Volume 14
Number 1

June *2012*
(ISSN 1109-1606)

Journal of APPLIED ELECTROMAGNETISM

JAE

Institute of Communication and Computer Systems

Athens - GREECE

Volume 14 **June 2012**
Number 1 **(ISSN 1109-1606)**

JOURNAL OF APPLIED ELECTROMAGNETISM

Institute of Communication and Computer Systems

Athens - GREECE

Volume 14 **Number 1**

June 2012

TRANS BLACK SEA REGION UNION OF APPLIED ELECTROMAGNETISM (BSUAE)

JOURNAL OF APPLIED ELECTROMAGNETISM

Institute of Communication and Computer Systems

Athens - GREECE

Journal of Applied Electromagnetism

Copyright Form

The undersigned I confirm that I agree the publication of the article

in the Journal of Applied Electromagnetism and the copyright to belong to Trans Black Sea Union of Applied Electromagnetism. I understand that I have the full right to reuse this manuscript for my own purposes.

Name:

Surname:

Address:

E-mail:

Signed:

***Please send the previous form signed either by e-mail to pfrangos@central.ntua.gr , or by fax to the fax number: +30 210 772 2281, attention of Prof. P. Frangos.**

Address

Institute of Communication and Computer Systems

National Technical University of Athens

9, Iroon Polytechniou str.

15773 Athens - GREECE

Tel: (+30) 2107723694

Fax: (+30) 2107722281, attention of Prof. P. Frangos

e-mail: pfrangos@central.ntua.gr

Web site: **http://jae.ece.ntua.gr**

TRANS BLACK SEA REGION UNION OF APPLIED ELECTROMAGNETISM (BSUAE)

JOURNAL OF APPLIED ELECTROMAGNETISM (JAE)

Volume 14 Number 1 June 2012

CONTENTS

The Riemann-Liouville operator for fractional differentiation is applied for a generalization of the Coulomb wave equation, named Fractional Coulomb equation. By means of the Fröbenius' method, the regular solution of the latter is obtained in a power series form. It is shown that the same contains the respective solution of the classical Coulomb equation as a special case.

The electrodynamic theory of the square loop antenna is suggested. In an analytical form its radiation pattern is presented. The problem of electromagnetic compatibility of this antenna is studied. The numerical data are given, and the way of selection of the optimal regime of the antenna functioning is stated.

We demonstrate here the existence of a local Darboux chart for the Manev model such that its dynamics becomes locally equivalent to Kepler model. This explains lot of similarities between these two models and especially why they share common invariants and symmetry algebras. We also discuss the role of a certain discrete symmetry appearing in the case of a bounded motion.

ON THE FRACTIONAL COULOMB EQUATION

Y. Nikolova[1], L. I. Boyadjiev[1], G. N. Georgiev[2], and M. N. Georgieva-Grosse[3]

[1] Department of Applied Mathematics and Informatics,
Technical University of Sofia, BG–1000 Sofia, Bulgaria
e-mail: jvr@abv.bg

[2] Faculty of Mathematics and Informatics, University of Veliko Tirnovo
"St. St. Cyril and Methodius", BG–5000 Veliko Tirnovo, Bulgaria

[3] Consulting and Researcher in Physics and Computer Sciences,
Meterstrasse 4, D–70839 Gerlingen, Germany

Abstract

The Riemann-Liouville operator for fractional differentiation is applied for a generalization of the Coulomb wave equation, named Fractional Coulomb equation. By means of the Fröbenius' method, the regular solution of the latter is obtained in a power series form. It is shown that the same contains the respective solution of the classical Coulomb equation as a special case.

1. INTRODUCTION

Fractional are called differential equations which involve the Riemann-Liouville fractional derivative ${}_0D_x^{\lambda}$ of order λ, where λ is a positive real number [1,2]. They are an object of the Fractional Calculus, a branch of the mathematical analysis, dealing with differentiation and integration of an arbitrary order [1-4]. These equations turned out to be a suitable means for investigating various tasks of physics, electrochemistry and mathematics [1-7].

In this study the notion *Fractional Coulomb equation* is advanced, as an equation of the aforesaid type that contains the derivative ${}_0D_x^{2\lambda}$ of order 2λ, with $0<\lambda\le 1$. It is considered as a fractional generalization of the Coulomb wave equation [8-11], correspondingly of the Kummer and Whittaker confluent hypergeometric ones [8-18]. The Fröbenius' method [19] is applied to determine the fractional analog of the regular solution of the Coulomb wave equation. It is shown that the latter (resp. its solution) is recovered from the fractional one when $\lambda=1$.

An example for putting into practice of the special case of the solution of fractional equation – the one of the classical Coulomb equation, is the problem for propagation of normal TE_{0n} modes in the azimuthally magnetized circular ferrite waveguide [20].

This structure is of particular interest in designing digital nonreciprocal phase shifters for electronically scanned antenna arrays, operating in the normal TE_{01} mode [21-34].

2. KUMMER AND WHITTAKER CONFLUENT HYPERGEOMETRIC EQUATIONS

The equation [8,9,11-16,18,19,21,24]:

$$x\frac{d^2y}{dx^2}+(c-x)\frac{dy}{dx}-ay=0 \tag{1}$$

is called confluent hypergeometric or Kummer equation. It possesses a regular and irregular singularities at $x=0$ and $x=\infty$, resp. By means of the power series method (the Fröbenius' method), its two solutions in the neighbourhood of the first of these points [8,9,11-16,18,19]:

$$y_1=\Phi(a,c;x) \tag{2}$$

and

$$y_2=x^{1-c}\Phi(a-c+1,2-c;x) \tag{3}$$

can be found. $\Phi(a,c;x)$ is known as confluent hypergeometric or simply Kummer function. It is defined for all real or complex a, c and x, except $c=0,-1,-2,...$ through the expression [8,9,11-16,18,19,21,24]:

$$\Phi(a,c;x)=\sum_{\nu=0}^{\infty}\frac{(a)_\nu}{(c)_\nu}\frac{x^\nu}{\nu!} \tag{4}$$

in which $(a)_\nu$ denotes the Pochhammer symbol [8,9,11-16,18,19,21,24]:

$$(a)_\nu=\frac{\Gamma(a+\nu)}{\Gamma(a)}=a(a+1)...(a+\nu-1) \tag{5}$$

with $\Gamma(a)$ – the Euler gamma function [8,9,11-16,18,19]. $\Phi(a,c;x)$ is regular at zero, entire and single-valued (analytic in the whole x – plane). Because it might be considered as a special case of the generalized hypergeometric function ${}_2F_1(a,b;c;x)$, it is often written in the form ${}_1F_1(a,c;x)$ [8]. Other notations used for it are $M(a,b,x)$ and $F(\alpha,\beta,x)$, as well [9,19].

Obviously, y_2 has sense, if $2-c\neq 0,-1,-2,...$, i.e., if $c\neq 2,3,4,...$ For $c=1$, y_1 and y_2 coincide. Thus, the solutions presented are linearly independent on condition that c is different from a positive, negative integer or zero. When this parameter acquires

the values pointed out, the general integral of Eq. (1) instead by y_1 and y_2 could be constructed in terms of: *i*) y_1 and the Tricomi confluent hypergeometric function $\Psi(a,c;x)$, stipulating that $a \neq 0,-1,-2,...$; *ii*) y_1 and $e^x\Psi(c-a,c;-x)$, provided $c-a \neq 0,-1,-2,...$; *iii*) y_2 and $\Psi(a,c;x)$, assuming that $c-a \neq 1,2,3,...$; *iv*) y_2 and $e^x\Psi(c-a,c;-x)$, when $a \neq 1,2,3,...$ and *v*) $\Psi(a,c;x)$ and $e^x\Psi(c-a,c;-x)$ for all a, c and $c-a$ without any restrictions [8,9,13,15,16,18]. Besides, it should be noted that if $c=n+1$, $n=0,1,2,...$, the logarithmic representation of $\Psi(a,c;x)$ is employed [8,9]. Further, on the understanding that $c=1-n$, the formula [9]:

$$\Psi(a,1-n;x)=x^n\Psi(a+n,1+n;x) \tag{6}$$

reduces the debate to the previous case.

