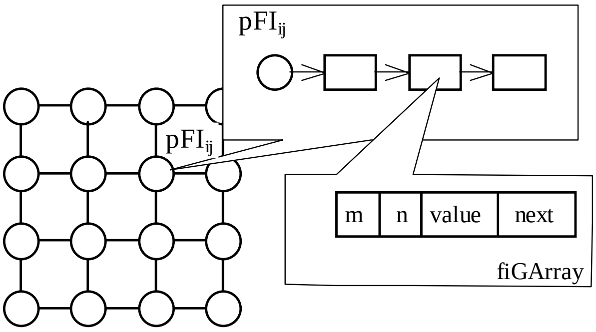


Fig. 3.1. Structure for keeping the prototype information.

In the scope of the Project, we have presented a theoretical basis for the projection method used.

Theorem 3.5. Let the conditions of Theorem 3.3 hold. Then:

1. For every sufficiently small temporal step $\tau > 0$, Eq. (3.12) have a unique solution. The iterations $w^{(s)}$ given by (3.13) converge to the exact solution w^m of Eq. (3.12) in $L^2(\Omega)$, according to a geometric progression with the denominator value being proportional to $\tau^{1/2}$;
2. The estimate of convergence, $\| u - w \|_{L^2(Q_T)} \le C\left(\tau^{1/4-\alpha} + |h|^{1/2-2\alpha}\right)$, is valid with the constant C being independent of τ, h_1, h_2.

By analogy with Eq. (3.12), a projection-difference scheme has been suggested and elaborated for numerically solving the conjugate problem (3.6). In the operator form, the projection-difference scheme for searching $\xi^m \in S_1$ can be expressed as

$$\left((1+\tau)B + \tau D\Lambda\right)\xi^{m-1} = f^m, \qquad f^m = B(2\tau\rho(t_m)(w^m - u_1^m) + \xi^m) + \Psi^m, \qquad (3.13)$$

where Ψ^m is used for approximating the inner product $\left\langle F'(w^{m-1})\xi^m(\mathbf{g}), \varphi\right\rangle$. To implement the scheme, it is essential to use the information on the mutual correspondence between the prototype of points and spline supports, which is built during solving the direct problem and is stored in the fiGArray structure described above.

2.3.6. Testing the simplified gradient projection method

Algorithm (3.10) was first tested with spatially 1D target functions for the problem of phase distortion suppression, described by Eq. (3.1) subject to 2π-periodic boundary conditions. Such a distortion was introduced to a spatially inhomogeneous (with respect to argument x_1) input phase profile:

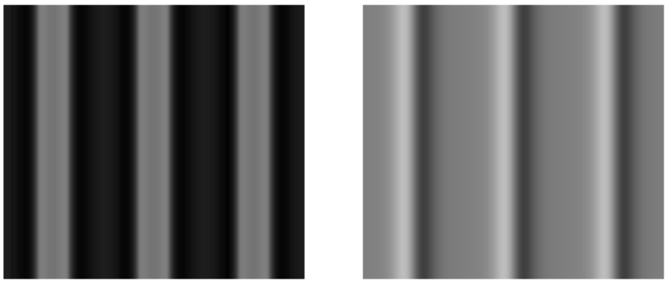
$$u_0(x_1, x_2) = 2 + 0.3\left(\exp\left(-15(x_1 - \pi)^2\right) + \exp\left(-15(x_1 - \pi/3)^2\right) + \exp\left(-15(x_1 - 5\pi/3)^2\right)\right).$$

The suppression was ensured by the feedback loop of a nonlinear interferometer with the nonlinearity $F(u) = K(1 + \gamma\cos u)$, where K is a coefficient proportional to the intensity of the input light wave; and γ is the contrast of the interference pattern. In the computation performed, the parameters of the nonlinearity and equation (3.1) were set to $K = 3.6$, $\gamma = 1$, $D = 0.07$. It was necessary to find the argument transformation, $\mathbf{g}(\mathbf{x})$, which ensures the minimum deviation of the phase profile $u = u(\mathbf{x}, t)$ from its initial (at $t = 0$) distribution $u_0(\mathbf{x})$ within the whole time interval $t \in [0, T]$ for $T = 3$. Such a statement is formulated in the form of the minimization problem (3.4)-(3.5) with the stationary target profile $u_1(\mathbf{x},t) \equiv u_0(\mathbf{x})$. The weight function contained in (3.4) had the following form:

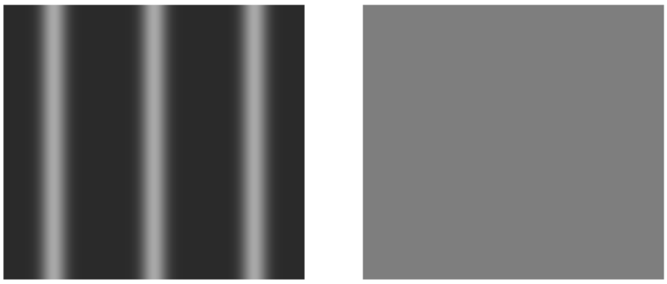
$$\rho(\mathbf{x},t) = \exp\left(1 - T^2\left(T^2 - (t-T)^2\right)^{-1}\right),$$

which takes into account that the contribution of the final portion of the trajectory to the functional value exceeds that of the initial portion. As an initial approximation for the control law in the gradient method (15), the identical transform, $\mathbf{g}^0(\mathbf{x}) = \mathbf{x}$, was taken. By the final point of time, the solution became spatially homogeneous, with the functional $J(\mathbf{g}^0) = 0.79$.

As a result of two steps of the gradient method (3.14), the value of the functional decreased down to $J(\mathbf{g}^2) = 0.30$, that is, by the factor of more than 2.5. Figs. 3.2a-b show the profiles of the target function and the optimized solution obtained. Cross-sections of these profiles are shown in Fig. 3.3a. It can be seen that the optimization procedure has succeed in substantially reflecting the main features of the target function. One can notice certain symmetry peculiar to the optimization problem: if the initial approximation of the transform does not differ from the identical with respect to variable x_2 and if the target profile is independent of this variable, then the solution of the conjugate problem (and, consequently, the functional gradient as well) is also independent of x_2. Therefore, the gradient procedure (3.14) generates a sequence of argument transforms for which the first component, $g_1(\mathbf{x})$, is responsible for spatial nonhomogeneity of the target profile, while the second component does not change: $g_2(\mathbf{x}) \equiv x_2$. Figs. 3.2c-d show the profiles of the deviations of the corresponding components, $g_1(\mathbf{x})$ and $g_2(\mathbf{x})$, of the second iteration of the argument transform, $\mathbf{g}^2(\mathbf{x}) = (\mathrm{g}_1(\mathbf{x}), \mathrm{g}_2(\mathbf{x}))$, from the identical transform, $\mathbf{g}^0(\mathbf{x}) = \mathbf{x}$. The cross-section of $g_1(\mathbf{x})$ is shown in Fig. 3.3b.



a. Phase profile $u(\mathbf{x},T)$.   c. Profile $g_1(\mathbf{x})$-x_1.



b. Target profile $u_1(\mathbf{x},T)$.   d. Profile $g_2(\mathbf{x})$-x_2.

Fig. 3.2. Results of optimization for the problem of phase distortion suppression.