Setting $y=e^{x/2}x^{-c/2}w(x)$ in the Kummer equation yields [8,12,13,17,18]:

$$\frac{d^2w}{dx^2}+\left[-\frac{1}{4}+\left(\frac{c}{2}-a\right)\frac{1}{x}+\frac{c}{2}\left(1-\frac{c}{2}\right)\frac{1}{x^2}\right]w=0\,. \tag{7}$$

With the help of the simple substitutions: $c=2m+1$ and $a=(1/2)+m-\kappa$, the same is transformed into [8-13,17,18]:

$$\frac{d^2w}{dx^2}+\left[-\frac{1}{4}+\frac{\kappa}{x}+\frac{\frac{1}{4}-m^2}{x^2}\right]w=0\,, \tag{8}$$

termed Whittaker equation. Bearing in mind the above discussion, in view of the relation between Eqs. (1) and (8) it follows that, if $2m$ is not zero or an integer, as two linearly independent solutions of the latter at the point $x=0$, the functions:

$$M_{\kappa,m}(x)=x^{\frac{1}{2}+m}e^{-\frac{x}{2}}\Phi\left(\frac{1}{2}+m-\kappa,1+2m;x\right) \tag{9}$$

and

$$M_{\kappa,-m}(x)=x^{\frac{1}{2}-m}e^{-\frac{x}{2}}\Phi\left(\frac{1}{2}-m-\kappa,1-2m;x\right) \tag{10}$$

could be chosen [11-13,17,18]. When the opposite holds, the complete integral of the equation regarded, as a substitute of the pair $M_{\kappa,m}(x)$ and $M_{\kappa,-m}(x)$, the next couples of functions are used: *i*) $M_{\kappa,m}(x)$ and $W_{\kappa,m}(x)$, provided $2(m-\kappa)\neq -1,-3,-5,...$; *ii*) $M_{\kappa,m}(x)$ and $W_{-\kappa,m}(-x)$, when $2(m+\kappa)\neq -1,-3,-5,...$; *iii*) $M_{\kappa,-m}(x)$ and $W_{\kappa,m}(x)$, if $2(m+\kappa)\neq 1,3,5,...$; *iv*) $M_{\kappa,-m}(x)$ and $W_{-\kappa,m}(-x)$, presuming that

$2(m-\kappa)\neq 1,3,5,...$; *v*) $W_{\kappa,m}(x)$ and $W_{-\kappa,m}(-x)$, without any limitations for the numerical equivalents of κ, m ($2(m\pm\kappa)$) and x [8,12,13,17,18]. In the lieu of $W_{\kappa,m}(x)$ and $W_{-\kappa,m}(-x)$, the functions $W_{\kappa,-m}(x)$ and $W_{-\kappa,-m}(-x)$, resp. can be employed, as well [18]. $M_{\kappa,m}(x)$ and $W_{\kappa,m}(x)$ are named Whittaker first and second function. They are multiple-valued in the complex x – plane with the origin as a branch point. Their principle branch is determined by the condition $-\pi<\arg x<\pi$ [8,9,12,13]. $W_{\kappa,m}(x)$ is connected with $\Psi(a,c;x)$ through the relation [8,18]:

$$W_{\kappa,m}(x)=x^{\frac{1}{2}+m}e^{-\frac{x}{2}}\Psi\left(\frac{1}{2}+m-\kappa,1+2m;x\right). \tag{11}$$

In case $2m=0,\pm1,\pm2,...$ $W_{\kappa,m}(x)$ is taken in its logarithmic form. $\Phi(a,c;x)$ and $M_{\kappa,m}(x)$ are finite at zero, while $\Psi(a,c;x)$ and $W_{\kappa,m}(x)$ tend to infinity, if $x\to 0$ [8,9,12,13].

According to Whittaker and Watson [17], the greater symmetry of $M_{\kappa,m}(x)$ and $W_{\kappa,m}(x)$ is their main advantage. Really, the various partial integrals of Eq. (8) are obtained from each other just by changing the signs of parameters and variable. Further, the graphically presented results for $M_{\kappa,m}(x)$ in case $m=1$ $\kappa=jk$, $k=0,\pm0.1,\pm0.3,\pm0.5$ and $x=jz$, z – real, positive, have shown that $\mathrm{Re}\,M_{\kappa,m}(x)=\mathrm{Im}\,M_{\kappa,m}(x)$ [21]. Simultaneously, though not possessing the property mentioned, $\Phi(a,c;x)$ and $\Psi(a,c;x)$ are also attractive in the applications, especially from a computation point of view, since the expressions, yielding them are much simpler and $\Phi(a,c;x)$ is single-valued.

3. COULOMB WAVE EQUATION

The equation:

$$\frac{d^2v}{d\rho^2}+\left[1-\frac{2\eta}{\rho}-\frac{L(L+1)}{\rho^2}\right]v=0 \tag{12}$$

in which ρ and η are real, $\rho>0$, $-\infty<\eta<\infty$ and L is a non-negative integer ($L=0,1,2,...$), is referred to as a Coulomb wave equation [8-10]. It has a regular singularity with indexes $L+1$ and $-L$ at $\rho=0$ and an irregular one at $\rho=\infty$. Its general solution is presented as a linear combination of the regular and irregular (logarithmic) Coulomb wave functions $F_L(\eta,\rho)$ and $G_L(\eta,\rho)$, resp. The first of them is given by the expression [9]:

$$F_L(\eta,\rho) = C_L(\eta)\rho^{L+1}\Phi_L(\eta,\rho) \tag{13}$$

where

$$C_L(\eta) = \frac{2^L e^{-\pi\eta/2}\left|\Gamma(L+1+j\eta)\right|}{\Gamma(2L+2)} \tag{14}$$

and

$$\Phi_L(\eta,\rho) = \sum_{q=L+1}^{\infty} A_q^L \rho^{q-L-1}\,. \tag{15}$$

For the coefficients A_q^L it holds:

$$A_{L+1}^L = 1, \tag{16}$$

$$A_{L+2}^L = \frac{\eta}{L+1}\,, \tag{17}$$

$$(q+L)(q-L-1)A_q^L = 2\eta\, A_{q-1}^L - A_{q-2}^L,\ (q > L+2)\,. \tag{18}$$

The series representation (13) provides a convenient way for numerical evaluation of $F_L(\eta,\rho)$. Putting $x = 2j\rho$, $c = 2L+2$, $a = L+1-j\eta$ or $x = 2j\rho$, $\kappa = j\eta$, $m = L+1/2$, the Coulomb equation is reduced to the Kummer or Whittaker one. Accordingly, another option for computation of the solution considered of Eq. (12), is to apply its expansion in terms of the Kummer or Whittaker confluent hypergeometric function [10]:

$$F_L(\eta,\rho) = C_L(\eta)\rho^{L+1}e^{-j\rho}\,\Phi(L+1-j\eta,\,2L+2;\,2j\rho)\,, \tag{19}$$

$$F_L(\eta,\rho) = \frac{1}{2}\left|\frac{\Gamma(L+1\pm j\eta)}{\Gamma(2L+2)}\right| e^{-\frac{1}{2}j\pi(L+1-j\eta)}\, M_{j\eta,L+\frac{1}{2}}(2j\rho)\,. \tag{20}$$

The relation (19) can be used as a definition of the regular Coulomb wave function in the sense of Thompson and Barnett [35], valid for all complex ρ, η and L. In particular, it allows to determine the same, if ρ and η are real, $\rho > 0$, $-\infty < \eta < \infty$ and $L = \pm 0.5$. The numerical analysis [22] has shown that for the values of parameters indicated $F_L(\eta,\rho)$ can be reckoned in two ways: *i*) using the representation (13), though initially defined for $L = 0,1,2,...$ and *ii*) employing Eq. (19). It turned out that the first approach is preferable, since: *i*) the power series operates with real quantities; *ii*) the variable ρ is twice smaller than the argument of Kummer function; *iii*) the power series (15) is more rapidly converging than the Kummer one, especially for large arguments [22]. Graphs, based on the numerical

analysis, illustrating the behaviour of $F_L(\eta,\rho)$ for the case considered have been depicted in Ref. [22].

The notions "*Coulomb wave equation*" and "*Coulomb wave functions*", however, are not unique. The theory presented follows the approach of M. Abramowitz [9] and is based also on the well-known survey of the same by M.K. Kerimov [10]. In a comprehensive study of the problem [10] Curtis revealed that beside Eq. (12) and the functions $F_L(\eta,\rho)$ and $G_L(\eta,\rho)$ there exist also other representations of the same. Such is the equation [10]:

$$\frac{d^2Y}{dt^2}+\left[b+\frac{2}{t}-\frac{L(L+1)}{t^2}\right]Y=0, \tag{21}$$

investigated by Miller, on condition that t is real, positive, b is real and L equals zero or a positive integer who integrated it in terms of the functions: $P_L(b,t)$ and $Q_L(b,t)$ when $b>0$ and $U_L(b,t)$ and $V_L(b,t)$ when $b<0$. Instead of the second couple Jeffreys and Jeffreys introduced the one $U(\alpha,\gamma,Z)$ and $V(\alpha,\gamma,Z)$ [10]. Hartree considered the wave equation of the form [10]:

$$\frac{d^2P}{d\sigma^2}+\left[\frac{1}{\sigma}-\frac{1}{4n^2}-\frac{L(L+1)}{\sigma^2}\right]P=0 \tag{22}$$

and found its two solutions: $G_L(\sigma)$ and $H_L(\sigma)$.

Treating a more universal equation, similar to the Whittaker one and its solution, written in a general form, L.J. Slater reduced it in a special case to eq. (12) and presented the solutions of the latter as [8]:

$$v_1=(2j\rho)^{L+1}e^{-j\rho}\,\mathcal{F}(L+1-j\eta\,,\,2L+2;\,2j\rho), \tag{23}$$

$$v_1=\mathcal{M}_{j\eta,L+\frac{1}{2}}(2j\rho). \tag{24}$$

Here the symbols $\mathcal{F}$ and $\mathcal{M}$ are used to denote any general integral of the Kummer and Whittaker equation, resp. Usually however, she continued, as such are taken the following two Coulomb functions [8]:

$$F_L(\eta,\rho)=\frac{A}{\Gamma(2L+2)}M_{j\eta,L+\frac{1}{2}}(2j\rho), \tag{25}$$

$$G_L(\eta,\rho)=BW_{j\eta,L+\frac{1}{2}}(2j\rho)+\frac{C}{\Gamma(2L+2)}M_{j\eta,L+\frac{1}{2}}(2j\rho) \tag{26}$$

where the constants A, B and C are determined through the relations [8]:

$$A = -jC = \frac{1}{2}|\Gamma(L+1\pm j\eta)|\exp\left[-\frac{1}{2}\pi\eta - \frac{1}{2}\pi j(L+1)\right], \tag{27}$$

$$B = \exp\left(-j\delta + \frac{1}{2}\pi\eta + \frac{1}{2}\pi jL\right) \tag{28}$$

in which

$$e^{\pm j\delta}|\Gamma(L+1\pm j\eta)| = \Gamma(L+1\pm j\eta). \tag{29}$$

Here $\eta > 0$, $L = 0,1,2,\ldots$. Buchholz represented Eqs. (25) and (26) through the introduced by him new function [12]:

$$\mathcal{M}_{\kappa,m}(x) = \frac{M_{\kappa,m}(x)}{\Gamma(1+2m)}, \tag{30}$$

called after his name ($m = L+1/2$). Accordingly, he wrote [12]:

$$F_L(\eta,\rho) = A\,\mathcal{M}_{j\eta,L+\frac{1}{2}}(2j\rho), \tag{31}$$

$$G_L(\eta,\rho) = BW_{j\eta,L+\frac{1}{2}}(2j\rho) + A\,\mathcal{M}_{j\eta,L+\frac{1}{2}}(2j\rho). \tag{32}$$

Considering the radial part of the Schrödinger equation for the electron in a central symmetric potential field, Morse and Feshbach have shown that it reduces to Eq. (1) with $a = L+1-1/\mathrm{K}$, $c = 2L+2$ and $x = 2\mathrm{K}s$. They suggested the finite at $s = 0$ solution of the latter [14]:

$$C_L(\mathrm{K},s) = \frac{\sqrt{\pi}(2\kappa s)^L e^{-\mathrm{K}s}}{2^{2L+1}\Gamma\left(L+\frac{3}{2}\right)}\Phi\left(L+1-\frac{1}{\mathrm{K}}, 2L+2; 2\mathrm{K}s\right) \tag{33}$$

in which K might be real or purely imaginary. When $\mathrm{K} = jK$ (K is real) it holds [14]:

$$\begin{aligned} C_L(jK,s) = (2jKs)^L\Gamma(L+1)\Bigg[&\frac{1}{\Gamma(L+1+(j/K))}U_1\left(L+1+\frac{j}{K}, 2L+2; 2jKs\right) + \\ &\frac{1}{\Gamma(L+1-(j/K))}U_2\left(L+1+\frac{j}{K}, 2L+2; 2jKs\right)\Bigg]e^{-jKs} \end{aligned} \tag{34}$$

where $U_1(L+1+j/K, 2L+2; 2jKs)$ and $U_2(L+1+j/K, 2L+2; 2jKs)$ are expressed by means of y_1 and y_2 [14]. In addition to $C_L(jK;s)$ the function [14]:

$$D_L(jK,s)=\frac{\sqrt{\pi}(2jKs)^L e^{-jKs}}{2^{2L+1}\Gamma\left(L+\frac{3}{2}\right)}G\left(L+1+\frac{j}{K},2L+2;2jKs\right), \tag{35}$$

with $G(L+1+j/K,2L+2;2jKs)$ – the Gordon one has been advanced [14]. $C_L(\mathrm{K};s)$ ($C_L(jK;s)$) and $D_L(jK;s)$ are Coulomb wave functions in the Morse-Feshbach form [14].

F.G. Tricomi termed in the same way the function [16]:

$$H(L,\eta;\rho)=e^{j\rho}\Phi(1+L+j\eta,2+2L;-2j\rho)\ . \tag{36}$$

An application of the first Kummer theorem [9] allows to establish that $H(L,\eta;\rho)$ coincides with its complex conjugate, i.e. it is real and for this reason it does not involve in explicit form the imaginary unit j.