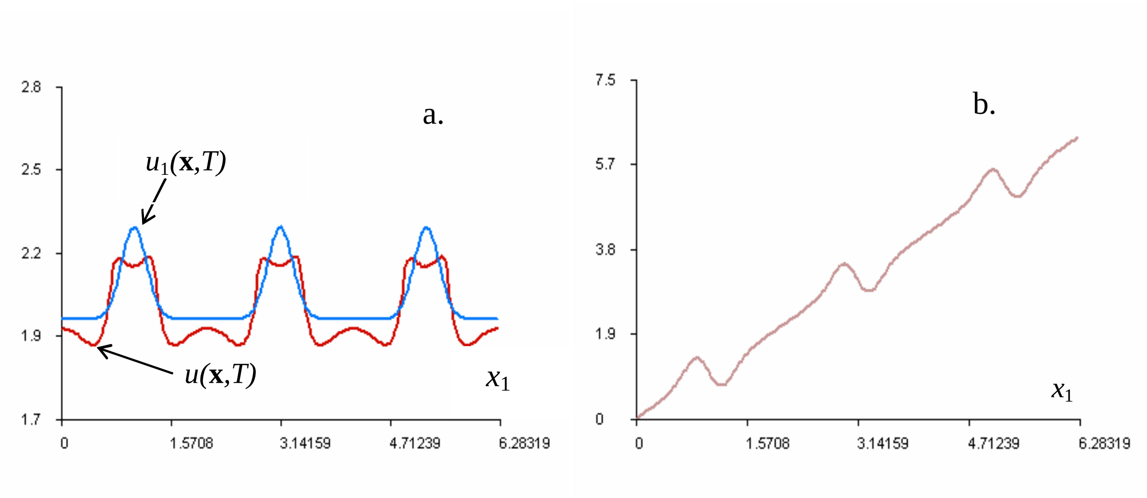
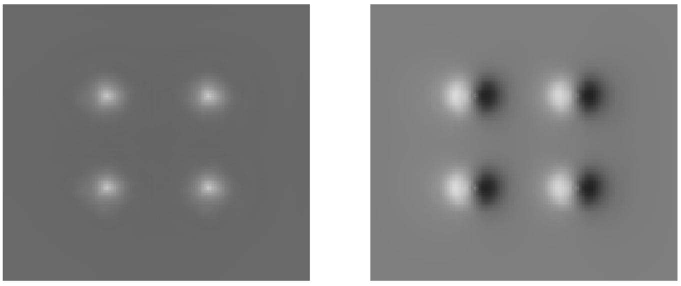


Fig. 3.3. Cross-sections of $x_2 = const$: $u(\mathbf{x},T)$ and $u_1(\mathbf{x},T)$ from Fig.2a-b (a); $g_1(\mathbf{x})$ component of the optimized transform (b).

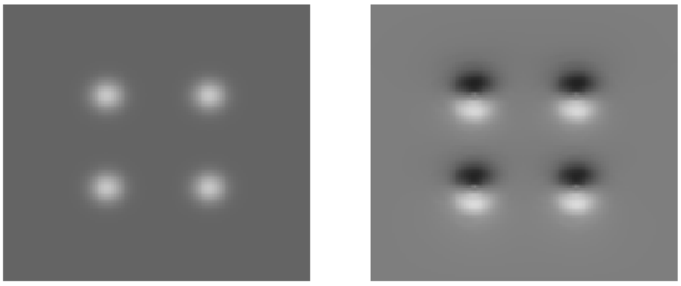
In the second example, the control target was switching system (3.1) from the initial spatially homogeneous state $u(\mathbf{x},t=0)=u_0(\mathbf{x})=2$ with a constant (for each point $\mathbf{x}$) velocity to the final (at $T=2$) state defined by a phase profile having four localized peaks (see Fig.3.4). In the corresponding functional (3.4), the target state was defined by the following formula:

$$u_1(x_1,x_2)=$$
$$=2+t\cdot\left(\exp\left(-9(x_1-2\pi/3)^2-9(x_2-2\pi/3)^2\right)+\exp\left(-9(x_1-4\pi/3)^2-9(x_2-2\pi/3)^2\right)\right)+$$
$$+t\cdot\left(\exp\left(-9(x_1-2\pi/3)^2-9(x_2-4\pi/3)^2\right)+\exp\left(-9(x_1-4\pi/3)^2-9(x_2-4\pi/3)^2\right)\right)$$

In the case of the identical transformation $\mathbf{g}^0(\mathbf{x})=\mathbf{x}$, the solution of problem (3.1) remained spatially homogeneous within the whole time interval $t\in[0,T]$, which led to the functional value of $J(\mathbf{g}^0)=1.63$ (for the following parameters: $K=4.3$, $\gamma=1$, $D=0.07$). As a result of five steps of the gradient method (3.14), the functional decreased down to $J(\mathbf{g}^5)=0.37$, that is, more than 4 times. The optimized solution profile $u(\mathbf{x},T)$ well corresponded with the target profile $u_1(\mathbf{x},T)$ (see Fig.3.4a-b). The deviations of the obtained components, $\mathrm{g}_1(\mathbf{x})$ and $\mathrm{g}_2(\mathbf{x})$, of the transform $\mathbf{g}^5(\mathbf{x})=(\mathrm{g}_1(\mathbf{x}),\mathrm{g}_2(\mathbf{x}))$ are shown in Fig.3.4c-d. It can be seen that the x_1-cross-section of the $\mathrm{g}_1(\mathbf{x})$ profile and the x_2-cross-section of the $\mathrm{g}_2(\mathbf{x})$ profile look very similar to those plotted in Fig.3.3, especially at their localization peaks.



a. Phase profile $u(\mathbf{x},T)$. c. Profile $g_1(\mathbf{x})$-x_1.



b. Target profile $u_1(\mathbf{x},T)$. d. Profile $g_2(\mathbf{x})$-x_2.

Fig. 3.4. Optimization results for a 2-D localized target function.

Basing on the above calculations, the following conclusions can be made on the features of formation of localized solutions by means of the spatial argument transform control:

3. Localized structures can be obtained by means of argument transforms that are not invertible in the vicinity of each localization peak. A similar conclusion has been earlier made for the spatially homogeneous case.
4. If the initial and the target state of the system are independent of one of the variables, then the corresponding component of the optimized transform can be considered identical with this variable.
5. Simultaneous localization with respect to two variables leads to an optimized transform, each component of which is mainly responsible for the localization along the corresponding variable.

2.3.7. Testing the conditional gradient method in contrast with the simplified method

We have tested and compared the gradient projection and the conditional gradient method for the problem of constructing a phase profile being maximally close to the desired target profile at the final time point T. For this model, it is convenient to use as a weight function $\rho(t)$ some function that is nonzero only within a portion of the segment and equals unity at the final time point. Steeper functions, such as super-Gaussian, are more advantageous for this purpose. When modeling the optimization problem with the following parameters

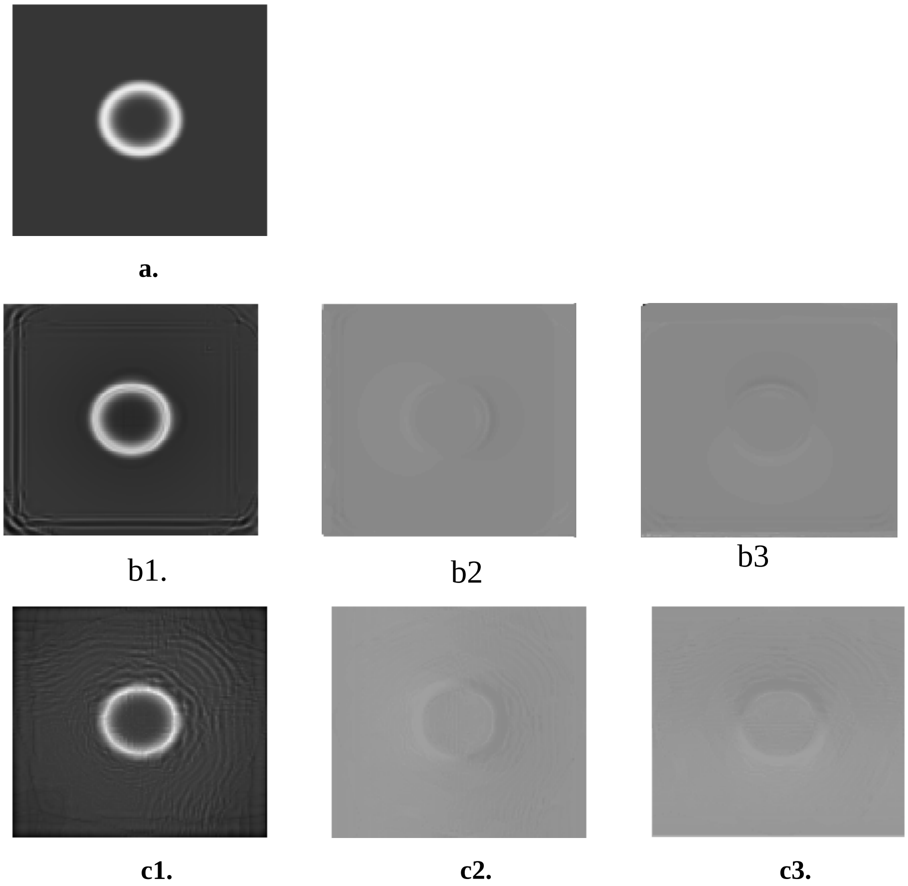
$$u_1(x_1,x_2,t) = 2.02+0.25t(x_1{}^2+x_2{}^2)^3\exp(-2\cdot(x_1{}^2+x_2{}^2)^4),$$

$$D = 0.02,\quad K = 3.6,\quad \gamma = 1,$$

$$u_0(x_1,x_2) = 2.02 + 0.5(x_1{}^2 + x_2{}^2)^3\exp(-2\cdot(x_1{}^2 + x_2{}^2)^4)$$

$$\mathbf{g}^0(\mathbf{x}) = \mathbf{x},\ T = 2,$$

by means of the gradient projection method, the target functional finally decreased by 90.8%. When using the conditional gradient method, the target functional decreased by 70.9%.



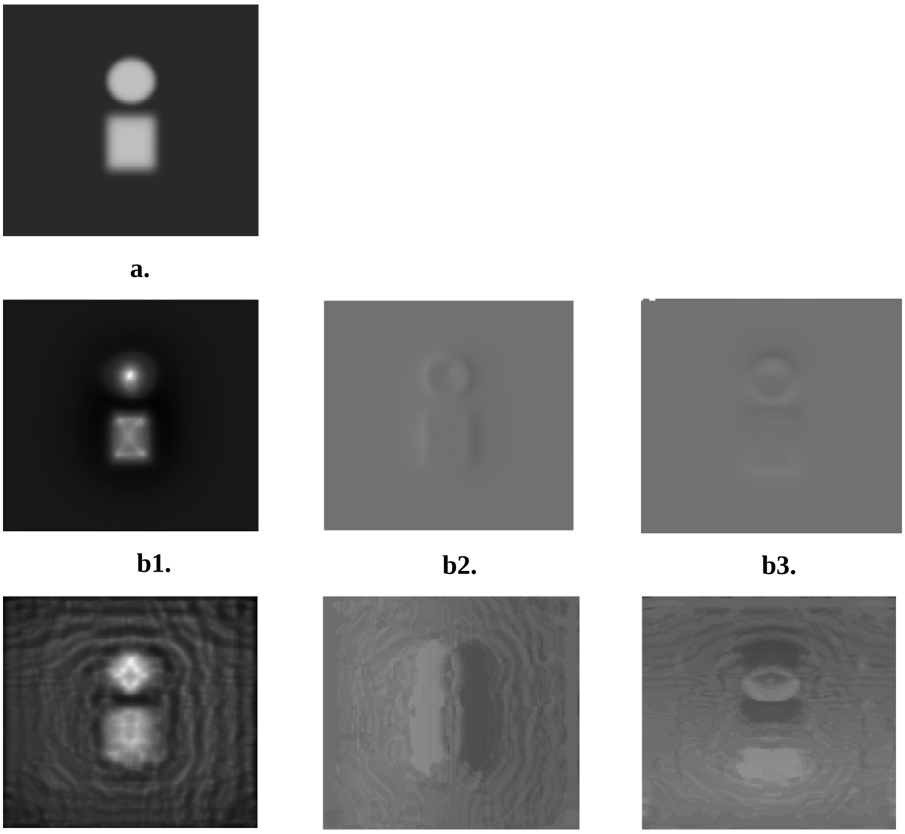
a.

b1. b2 b3

c1. c2. c3.

Fig.3.5. Optimization results for an annular target profile: target function (**a**); solution and transform components obtained by the gradient projection method (**b1-b3**); solution and transform components obtained by the conditional gradient method (**c1-c3**).

We also studied possibilities to model complex structures containing both rectangular and circular regions. In this case, the conditional gradient method shows better results. For the example described below, this method yielded a 65.2% target functional decrease, while the gradient projection method yielded 61.5%.



a.

b1. b2. b3.

c1. c2. c3.

Fig.3.6. Optimization results for a target profile consisting of a rectangular and a circular region: target function (**a**); solution and transform components obtained by the gradient projection method (**b1-b3**); solution and transform components obtained by the conditional gradient method (**c1-c3**).

The parameters used for the simulations, the results of which are shown in Fig.3.6, were the following:

$D = 0.01, \quad K = 3.6, \quad \gamma = 0.8,$

$$u_0(x_1, x_2) = 2 + \exp(-50(x_1{}^2 + (0.1x_2 - \pi/3)^2)(x_1{}^2 + (x_2 - \pi/3)^2)^4) + \exp(-18x_1{}^6 - 5(x_2 + \pi/5)^6)$$

$\mathbf{g}^0(\mathbf{x}) = \mathbf{x}, \; T = 2,$

$$u_1(x_1, x_2, t) = 2 + 0.5t(\exp(-50(x_1{}^2 + (0.1x_2 - \pi/3)^2)(x_1{}^2 + (x_2 - \pi/3)^2)^4) + \exp(-18x_1{}^6 - 5(x_2 + \pi/5)^6)) .$$

Although the solution obtained by the conditional gradient method better reproduces the shape of the target profile, this solution is well perturbed (like in the previous example).

2.3.8. Testing special gradient projection method in contrast with the simplified method

Example 1. In the numerical simulation series discussed below, the target function was defined by the following formula:

$$u_1(x_1,x_2,t) = 2 + \exp(-((x_1 - \pi)^2 + (x_2 - \pi)^2))\cos((x_1 - \pi) - (x_2 - \pi)((x_1 - \pi)^2 + (x_2 - \pi)^2)^2) .$$

A plot of the target phase distribution for the moment $t = T$ is shown in **Fig. 3.7.**

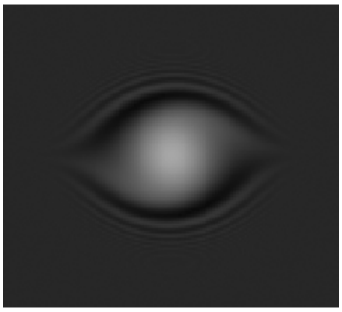


Fig. 3.7. Target function $u_1(x_1,x_2,t=T)$

The parameters used for the simulations, the results of which are shown in Fig.3.8, were the following:
$u_1(x_1,x_2,t)=2+\exp(-((x_1-\pi)^2+(x_2-\pi)^2))\cos((x_1-\pi)-(x_2-\pi)((x_1-\pi)^2+(x_2-\pi)^2)^2)$,
$u_0(x_1,x_2)=2, D=0.01, \ K=3.2, \ \gamma=0.7, \ \mathbf{g}^0(\mathbf{x})=\mathbf{x}, \ T=4$.