Similarly, rearranging the right-hand side of Eq. (23), it is postulated:

$$T(L,\eta,\rho)=2^{L+1}e^{\frac{1}{2}j\pi(L+1)}\rho^{L+1}e^{-j\rho}\Phi(L+1-j\eta\ ,2L+2;2j\rho) \tag{37}$$

Obviously, it holds:

$$A=2^{-(L+1)}e^{-\frac{1}{2}j\pi(L+1)}\Gamma(2L+2)C_L(\eta)\,, \tag{38}$$

$$F_L(\eta,\rho)=\begin{cases}C_L(\eta)\rho^{L+1}e^{-j\rho}\Phi(L+1-j\eta,2L+2;2j\rho)\\ C_L(\eta)\rho^{L+1}\Phi_L(\eta,\rho)\\ C_L(\eta)\rho^{L+1}H(L,\eta,\rho)\\ 2^{-(L+1)}e^{-\frac{1}{2}j\pi}C_L(\eta)T(L,\eta,\rho)\\ \frac{1}{2}\left|\frac{\Gamma(L+1\pm j\eta)}{\Gamma(2L+2)}\right|e^{-\frac{1}{2}j\pi(L+1-j\eta)}M_{j\eta,L+\frac{1}{2}}(2j\rho),\\ \frac{1}{2}\left|\frac{\Gamma(L+1\pm j\eta)}{\Gamma(2L+2)}\right|e^{-\frac{1}{2}j\pi(L+1-j\eta)}T(L,\eta,\rho)\\ \frac{A}{\Gamma(2L+2)}M_{j\eta,L+\frac{1}{2}}(2j\rho)\\ \frac{A}{\Gamma(2L+2)}T(L,\eta,\rho)\end{cases} \tag{39}$$

$$\Phi_L(\eta,\rho)=\begin{cases}e^{-j\rho}\Phi(1+L-j\eta,2L+2;2j\rho)\\ 2^{-(L+1)}e^{-\frac{1}{2}j\pi(L+1)}\rho^{-(L+1)}M_{j\eta,L+\frac{1}{2}}(2j\rho)\end{cases}, \tag{40a}$$

$$\Phi_L(\eta,\rho)=\begin{cases} C_L(\eta)^{-1}\rho^{-(L+1)}F_L(\eta,\rho) \\ 2^{-(L+1)}e^{-\frac{1}{2}j\pi(L+1)}\rho^{-(L+1)}T(L,\eta,\rho), \\ H(L,\eta,\rho) \end{cases} \quad (40b)$$

$$T(L,\eta,\rho)=\begin{cases} 2^{L+1}e^{\frac{1}{2}j\pi(L+1)}\rho^{L+1}e^{-j\rho}\Phi(L+1-j\eta,2L+2,2j\rho) \\ M_{j\eta,L+\frac{1}{2}}(2j\rho) \\ 2^{L+1}e^{\frac{1}{2}j\pi(L+1)}C_L(\eta)^{-1}F_L(\eta,\rho) \\ 2^{L+1}e^{\frac{1}{2}j\pi(L+1)}\rho^{L+1}\Phi_L(\eta,\rho) \\ 2^{L+1}e^{\frac{1}{2}j\pi(L+1)}\rho^{L+1}H(L,\eta,\rho) \end{cases}, \quad (41)$$

$$M_{j\eta,L+\frac{1}{2}}(2j\rho)=\begin{cases} 2^{L+1}e^{\frac{1}{2}j\pi(L+1)}\rho^{L+1}e^{-j\rho}\Phi(1+L-j\eta,21+2L;2j\rho) \\ 2^{L+1}e^{\frac{1}{2}j\pi(L+1)}\rho^{L+1}H(L,\eta,\rho) \\ 2^{L+1}e^{\frac{1}{2}j\pi(L+1)}\rho^{L+1}\Phi_L(\eta,\rho) \\ T(L,\eta,\rho) \\ 2\left|\dfrac{\Gamma(2L+2)}{\Gamma(L+1\pm j\eta)}\right|e^{\frac{1}{2}j\pi(L+1-j\eta)}F_L(\eta,\rho) \end{cases}, \quad (42)$$

$$\Phi(L+1-j\eta,2L+2;2j\rho)=\begin{cases} 2^{-(L+1)}e^{-\frac{1}{2}j\pi(L+1)}\rho^{-(L+1)}e^{j\rho}M_{j\eta,L+\frac{1}{2}}(2j\rho) \\ C_L(\eta)^{-1}\rho^{-(L+1)}e^{j\rho}\,F_L(\eta,\rho) \\ e^{-j\rho}\,\Phi_L(\eta,\rho) \\ e^{-j\rho}\,H(L,\eta,\rho) \\ 2^{-(L+1)}e^{-\frac{1}{2}j\pi(L+1)}\rho^{-(L+1)}e^{j\rho}T(L,\eta,\rho) \end{cases}. \quad (43)$$

$H(L,\eta,\rho)$ is given by the set of formulae (40a,b), if in the latter this symbol and the one $\Phi_L(\eta,\rho)$ are interchanged wherever they are met. Any of the functions $2^{L+1}e^{(1/2)j\pi(L+1)}\rho^{L+1}e^{-j\rho}\Phi(L+1-\eta,2L+2;2j\rho)$, $M_{j\eta,L+1/2}(2j\rho)$, $\mathcal{M}_{j\eta,L+1/2}(2j\rho)$, $T(L,\eta,\rho)$, $F_L(\eta,\rho)$, $\Phi_L(\eta,\rho)$, $e^{\pm j\rho}\Phi(1+L\pm j\eta,2+2L;\mp 2j\rho)$, $H(L,\eta,\rho)$ and $C_L(jK,s)$ could be called Coulomb wave function in the sense of Slater, Buchholz, Abramowitz, Tricomi, Morse-Feshbach, etc. The first four represent the

same function which differs from $F_L(\eta,\rho)$ by the factor $A/\Gamma(2L+2)$ ($2^{-(L+1)}e^{-(1/2)j\pi(L+1)}C_L(\eta)$), resp. A only. The next three are identical and are distinguished from $F_L(\eta,\rho)$ with the multiple $\rho^{(L+1)}C_L(\eta)$. There are also simple links of a similar nature between the Miller function $P_L(a,x)$, that of Hartree $G_L(\sigma)$ and the Abramowitz one $F_L(\eta,\rho)$ [10]. The correlation between the latter and $C_L(jK;s)$, however, is more complicated.

4. FRACTIONAL CALCULUS

The fractional Calculus (differentiation and integration of an arbitrary order) is nowadays one of the most intensively developing areas of the mathematical analysis. Its areas of putting into practice range from biology through physics and electro-chemistry to economics, probability theory and statistics. The fractional derivatives provide an excellent instrument for the modeling of memory and hereditary properties of various materials and processes. Half-order derivatives and integrals prove to be more advantageous for the formulation of certain electrochemical problems than the classical methods [5]. The modeling of diffusion in a specific type of porous medium is one of the most significant applications of the fractional derivatives [2], [6]. Recently, fractional differentiation and integration operators are also used for extension of the temperature field problem in oil strata [7]. In the special treatises as [1], [3] and [4], the mathematical aspects and employments of the fractional calculus are extensively thrashed out.

In view of the purpose of the study, in this paper the Riemann-Liouville fractional derivative of a function $f(x)$ of order $\lambda > 0$ is adopted, defined by [1]:

$$ {}_0D_x^{\lambda} f(x) = \frac{d^l}{dx^l}\left[{}_0D_x^{-(l-\lambda)} f(x)\right] \tag{44} $$

where $l > 0$ is a positive integer, $l-1 \le \lambda < l$ and

$$ {}_0D_x^{-(l-\lambda)} f(x) = \frac{1}{\Gamma(l-\lambda)}\int_0^x (x-\tau)^{l-\lambda-1} f(\tau)\, d\tau \tag{45} $$

is the Riemann-Liouville fractional integral of $f(x)$ of order $l-\lambda$.