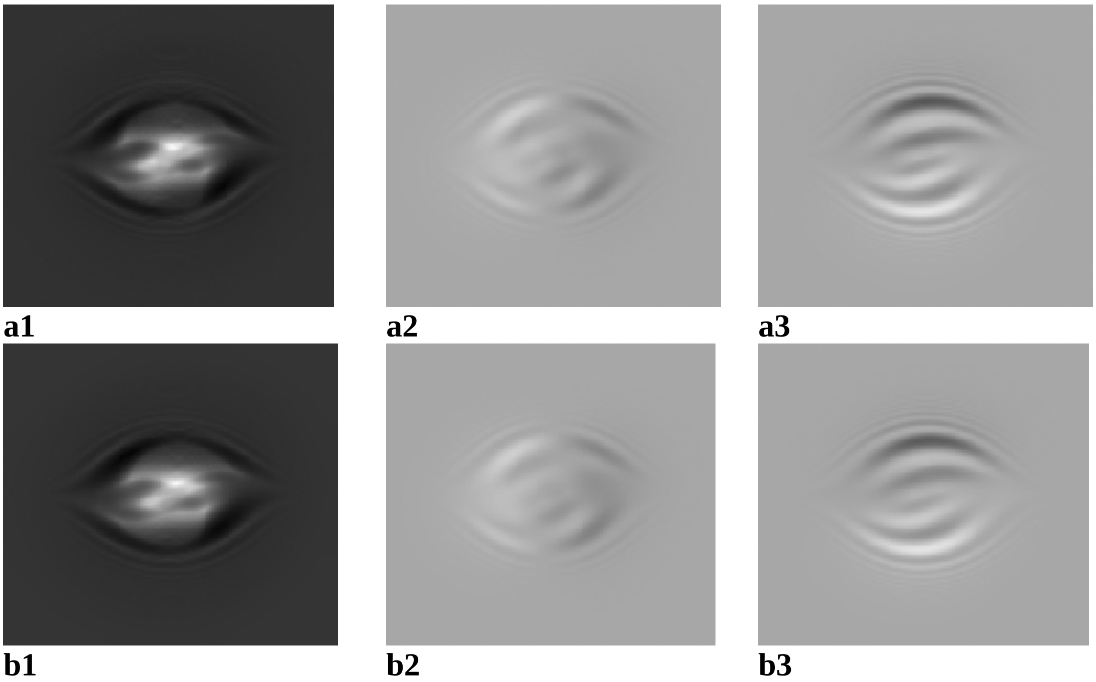


Fig. 3.8. Optimization results: simplified projection gradient method (**a1-a3**) and special projection gradient method (**b1-b3**).

In this series of numerical simulations, both methods showed similar results at different values of the step splitting factor of the gradient method (the system parameters being the same). After four steps of the gradient method, the classical gradient projection method yielded a 71,48% decrease of the, while the modified method ensured a 71,65% decrease. As it can be seen from **Fig. 3.8,** the results are hardly distinguishable visually**.**

Example 2. In the next series of numerical simulations, the target function was defined as $u_1(x_1,x_2,t)=2.2+0.22\sin((x_1-\pi)+1.4\cos(3(x_2-x_1)))$. A plot of the target phase profile is shown in **Fig.3.9.**

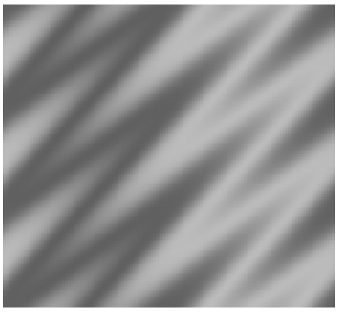
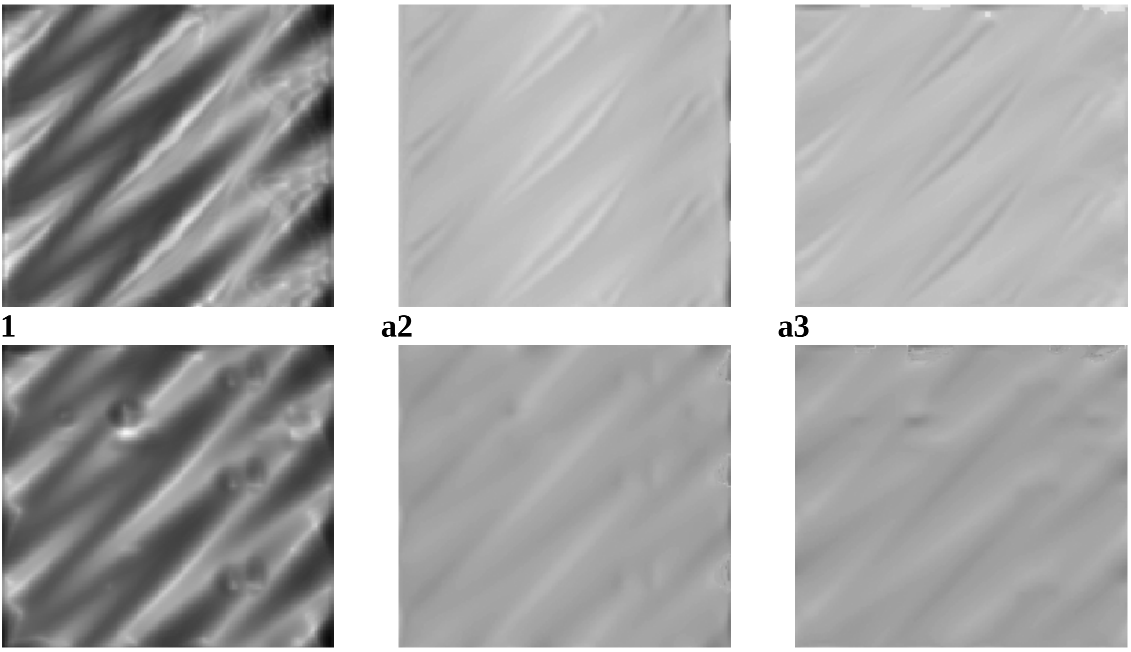


Fig.3.9. Target function $u_1(x_1,x_2,t=T)$

The parameters used for the simulations, the results of which are shown in Fig.3.10, were the following:
$D=0.05,\ \ K=4,\ \ \gamma=1,\ \ u_0(x_1,x_2)=2.2{+}0.22\sin((x_1-\pi)+1.4\cos(3(x_2-x_1))),\ \mathbf{g}^0(\mathbf{x})=\mathbf{x},\ T=4$,
$u_1(x_1,x_2,t)=2.2{+}0.22\sin((x_1-\pi)+1.4\cos(3(x_2-x_1)))$.



a1 **a2** **a3**

b1 **b2** **b3**

Fig. 3.10. Optimization results: simplified projection gradient method (**a1-a3**) and special projection gradient method (**b1-b3**).

After 9 iterations, the simplified method yielded a 52,98% decrease of the target functional with respect to its value before the optimization. At the same time, the special method was capable of decreasing the functional by only by 40,79%, after 3 iterations. Besides, in this case the special method appeared to be less flexible: with its use, the optimization results remained almost the same regardless of the step splitting, while with the simplified method allowed us to improve the optimization results by step splitting.

Example 3. In this series of numerical simulations, the target function was defined as
$u_1(x_1,x_2,t)=1.8+0.3\exp(\sin((x_1-\pi)(x_2-\pi)(5(x_1-\pi)-4(x_2-\pi))+(x_1-\pi)^2+(x_2-\pi)^2)-1)$.
A plot of the target phase profile is shown in **Fig.3.11**:

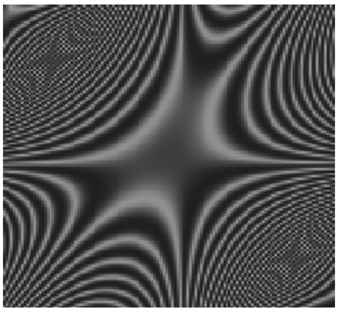


Fig.3.11. Target function $u_1(x_1,x_2,t=T)$

The parameters used for the simulations, the results of which are shown in Fig.3.12, were the following:

$$D = 0.02,\ \ K = 3.2,\ \ \gamma = 1,\ \ u_0(x_1, x_2) = 2,\ \mathbf{g}^0(\mathbf{x}) = \mathbf{x},\ \ T = 4,$$
$$u_1(x_1,x_2,t) = 1.8 + 0.3\exp(\sin((x_1-\pi)(x_2-\pi)(5(x_1-\pi)-4(x_2-\pi)) + (x_1-\pi)^2 + (x_2-\pi)^2) - 1).$$

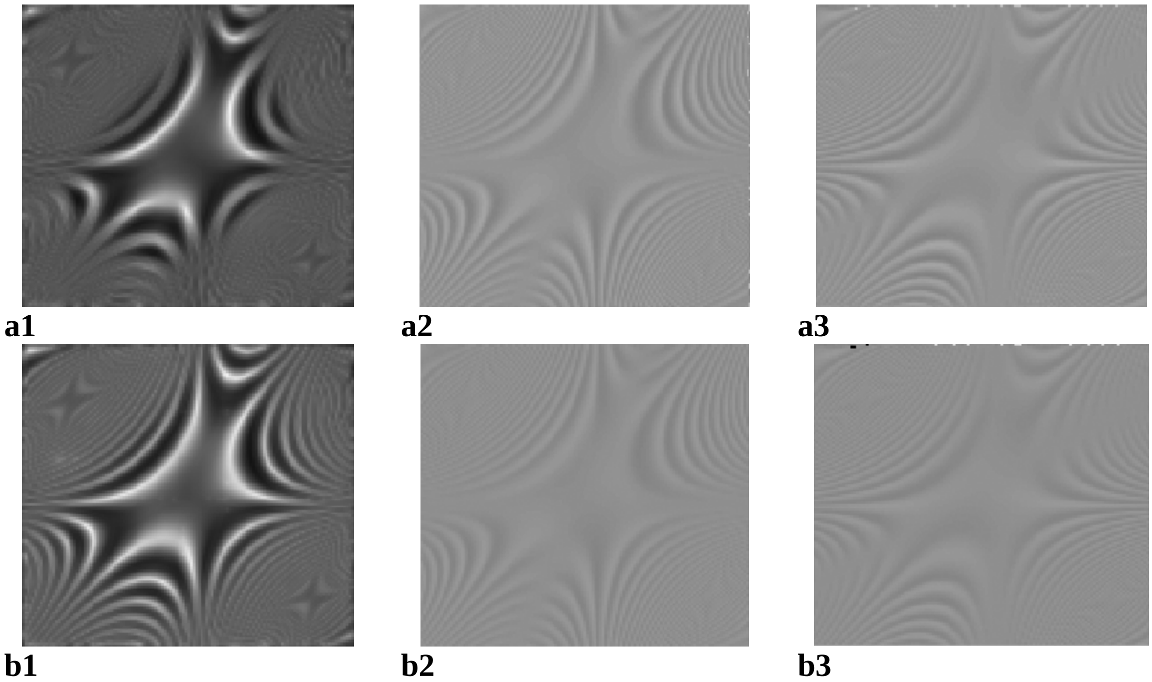


Fig. 3.12. Optimization results: simplified projection gradient method (**a1**-**a3**) and special projection gradient method (**b1**-**b3**).

In this series of numerical simulations, the special method shoed better performance, yielding a 41,84% decrease of the target functional after only 2 iterations. The same 2 iterations performed for the simplified method yielded only a 21,27% decrease of the functional. Visually comparing the results, one may notice that the special method provides better resolution of fine details (see **Fig.3.12b1**), while the simplified method blurred these (see **Fig.3.12a1**). .

The results of the above simulations demonstrate that no universal gradient-type method obviously overscoring the other methods does exist. The special gradient projection method is theoretically convergent regardless of the initial iteration and is more formalized. However, it has fewer fitting parameters and therefore is less flexible. On the contrary, no convergence is theoretically guaranteed for the simplified method. However, the wide possibilities of step splitting allow one to achieve good results with this method.

1.0.0. Static and iterative Tikhonov regularization methods.

In the studies of the conditional gradient method, the shape of the solutions was perturbed, especially within constant areas. To obtain more smooth solutions, we propose to employ regularization. To investigate its capabilities, we used a regularized version of the conditional gradient method. The functional to minimize is the Tikhonov functional

$$T_k(\mathbf{g}) = J(\mathbf{g}) + \alpha_k \|\mathbf{g}\|^2_{L^2(\Omega)} \to \inf_G \,, \qquad (3.14)$$

while for its gradient the following formula is valid: $T_k'(\mathbf{g}) = J'(\mathbf{g}) + 2\alpha_k \mathbf{g}$, where the target functional gradient is calculated according to (3.7) with the nonlinearity $F(u) = K(1+\gamma \cos u)$, which corresponded to the nonlinear interferometer model, while the computation time interval was $T = 1.5$ (in the units of the response time of the nonlinear medium).

First, we employed **static regularization** with the regularization parameter α_k being constant, regardless of the iteration number. These studies showed that the solution obtained with the regularization method has a more smooth profile. However, such regularization reduced the "flexibility" of the gradient method, that is, increased the number of the steps required and decreased the relative reduction of the functional (relatively to the case of no regularization).

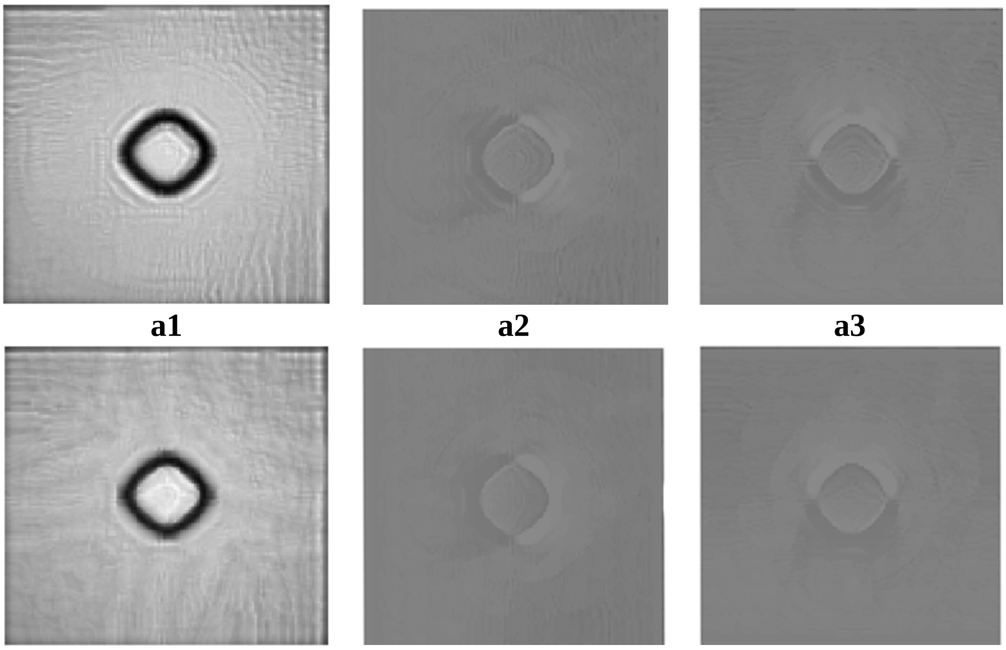


Fig. 3.13. Solution and transform components obtained by the conditional gradient method without regularization (**a1-**a**3**) and with regularization (**b1-b3**).

The parameters used for the simulations, the results of which are shown in Fig.3.13, were the following:

$D = 0.02,\ \ K = 4.2,\ \ \gamma = 0.9,\ \ u_0(x_1, x_2) = 2.2\,,$

$u_1(x_1, x_2, t) = 2.2 + 0.8(\exp(-15(x_1^{\ 4} + x_2^{\ 4} + 4x_1^{\ 2}x_2^{\ 2})^2) + \exp(-5(x_1^{\ 4} + x_2^{\ 4} + 4x_1^{\ 2}x_2^{\ 2})^2)\,.$

In the case of no regularization, 14 steps of the gradient method were made, which led to 27% reduction of the target functional. When we "turn on" the regularization, the functional value decreased by 23.75% after 9 steps. Such a regularity can be explained by the fact that the functional contains an additional summand whose contribution to the whole functional increases as the solution approaches the target function. This reduces the possibilities of perfectly matching the solution with the target profile.

Some numerical studies were performed in order to reveal the influence of this summand (namely – the regularization parameter) on the matching quality. The results of these studies showed that increasing the regularization parameter decreased both the number of the gradient method steps and the matching quality. The latter fact shows up in the relative decrease of the functional value, as it is illustrated in the table below.

Functional decreasing, %	Regularization parameter	Number of steps
32,16	0,00005	7
30,94	0,0001	6
26,07	0,0005	4

Tab.3.1. Influence of the static regularization parameter on the optimization capabilities.