It is worthwhile to mention that in a contrast to the classical calculus, if $\lambda > 0$, $x > 0$ and $\varsigma > -1$, then the fractional derivative of the power function x^{ς} is given by [1]:

$$ {}_0D_x^{\lambda} x^{\varsigma} = \frac{\Gamma(\varsigma+1)}{\Gamma(\varsigma-\lambda+1)} x^{\varsigma-\lambda}, \tag{46} $$

that in the particular case $\varsigma = 0$ and $0 < \lambda < 1$ implies:

$$ {}_0D_x^{\lambda} 1 = \frac{x^{-\lambda}}{\Gamma(1-\lambda)}. \tag{47} $$

5. FRACTIONAL COULOMB EQUATION AND ITS REGULAR SOLUTION

The fractional differential equation of the form

$$ \rho^{2\lambda} {}_0D_{\rho}^{2\lambda} u(\rho) + (\rho^{2\lambda} + g\rho^{\lambda} + h)u(\rho) = 0 \tag{48} $$

is considered in which ρ, λ, g and h are real, $\rho > 0$, $0 < \lambda \le 1$, and ${}_0D_{\rho}^{2\lambda}$ is the Riemann-Liouville fractional derivative of order 2λ. The reason to deal with it is that it appears as a fractional generalization of the Coulomb wave equation. Indeed, if it is set $\lambda = 1$, $g = -2\eta$ and $h = -L(L+1)$, $L = 0,1,2,\ldots$, Eq. (48) reduces to the latter.

Here the discussion is restricted to determining the fractional analog of the regular solution of Eq. (12). For this purpose, the Fröbenius' method is applied to Eq. (48).

Theorem 1: Let $0 < \lambda \le 1$, $\rho > 0$, $g, h \in R$ and δ is such that $\delta - \lambda > -1$ and $\frac{\Gamma(\lambda+\delta+1)}{\Gamma(-\lambda+\delta+1)} + h = 0$. Then the equation (48) is solvable and its solution has the form:

$$ u(\rho) = \sum_{p=1}^{\infty} a_p(\lambda,\delta) \rho^{p\lambda+\delta}, \tag{49} $$

where the coefficients $\alpha_p(\lambda,\delta)$ satisfy the recurrence relations

$$ a_2(\lambda,\delta)\left[\frac{\Gamma(2\lambda+\delta+1)}{\Gamma(\delta+1)} + h\right] + g a_1(\lambda,\delta) = 0 \tag{50} $$

and for $p \ge 3$,

$$ a_p(\lambda,\delta)\left\{\frac{\Gamma(p\lambda+\delta+1)}{\Gamma[(p-2)\lambda+\delta+1]} + h\right\} + g a_{p-1}(\lambda,\delta) + a_{p-2}(\lambda,\delta) = 0. \tag{51} $$

Proof: Following the basic idea of the power series method, the solution of Eq. (48) is searched in the representation (49). Accordingly, inserting this expression in the equation mentioned and taking into account the definition (46), it is obtained:

$$\sum_{p=1}^{\infty} a_p(\lambda,\delta)\frac{\Gamma(p\lambda+\delta+1)}{\Gamma[(p-2)\lambda+\delta+1]}\rho^{p\lambda+\delta}+\sum_{p=3}^{\infty} a_{p-2}(\lambda,\delta)\rho^{p\lambda+\delta}+$$
$$g\sum_{p=2}^{\infty} a_{p-1}(\lambda,\delta)\rho^{p\lambda+\delta}+h\sum_{p=1}^{\infty} a_p(\lambda,\delta)\rho^{p\lambda+\delta}=0 \quad . \tag{52}$$

Rearranging the terms in the above equality, permits to get:

$$\sum_{p=3}^{\infty}\left\{a_p(\lambda,\delta)\frac{\Gamma(p\lambda+\delta+1)}{\Gamma[(p-2)\lambda+\delta+1]}+a_{p-2}(\lambda,\delta)+ga_{p-1}(\lambda,\delta)+ha_p(\lambda,\delta)\right\}\rho^{p\lambda+\delta}+$$
$$\left[a_2(\lambda,\delta)\frac{\Gamma(2\lambda+\delta+1)}{\Gamma(\delta+1)}+ga_1(\lambda,\delta)+ha_2(\lambda,\delta)\right]\rho^{2\lambda+\delta}+ \tag{53}$$
$$\left[\frac{\Gamma(\lambda+\delta+1)}{\Gamma(-\lambda+\delta+1)}+h\right]a_1(\lambda,\delta)\rho^{\lambda+\delta}=0\,.$$

Eq. (53) implies directly the truthfulness of the recurrence relations (50) and (51) and thus the statement is proved.

Keeping to the Thompson and Barnett's concept [35], developed by Georgiev and Georgieva [22], the above Theorem could be generalized. In particular it is seen directly that it holds for any real L, for example for $L=\pm 0.5$.

6. COULOMB WAVE EQUATION AND ITS REGULAR SOLUTION AS SPECIAL CASES OF THE FRACTIONAL ONES

If $\lambda=1$, $g=-2\eta$, $\delta=L$, $L=0,1,2,...$ and $h=-L(L+1)$, the fractional Coulomb wave equation is reduced to the classical one, i.e. the second is a partial case of the first. A similar interdependence should exist between their solutions. Really, introducing the above values of parameters λ, g, δ and h in the expressions (49)-(51) yields:

$$u(\rho)=\sum_{p=1}^{\infty} a_p(1,L)\rho^{p+L}, \tag{54}$$

$$a_2(1,L)\left[\frac{\Gamma(L+3)}{\Gamma(L+1)}-L(L+1)\right]-2\eta a_1(1,L)=0, \tag{55}$$

$$a_p(1,L)\left\{\frac{\Gamma(p+L+1)}{\Gamma[(p+L-1]}-L(L+1)\right\}-2\eta a_{p-1}(1,L)+a_{p-2}(1,L)=0, \quad p\geq 3. \tag{56}$$

Rewriting Eq. (54) in a new form and bearing in mind the properties of gamma function results in:

$$u(\rho) = \rho^{L+1} \sum_{p=1}^{\infty} a_p(1,L) \rho^{p-1}, \tag{57}$$

$$a_2(1,L) = \frac{\eta}{L+1} a_1(1,L), \tag{58}$$

$$(p+2L)(p-1) a_p(1,L) = 2\eta a_{p-1}(1,L) - a_{p-2}(1,L), \quad p \geq 3. \tag{59}$$

Putting:

$$a_p(1,L) = C_L(\eta) A_{L+p}^L, \quad p \geq 1 \tag{60}$$

in Eqs. (57)-(59) gives:

$$u(\rho) = C_L(\eta) \rho^{L+1} \sum_{p=1}^{\infty} A_{L+p}^L \rho^{p-1}, \tag{61}$$

$$A_{L+2}^L = \frac{\eta}{L+1} A_{L+1}^L \tag{62}$$

$$(p+2L)(p-1) A_{L+p}^L = 2\eta A_{L+p-1}^L - A_{L+p-2}^L, \quad p \geq 3. \tag{63}$$

Assuming that $C_L(\eta)$ and A_{L+1}^L are represented by formulae (14) and (15), resp., and setting $p = q - L$, it is readily seen that Eqs. (61), (62) and (63) coincide with Eqs. (13), (17) and (18), resp. in the case considered it holds:

$$u(\rho) \equiv F_L(\eta, \rho). \tag{64}$$

In view of what has been said at the end of the previous Section, the last relation is true for any real L, including for $L = \pm 0.5$.

7. CONCLUSION

A definite fractional differential equation is considered and its regular partial integral is found, applying the power series method. It is demonstrated that under certain conditions the equation referred to and its solution reduce to the Coulomb wave equation and the regular Coulomb wave function $F_L(\eta, \rho)$, resp. For this reason they could be termed as *Fractional Coulomb equation* and regular *Fractional Coulomb function*, resp. The results obtained allow the development of a new approach to the analysis of problems, attacked until now by the classical Coulomb or by the confluent hypergeometric functions, e.g. the one for normal TE_{0n} modes in the circular waveguide, entirely filled with azimuthally magnetized ferrite.