The results contained in this table correspond to the following parameters of the model:

$$D = 0.007, \quad K = 4, \quad \gamma = 1,$$

$$u_0(x_1, x_2) = 2 +$$

$$+ \exp(-9(x_1^2 + (0.1x_2^2 + \pi/7)^2 (x_1^2 + (x_2^2 + \pi/7)^2 (x_1^2 + (x_2^2 + \pi/7)^2)(x_1^2 + (x_2^2 + \pi/7)^2))))$$

$$u_1(x_1, x_2, t) = 2 +$$

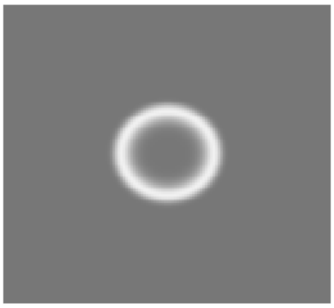
$$+ \exp(-9(x_1^2 + (0.1x_2^2 + \pi/7)^2 (x_1^2 + (x_2^2 + \pi/7)^2 (x_1^2 + (x_2^2 + \pi/7)^2)(x_1^2 + (x_2^2 + \pi/7)^2)))) .$$

To solve the optimization problem more efficiently, **iterative regularization** was employed. As the iterative sequence, we chose the following one: $\alpha_k = a_0(k+1)^{-\gamma}$, where k is the number of the current step of the gradient method, while a_0 and γ are parameters of the numerical experiment. Thus, the influence of the additional summand decreases as the step number (iteration number) increases. The below example, which corresponds to the following values of the parameters

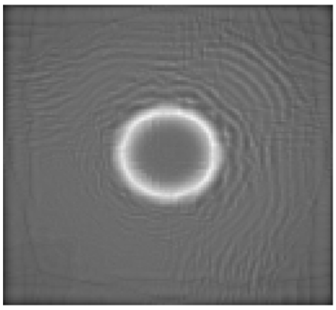
$$D = 0.02, \quad K = 3.6, \quad \gamma = 1, \quad u_0(x_1, x_2) = 2.02 + (x_1^2 + x_2^2)^3 \exp(-2 \cdot (x_1^2 + x_2^2)^4) ,$$

$$u_1(x_1, x_2, t) = 2.02 + 0.5t(x_1^2 + x_2^2)^3 \exp(-2 \cdot (x_1^2 + x_2^2)^4) ,$$

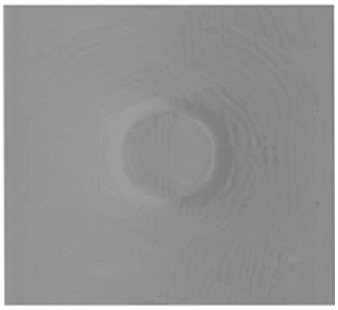
allows one to compare the results of numerical experiments involving no regularization, static regularization, and iterative regularization.



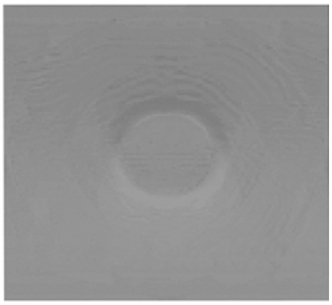
A



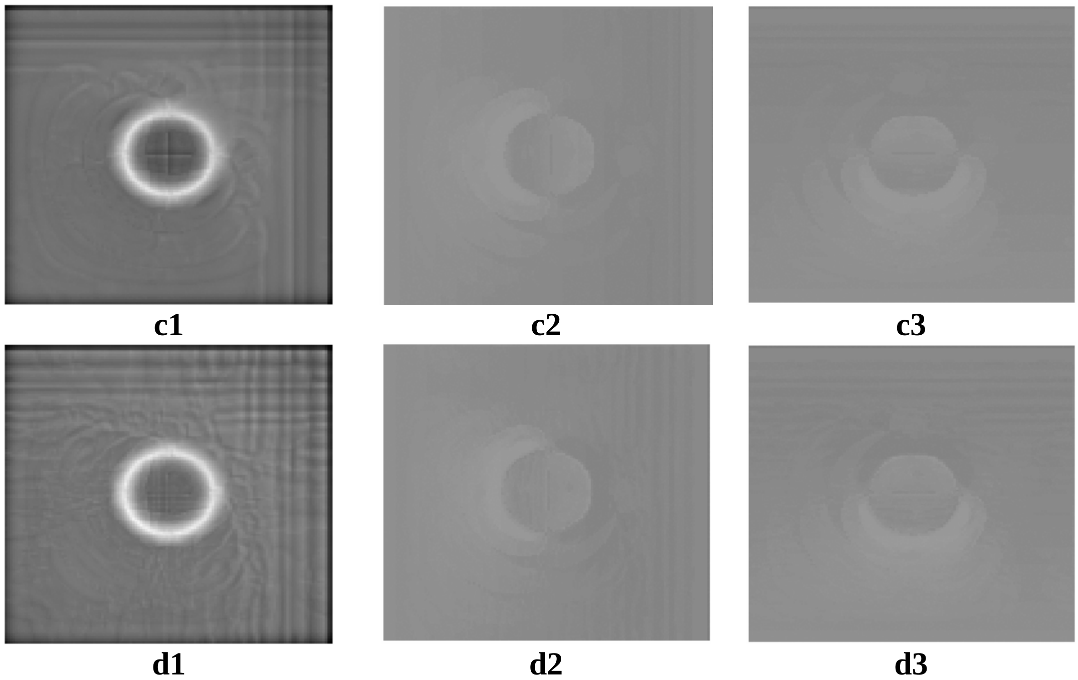
b1



b2



b3



c1 **c2** **c3**

d1 **d2** **d3**

Fig.3.14. Optimization results for an annular target profile: target function (**a**); solution and transform components obtained by the gradient projection method without regularization (**b1-b3**); using static regularization (**c1-c3**); using iterative regularization (**d1-d3)**.

In the case of static regularization, the functional decreased by 26.4%; in the case of iterative regularization – by 50%; in the case of no regularization – by 70.9%.

Thus, in the framework of this direction of investigations, we studied the regularization method aimed at obtaining a more smooth solution. For this purpose, methods of both static and iterative regularization were employed. Our studies showed that although static regularization yielded a smoother solution, the presence of the addition summand in the target functional somewhat reduced the optimization capabilities. In our opinion, the best alternative to the static regularization is iterative regularization that both allows one to reasonably control the contribution of the additional summand and yields a smoothed solution that is more close to the target function than the solution obtained with static regularization.

These studies showed that when we solved the minimization problem using the gradient projection method, the structure of the argument transform being the optimization result was close to the structure of the target function. When the conditional gradient method was used, the resultant argument transform was not too similar to the target profile structure. This was due to the specific character of forming a new approximation in the conditional gradient method. Thus, although the conditional gradient method yields more diverse possibilities to perform the optimization, the target profile structure is reproduced more precisely when using the gradient projection method. To join the potentials of the two methods, we proposed a new problem statement, in which some approximation of the argument transform, $\bar{\mathbf{g}}$, chosen on the basis of ***a priori*** **information**, participated in the stabilizing term $\| \mathbf{g} - \bar{\mathbf{g}} \|^2$ as an additional parameter. Thus, we suggest replacing functional (3.14) with

$$T_k(\mathbf{g}) = J(\mathbf{g}) + \alpha_k \, \| \mathbf{g} - \bar{\mathbf{g}} \|^2 \to \inf_G , \tag{3.15}$$

where $\bar{\mathbf{g}}$ is an approximation of the argument transform, chosen on the basis of *a priori* information. Our primary goal was to figure out how this parameter affects the optimization quality.

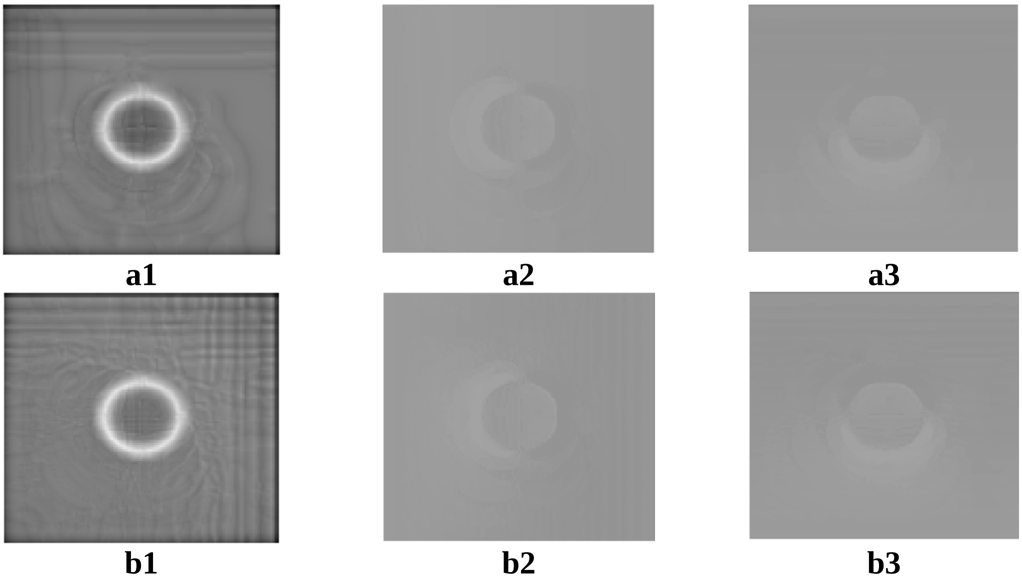
The influence of the parameter $\bar{\mathbf{g}}$ on the optimization quality was studied numerically. The corresponding numerical simulations involved a few stages. First, the problem was solved using the gradient projection method with no regularization. Then, the structure of the argument transform obtained was analyzed in order to build the approximation $\bar{\mathbf{g}}$ being the parameter of the stabilizing term. Afterwards, optimal control problem (3.15) was solved using the conditional gradient method for the case of iterative regularization.

Example 1. The parameters used for the simulations, the results of which are shown in Fig.3.15, were the following: $\mathbf{g}^0(\mathbf{x}) = \mathbf{x}$, $T = 2$.

$D = 0.02$, $K = 3.6$, $\gamma = 1$,

$$u_0(x_1, x_2) = 2.02 + ((x_1 - \pi)^2 + (x_2 - \pi)^2)^3 \exp(-2 \cdot ((x_1 - \pi)^2 + (x_2 - \pi)^2)^4)$$

$$u_1(x_1, x_2, t) = 2.02 + 0.5t((x_1 - \pi)^2 + (x_2 - \pi)^2)^3 \exp(-2 \cdot ((x_1 - \pi)^2 + (x_2 - \pi)^2)^4)$$



a1 a2 a3

b1 b2 b3

Fig.3.15. Optimization results for an annular target profile, obtained by the conditional gradient method: solution and transform components obtained with the additional parameter contained in the stabilization term (a1-a3) and without the additional parameter (b1-b3).

When we used no additional parameter in the stabilizing term, the target functional decreased by 50.1% after 7 steps of the conditional gradient procedure. When the additional parameter was “switched on”, the functional decreased by 74.65% after only 3 steps. The following approximation was taken as the additional parameter:

$$\bar{\mathbf{g}}(x_1) = x_1 + 0.2\sin(x_1)((x_1 - \pi)^2 + (x_2 - \pi)^2)^3 \exp(-2 \cdot ((x_1 - \pi)^2 + (x_2 - \pi)^2)^4)$$

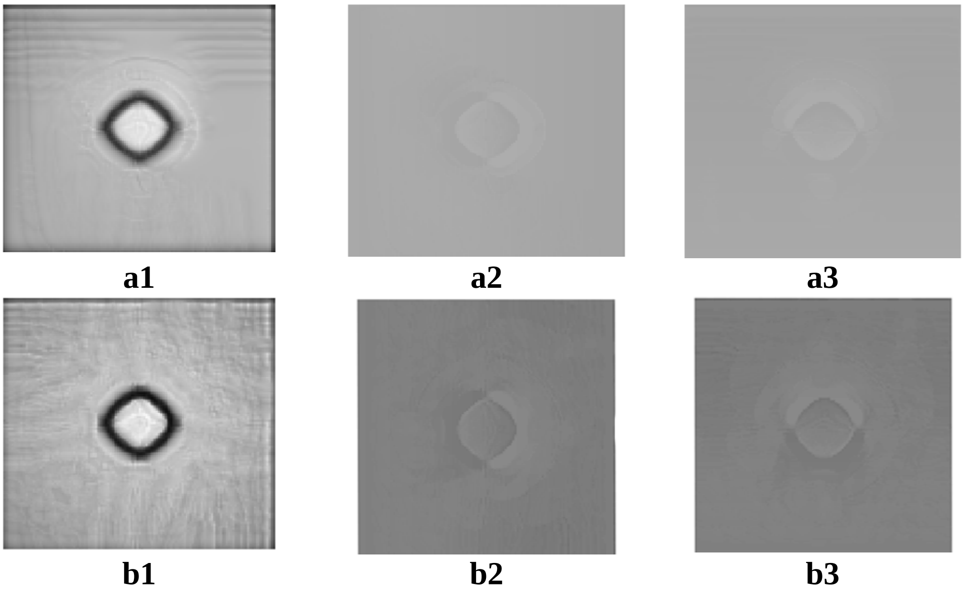
$$\bar{\mathbf{g}}(x_2) = x_2 + 0.2\sin(x_2)((x_1 - \pi)^2 + (x_2 - \pi)^2)^3 \exp(-2 \cdot ((x_1 - \pi)^2 + (x_2 - \pi)^2)^4)$$

Example 2. $\mathbf{g}^0(\mathbf{x}) = \mathbf{x}$, $T = 1.5$, $D = 0.02$, $K = 4.2$, $\gamma = 0.9$, $u_0(x_1, x_2) = 2.2$

$$u_1(x_1, x_2, t) = 2.2 + 0.8(\exp(-15((x_1 - \pi)^4 + (x_2 - \pi)^4 + 4(x_1 - \pi)^2(x_2 - \pi)^2)^2) + \\ + \exp(-5((x_1 - \pi)^4 + (x_2 - \pi)^4 + 4(x_1 - \pi)^2(x_2 - \pi)^2)^2).$$

$$\bar{\mathbf{g}}(x_1) = x_1 + 0.2\sin(x_1)(\exp(-15((x_1 - \pi)^4 + (x_2 - \pi)^4 + 4(x_1 - \pi)^2(x_2 - \pi)^2)^2) + \\ + \exp(-5((x_1 - \pi)^4 + (x_2 - \pi)^4 + 4(x_1 - \pi)^2(x_2 - \pi)^2)^2)$$

$$\bar{\mathbf{g}}(x_2) = x_2 + 0.2\sin(x_2)(\exp(-15((x_1 - \pi)^4 + (x_2 - \pi)^4 + 4(x_1 - \pi)^2(x_2 - \pi)^2)^2) + \\ + \exp(-5((x_1 - \pi)^4 + (x_2 - \pi)^4 + 4(x_1 - \pi)^2(x_2 - \pi)^2)^2)$$



a1 **a2** **a3**

b1 **b2** **b3**

Fig.3.16. Solution and transform components obtained by the conditional gradient method with regularization using the additional parameter in the stabilization term (**a1**-a**3**) and using no additional parameter (**b1-b3**).

Without the additional parameter, the target functional decreased by 23.75% after 9 steps of the conditional gradient procedure. When the additional parameter was “switched on”, the functional decreased by 35.8 % after 8 steps.

In both numerical experiments, a quite simple approximation of the argument transform was used: if $u_1(\mathbf{x},t) = C + f(t)u(\mathbf{x})$, where C is a constant and $f(t)$ is a time-dependent function, then $\overline{g}(\mathbf{x}) = \mathbf{x} + A\sin(\mathbf{x})u(\mathbf{x})$, where the amplitude A is chosen using the information on the transform, obtained with the gradient projection method.