ACKNOWLEDGEMENT by Y. Nikolova, L.I. Boyadjiev

This survey is partially supported by Project ID_09_0129 (ID № 02–17/2009): „Integral Transform Methods, Special Functions and Applications", National Science Fund – Ministry of Education and Science, Bulgaria.

ACKNOWLEDGEMENT by G.N. Georgiev, M.N. Georgieva-Grosse

We express our gratitude to our mother Trifonka Romanova Popnikolova and to our late father Nikola Georgiev Popnikolov for their self-denial and for their tremendous efforts to support all our undertakings.

REFERENCES

[1] K. Miller and B. Ross, *An Introduction to the Fractional Calculus and Fractional Differential Equations*. John Wiley & Sons, 1993.

[2] I. Podlubny, *Fractional Differential Equations* (*An Introduction to Fractional Derivatives, Fractional Differential Equations, to Methods of Their Solution and some of Their Applications*). San Diego: Academic Press, 1999.

[3] K. Oldham and J. Spanier, *The Fractional Calculus: Theory and Applications of Differentiation and Integration to Arbitrary Order*. Mathematics in Science and Engineering, vol. V, Academic Press, 1974.

[4] S.G. Samko, A.A. Kilbas and O.I. Marichev, *Fractional Integrals and Derivatives, Theory and Applications*. Amsterdam: Gordon and Breach, 1993.

[5] J. Crank, *The Mathematics of Diffusion*. 2nd Ed., Oxford: Clarendon Press, 1979.

[6] B. Mandelbrot, *The Fractal Geometry of Nature*. San Francisco: Freeman, 1982.

[7] L. Boyadjiev and R. Scherer, "Fractional extensions of the temperature field problem in oil strata," *Kuwait J. Sci. Eng.*, vol. 31(2), pp. 15-32, 2004.

[8] L.J. Slater, *Confluent Hypergeometric Functions.* Cambridge, UK: Cambridge Univ. Press, 1960.

[9] M. Abramowitz and I. Stegun, Eds., *Handbook of Mathematical Functions with Formulas, Graphs and Mathematical Tables.* Applied Mathematics Series 55. Washington, D.C.: National Bureau of Standards, 1964.

[10] A.R. Curtis, *Coulomb Wave Functions*. Moscow, USSR: Computational Center of the Academy of Sciences of USSR, 1969, (in Russian).

[11] Z.X. Wang and D.R. Guo, *Special Functions*. World Scientific, 1989.

[12] H. Buchholz, *Die Konfluente Hypergeometrische Funktion mit Besonderer*

Berücksichtigung Ihrer Anwendungen. Berlin, Göttingen, Heidelberg, Germany: Springer-Verlag, 1953.

[13] A. Erdélyi, W. Magnus, F. Oberhettinger and F.G. Tricomi, *Higher Transcendental Functions.* Bateman project. vol. I, New York, Toronto, London: McGraw-Hill, 1953.

[14] Ph.M. Morse and H. Feshbach, *Methods of Theoretical Physics.* Part I, New York: McGraw-Hill, 1953.

[15] F.G. Tricomi, *Funzioni Ipergeometriche Confluenti.* Rome, Italy: Edizioni Cremonese, 1954.

[16] F.G. Tricomi, *Fonctions Hypergéométriques Confluentes.* Paris, France: Gauthier-Villars, 1960.

[17] E.T. Whittaker and G.N. Watson, *A Course of Modern Analysis.* Cambridge, UK: Cambridge Univ. Press, 1965.

[18] W. Magnus, F. Oberhettinger and R.P. Soni, *Formulas and Theorems for the Special Functions of Mathematical Physics*. Berlin, Heidelberg, New York: Springer-Verlag, 1966.

[19] E.L. Ince, *Ordinary Differential Equations.* New York: Dover Publications, 1956.

[20] M.N. Georgieva-Grosse, G.N. Georgiev, L.I. Boyadjiev and Y. Nikolova, "On an application of the regular fractional Coulomb wave function," *J. Appl. Electromagn.*, 2012, (in print).

[21] K.P. Ivanov and G.N. Georgiev, "On a class of electromagnetic wave functions for propagation along the circular gyrotropic waveguide," *IEEE Trans. Microwave Theory Tech*., vol. MTT-34, no. 8, pp. 853-862, Aug. 1986.

[22] G.N. Georgiev and M.N. Georgieva, "On several new applications of Coulomb wave functions," in *Proc. 1994 Int. Conf. Comput. Electromagn. Its Appl*. *ICCEA'94*, Beijing, China, pp. 163-166, 1-4 Nov. 1994.

[23] G.N. Georgiev and M.N. Georgieva-Grosse, "Formulae for differential phase shift computation in an azimuthally magnetized circular ferrite waveguide," in *Proc. Millenn. Conf. Antennas Propagat. AP-2000,* Davos, Switzerland, vol. I-Antennas, p. 517, in Abstracts, paper 1002, in CDROM, 9-14 Apr. 2000.

[24] G.N. Georgiev and M.N. Georgieva-Grosse, "A new property of the complex Kummer function and its application to waveguide propagation," *IEEE Antennas and Wireless Propagat. Lett.*, vol. AWPL-2, pp. 306-309, Dec. 2003.

[25] G.N. Georgiev and M.N. Georgieva-Grosse, "Iterative method for differential phase shift computation in the azimuthally magnetized circular ferrite waveguide," in *Proc. 27th Progr. In Electromagn. Res. Symp. PIERS 2010*, Xi'an, China, p. 186, in Abstracts, pp. 274-278, in *PIERS Proc.*, 22-26 Mar. 2010; *PIERS Online*, vol.6, no.4, pp. 365-369, 2010.

[26] M.N. Georgieva-Grosse and G.N. Georgiev, "Advanced studies of the differential phase shift in the azimuthally magnetized circular ferrite waveguide," in *Proc. 28th Progr. In Electromagn. Res. Symp. PIERS 2010*, Cambridge, MA, USA, p. 505, in Abstracts, pp. 841-845, in *PIERS Proc.*, 5-8 July 2010.

[27] G.N. Georgiev and M.N. Georgieva-Grosse, "An application of the zeros of Laguerre polynomials," in *Proc. Twelfth Int. Conf. Electromagn. Adv. Applicat. ICEAA'10 offshore*, Sydney, Australia, pp. 637-640, 20-24 Sept. 2010.

[28] M.N. Georgieva-Grosse and G.N. Georgiev, "On the class of A_1 numbers: Definition, numerical modeling, domain of existence, basic property and application," in *Proc. XXX URSI General Assembly*, Istanbul, Turkey, article ID BP1.47, 4 pages, in CDROM, 13-20 Aug. 2011.

[29] M.N. Georgieva-Grosse and G.N. Georgiev, "Comparative study on the computational modeling and application of the classes of A, B, C numbers," in *Proc. Thirteenth Int. Conf. Electromagn. Adv. Applicat. ICEAA'11*, Turin, Italy, pp. 660-663, 12-16 Sept. 2011, (Invited Paper in the Special Session "Future challenges in mathematical and computational electromagnetics and its applications," organized by G.N. Georgiev and M.N. Georgieva-Grosse).

[30] M.I. Andriychuk, N.N. Voitovich, P.A. Savenko and V.P. Tkachuk. *Synthesis of Antennas according to Amplitude Directivity Pattern: Numerical Methods and Algorithms*. Kiev: Naukova Dumka Pub., p. 256, 1993, (In Russian).