Thus, for the case of localized optical patterns with a fine structure of the solution profile, the proposed method allows one to obtain better result than that obtained with the conditional gradient method, both with regularization and without it. Moreover, the proposed method requires fewer iterations. A quite simple approximation can be taken as the additional parameter $\overline{\mathbf{g}}$. So far, we haven’t obtained comparably good results for spatially distributed structures and localized structures having large spatially uniform areas. This is because a more careful selection of the approximation is required in this case.

3. Results and Conclusions

3.1. In the scope of Task 1, the original projection method of integral Fourier transformation based on expanding the solution into a series of a full system of Fourier transform eigenfunctions (Hermite functions) shows very promising results in image filtering, texture discrimination, and noisy image edge detection. The joint use of this method with Tikhonov regularization algorithms aloows us to construct very stable methods for image processing and analysis.

3.2. In the scope of Task 2, a profound mathematical study has been accomplished for the problem of optimal discrete Fourier filtering. It covers novel results on the existence and uniqueness theorems for both the direct and the conjugate initial-boundary value problems for the functional differential parabolic equation. The solvability of the filter optimization problem has been proved for a Hilbert brick and for a ball of the radius $R > 0$ in the ℓ_2 space as admissible filter sets. The terminal functional gradient formula has been obtained along with the estimation of the remainder term. The projection finite-element scheme has been developed both for direct and conjugate problems in order to figure out the performance of well-known Fourier filters in optical systems aimed at suppressing phase distortions of the input light wave. Conditional gradient iterative procedure has been developed and applied for searching for the optimal Fourier filter to control the feedback optical system.

3.3. In the scope of Task 3, we have suggested and developed the mathematical statement of the optimal control problem for a distributed spatial argument transform. It is applied to study controlling optical systems with nonlocal transforms of a light wave in a feedback loop. The approach developed utilizes a generalized way of determining the argument transform. As an advantage, one can use a wide range of nonsmooth and irreversible argument transforms, and consequently achieve better results with the optimal control. Moreover, this approach allows us to develop projection finite-element methods for approximating both the direct and the conjugate problem in a similar way. As a result, we have elaborated and theoretically analyzed computational versions of the projection gradient and the conditional gradient method for target functional minimization.

6. Future Work Recommended

The efficiency of the original projection method of integral Fourier-Hermite transformation with Tikhonov regularization capability continued increasing as we proceed with the Project. Nevertheless, even the speed of calculations for the projection method was optimized during the Project, we see a principal possibility to construct a faster version of the algorithm. The idea is reducing the volume of calculations using a special grid for expansion coefficients calculation. The grid is to be formed using zeroes of Hermite functions, stored as an additional array. Gibbs (ringing) effect is the next unpleasant property of the oscillatory sets of basic functions. Construction of Gibbs effect suppression algorithms using the Tikhonov regularization method and other postprocessing methods will allow us to enhance the projection method.

The preliminary modeling of the system with an additional integral feedback loop, performed in the scope of the Project, showed efficiency of this system for the task of phase distortion suppression. The combined use of the integral feedback, Fourier filtering control, and argument transform, which were development in the framework of this Project, will allow one to considerably improve the quality of phase distortion suppression and beam profile shaping.

Huge potentials of improving the efficiency of the operation of optical feedback systems can be revealed by studying the time-dependent, continuously-corrected control, which purposefully correct the dynamics of the system in order to improve the control quality. Implementing such an approach, which is widely used in adaptive optics, with the use of time-dependent Fourier filters and nonlocal transforms of spatial arguments will allow considerably increasing the performance of the system and will require special theoretical and numerical studies.

5. Personnel Supported.

Razgulin	Alexander	Vitalievich	Project director, Researcher
Krylov	Andrey	Serdjevich	Researcher
Denisov	Alexander	Mikhailovich	Researcher
Vasiliev	Fedor	Pavlovich	Researcher
Potapov	Mikhail	Mikhailovich	Researcher
Larichev	Andrey	Viktorovich	Researcher
Nikolaev	Ilia	Petrovich	Researcher
Chushkin	Vladimir	Alexandrovich	Researcher
Savvina	Svetlana	Sergeevna	Researcher
Kutovoi	Andrey	Vyacheslavovich	Student
Tsibanov	Vladimir	Nikolaevich	Student

5. Technical Publications.

Papers:

1. A.V. Razgulin and S.S. Savvina. Numerical optimization of the two-dimensional transformation of arguments in the functional-differential diffusion equation // Computational Mathematics and Modeling. 2004. Vol. 15. No 4. P. 333-343.
2. A.M. Denisov, A.S. Krylov., and V.N. Tsibanov. Edge detection method by Tikhonov regularization // Proceedings of 14th International Conference GraphiCon’2004. P. 163-166.
3. A.V. Razgulin and V.A. Chushkin. On the optimal Fourier filtration for a class of models of nonlinear optical systems with feedback // Computational Mathematics and Mathematical Physics. 2004. V. 44. No. 9. P. 1528-1538.
4. V.A. Chushkin. On the solvability of the problem of Fourier filter optimization for one class of models of nonlinear optical systems with feedback // In «Modeling and analysis in decision-making problems». Moscow: A.A. Dorodnitsin Computation Center of RAS, 2004. P. 123-135. (In Russian).
5. V.A. Chushkin. On the attractor of the equation of optical Fourier-filtration // Vestnik Mosk. Universiteta., ser. 15, Computational Mathematics and Cybernetics. 2005. No 2. P. 16-25 (In Russian).
6. A.V. Razgulin. On parabolic functional-differential equations with controlled transforms of spatial arguments // Russian Academy Doklady. 2005. Vol. 403. No. 4 (In Russian).
7. A.V. Razgulin. The control problem of 2D transforms of spatial arguments for parabolic functional-differential equations // Differential Equations (In Russian, submitted).
8. A.V. Razgulin. Projection-difference scheme for parabolic functional differential equation with 2D transform of arguments // Computational Mathematics and Mathematical Physics (submitted).
9. A.V. Razgulin On a gradient projection method for quasi-differentiable functionals with Holder-continuous gradient // Moscow University Computational Mathematics and Cybernetics (submitted).

Abstracts:

1. V.A. Chushkin On optimization of a discrete Fourier filter // Abstracts of XXVI Conference of Young Scientists of Mechanics and Mathematics Faculty of MSU, Moscow, 2004. P. 140 (in Russian).
2. V.A. Chushkin and A.V. Razgulin. On the problem of Fourier filter optimization in a functional-differential diffusion equation // Modern Methods for the Theory of Boundary Value

Problems. Materials of Voronezh Spring School «Pontryagin Readings-XV». Voronezh, VSU, 2004. P. 235 (in Russian).

3. V.A. Chushkin. Optimal Control Problem by Discrete Fourier Filter // Abstracts of 4-th International conference on operations research. Moscow, September 21-24, 2004. Moscow: MAKS Press, 2004. P. 54-58.
4. V.A. Chushkin. On numerical optimization of a Fourier filter in a functional-differential equation of nonlinear optics // Abstracts of XX All-Russian Workshop «Analytical methods and optimization of processes in gas and liquid mechanics », Abrau-Durso, September 4-7, 2004. Novosibirsk: M.A. Lavrentyev Hydrodinamics Institute of SB RAS, 2004. P. 79. (In Russian).
5. V.A. Chushkin, A.V. Larichev., I.P. Nikolaev, and A.V. Razgulin. On application of Fourier filtering for forming a given phase distribution in nonlinear optical systems with feedback // Abstracts of the III International Conference «Basic Problems of Optics ». St. Petersburg. October 18-21, 2004. St. Petersburg: ITMO, 2004. P. 62-64. (In Russian)

6. Interactions/Transitions

One report on National University of Singapore – MSU forum (December, 2003)

A. Razgulin. Nonlinear functional-differential parabolic equations with transformed spatial arguments: stability, control and finite-dimensional approximations

7. Patent Disclosures. N/A

8. Honors. N/A