[31] R.C. Hansen, *Phased Array Antennas*. N.Y.: Wiley, 1998.

[32] R.J. Mailloux, *Phased Array Antenna Handbook*. Norwood, MA: Artech House, 1st Ed. 1994, 2nd Ed. 2005.

[33] V.A. Kashin and A.P. Safonov, "A monopulse transmit-receive phased array with polarization discrimination of targets in the main beam," *J. Commun. Technol. Electron.*, vol. 50, no. 8, pp. 853-862, Aug. 2005, (in Russian).

[34] S.P. Skobelev, *Phased Array Antennas with Optimized Element Patterns*. Norwood, MA: Artech House, 2011.

[35] I.J. Thompson and A.R. Barnett, "Coulomb and Bessel functions of complex arguments and order," *J. Comput. Phys.*, vol. 64, pp. 490-509, 1986.

ELECTRODYNAMIC ANALYSIS OF THE SQUARE LOOP ANTENNA AND ITS ELECTROMAGNETIC COMPATIBILITY

K.V. Kotetishvili*, G.Sh. Kevanishvili**, I.G. Kevanishvili **,

A.V. Asanidze, G.G. Chikhladze, and L.G. Kokilashvili

Georgian Technical University, 0175, 77, M. Kostava st., Tbilisi, Georgia,

K.V. Kotetishvili*, Professor, GTU, Faculty of Informatics & Control Systems, 0177, 20, Chikovani st., Tbilisi, Georgia, **ketinooo@hotmail.com,**

G.Sh. Kevanishvili**, Professor, GTU, Faculty of Power Engineering & Telecommunications, 0186, 83, Vaja Pshavela av., ap. 127, Tbilisi, Georgia,

I.G. Kevanishvili **, Professor, GTU, Faculty of Power Engineering & Telecommunications, 0186, 83, Vaja Pshavela av., ap. 127, Tbilisi, Georgia,

Abstract

The electrodynamic theory of the square loop antenna is suggested. In an analytical form its radiation pattern is presented. The problem of electromagnetic compatibility of this antenna is studied. The numerical data are given, and the way of selection of the optimal regime of the antenna functioning is stated.

1. INTRODUCTION

The square loop antennas formed of thin conductors are widely used in antenna techniques, while their correct electrodynamic theory is still absent in the scientific literature. As a result, it is impossible to carry out the engineer calculations of their basic electrodynamic characteristics (radiation patterns, ordered action coefficients and others) as well, as the theoretical investigations of their electromagnetic compatibility.

2. THE THEORY OF THE SQUARE LOOP ANTENNA

2. 1. Description of the antenna and setting of the problem

In figure 1 the square loop antenna presents itself in the rectangular coordinate system (XYZ). Here AB and DC are the horizontal conductors of the loop, while AD and BC – the vertical ones. DC conductor is cut in the center and in K and L points of the slit the harmonic e. m. force of the constant amplitude is applied, creating the alternating current of constant amplitude (I_0) in the loop. The action of this current tends to appearance of electromagnetic waves, penetrating in outer space.

The aim of the paper is determination of the structure of the electromagnetic field in arbitrary point of observation M, arranged in the far zone of the antenna, i.e. reception of the quantitative relations, making possible to calculate the components of electric and magnetic vectors $\vec{E}, \vec{H}$.

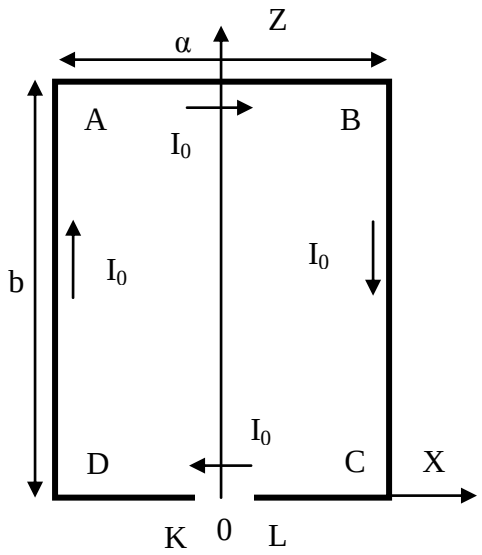


Figure 1. Orientation of the loop in XYZ coordinate system

2.2. The structure of the electromagnetic field of the antenna

The problem being set up may be solved step by step. First of all let us calculate the value (E_z) of the vertical component of the electric field created by the vertical conductors AD and BC in the point of observation M. Taking into account that, in AD and BC conductors the currents are directed oppositely, for the quantitative expression of E_z component yields (figure 2):

$$E_z = P\left(\int_0^b \frac{e^{-ikr_1}}{r_1}dz' - \int_0^b \frac{e^{-ikr_2}}{r_2}dz'\right), \qquad (1)$$

where $P = I_0 \frac{k^2 r_0}{4\pi\omega\varepsilon_0}$ V/m, $k = 2\pi/\lambda$, λ being the wavelength in vacuum, r_0 – the radius of the conductor, ω – the circular frequency of the field, $\varepsilon_0 = \frac{1}{36\pi}\cdot 10^{-9}$ F/m.

In the far zone

$$r_1 \cong R + a\cos\varphi\sin\theta - z'\cos\theta, \quad r_2 \cong R - a\cos\varphi\sin\theta - z'\cos\theta,$$

(R – the distance from the origin to M point)

Inserting them into (1) and carrying out the integration procedure, we get:

$$E_z = E_0(R)b\sin(\gamma\cos\varphi\sin\theta)\frac{\sin(\beta\cos\theta)}{\beta\cos\theta},\ \left(\gamma=\pi\frac{a}{\lambda},\ \beta=\pi\frac{b}{\lambda}\right) \tag{2}$$

where $E_0(R)=P\dfrac{e^{-ikR}}{R}$ is the amplitude of the electric field.

Further it becomes necessary utilization of the meridian component E_θ of the electric field strength (figure 2), which is the projection of $\vec{E}_z$ vector on the KK′ tangent to the meridian passing through M point; thus, $E_\theta=-E_z\sin\theta$, θ being the meridian angle of M point. Taking (2) into account, we get:

$$E_\theta = -E_0(R)b\sin(\gamma\cos\varphi\sin\theta)\frac{\sin(\beta\cos\theta)}{\beta\cos\theta}\sin\theta\,. \tag{3}$$

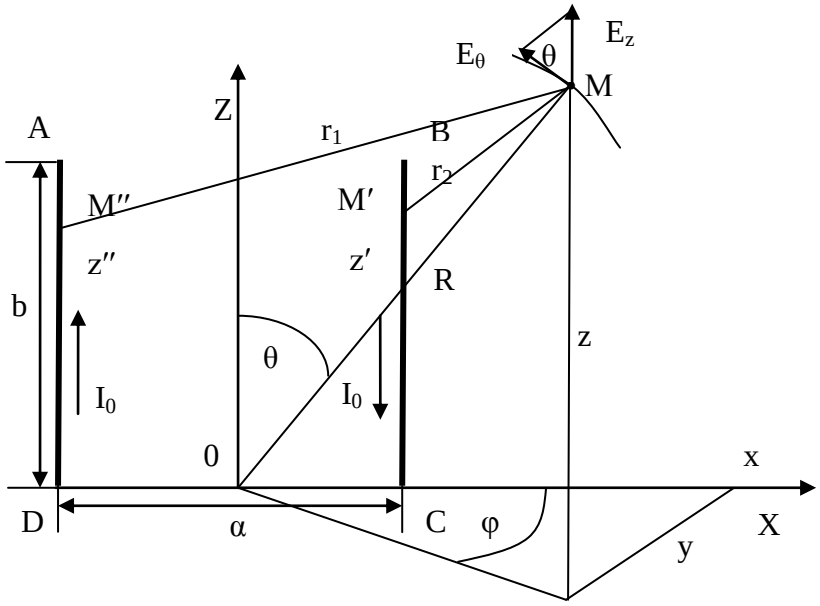


Figure 2. For the solution of (2) formula.

The azimuth component (H_φ) of the magnetic vector may now be calculated from the first or the second Maxwell's equations in the spherical coordinate system.

Now let us calculate the value (E_x) of the azimuth component of the electric field created by the horizontal conductors AB and DC of the loop in the same point of observation M. For this purpose let us turn to the figure 3 and take into account that, in the far zone the horizontal component of the electric field strength should be given as follows:

$$E_x = I_0 \frac{k^2 r_0}{4\pi i \omega \varepsilon_0} \int_{-a/2}^{a/2} \left(\frac{e^{-ikr_1}}{r_1} - \frac{e^{-ikr_2}}{r_2} \right) dx', \tag{4}$$

while in the far zone

$$r_1 \cong R + x' \cos\varphi \sin\theta - b\cos\theta, \quad r_2 \simeq R - x' \cos\varphi \sin\theta. \tag{5}$$

Inserting (5) into (4) and calculating the received integral yields:

$$E_x = E_0(R) a e^{i\beta\cos\theta} \sin(\beta\cos\theta) \frac{\sin(\gamma \sin\theta\cos\varphi)}{\gamma \sin\theta\cos\varphi}. \tag{6}$$

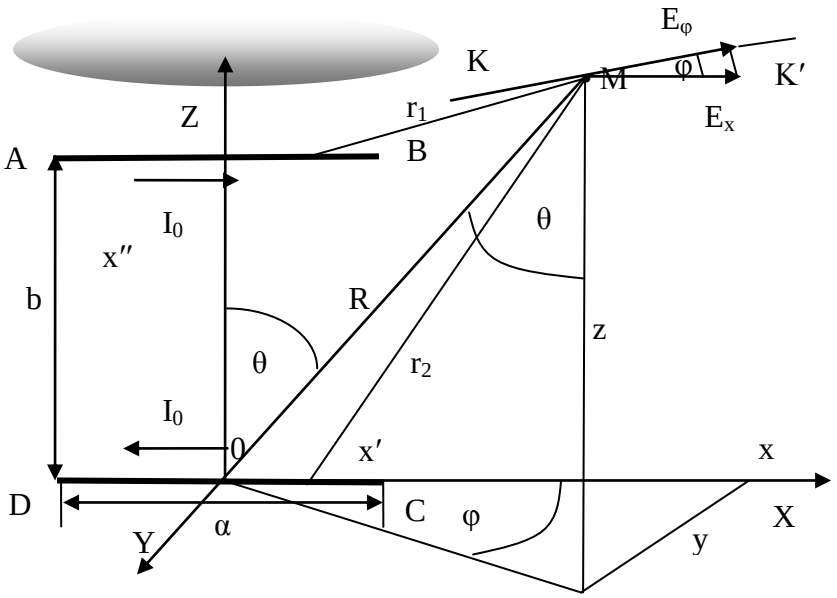


Figure 3. For the solution of (7) formula.
The darkening presents the plane, parallel to the equatorial plane.

The azimuth component of the field E_φ may be calculated from the expression $E_\varphi = -E_x \sin\varphi$, or

$$E_\varphi = -E_0(R) a e^{i\beta\cos\theta} \sin(\beta\cos\theta) \frac{\sin(\gamma \sin\theta\cos\varphi)}{\gamma \sin\theta\cos\varphi} \sin\varphi. \tag{7}$$

The radiation pattern of the loop may be calculated now from the following formula:

$$F(\theta,\varphi) = \sqrt{q^2 \left|F_{\parallel}(\theta,\varphi)\right|^2 + \left|F_{\perp}(\theta,\varphi)\right|^2}, \quad \left(q = b/a\right) \tag{8}$$

$$\left|F_{\parallel}(\theta,\varphi)\right| = \left| \sin(\gamma \sin\theta\cos\varphi) \frac{\sin(\beta\cos\theta)}{\beta\cos\theta} \sin\theta \right|, \tag{9}$$

$$\left|F_{\perp}(\theta,\varphi)\right| = \left|\sin(\beta\cos\theta)\frac{\sin(\gamma\sin\theta\cos\varphi)}{\gamma\sin\theta\cos\varphi}\sin\varphi\right|. \qquad (10)$$

Taking (9) and (10) into account, (8) presents the distribution of the electromagnetic field, radiated in outer space by the loop according to θ and φ angles.

The analysis of (8) formula is given below, that will make possible to determine the basic electromagnetic properties of the square loop antenna.

2.3. Special cases

1) Consider the situation of presence of the point of observation in the main meridian plane $(\varphi = 0)$, then from (8) follows:

$$F(\theta,0) = q\left|F_{||}(\theta,0)\right| = q\sin(\gamma\sin\theta)\frac{\sin(\beta\cos\theta)}{\beta\cos\theta}\sin\theta, \quad F_{\perp}(\theta,0) = 0. \qquad (11)$$

It is seen here that, when $\theta = 0$, then $F_{||}(0,0) = 0$ and, thus, in the vertical direction the antenna does not radiate, while, if $\theta = \pi/2$, corresponding to the horizontal direction, then we get:

$$F_{||}\left(\pi/2,0\right) = \sin\gamma = \sin\left(\pi a/\lambda\right). \qquad (12)$$

And now everything depends on the relation between the width a of the loop and the wavelength λ. If $a = n\lambda$ $(n = 1,2,3,\ldots)$, then $F_{||}\left(\pi/2,0\right) = 0$, i.e. when the width of the loop equals to the integral number of wavelengths, then it does not radiate in horizontal direction, while, if $a = (2n+1)\lambda/2$, i.e. the width equals to the odd number of wavelengths – the radiation in this direction is maximal $\left|F_{||}\left(\pi/2,0\right)\right| = 1$.

2) Assume now the point of observation be in the main equatorial plane, i.e. $\theta = \pi/2$, then

$$\left|F_{||}\left(\pi/2,\varphi\right)\right| = \sin(\gamma\cos\varphi), \quad F_{\perp}\left(\pi/2,\varphi\right) = 0 \quad \left(\gamma = \pi\frac{a}{\lambda}\right).$$

It is clear that, in directions, where $\varphi_n = \arccos\left(n\lambda/a\right)$, $(n = 1,2,3,\ldots)$, $n\lambda/a < 1$, the loop does not radiate, while in directions, where $\varphi_m = \arccos\left[(2m+1)\lambda/2a\right]$, $(m = 0,1,2,\ldots)$, the radiation is maximal.

3) When $\varphi = \pi/2$, i.e. the point of observation is located in the plane, normal to the plane of the loop itself, then $F_{]]}\left(\theta, \pi/2\right) = 0$ and $F_{\perp}\left(\theta, \pi/2\right) = \sin(\beta\cos\theta)$. Then it follows from it that, at $\theta_{||} = \arccos\left(n\lambda/b\right)$ $(n = 1,2,3,\ldots)$ angles the radiation is absent, while at $\theta_{||} = \arccos\left[(2m+1)\lambda/2b\right]$ – it is maximal.

4) If the parameter $q << 1$ $\left(b/a << 1\right)$ the square loop antenna transforms into the hinge (figure 4).

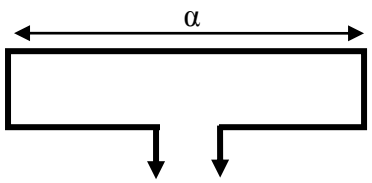


Figure 4. The hinge antenna.

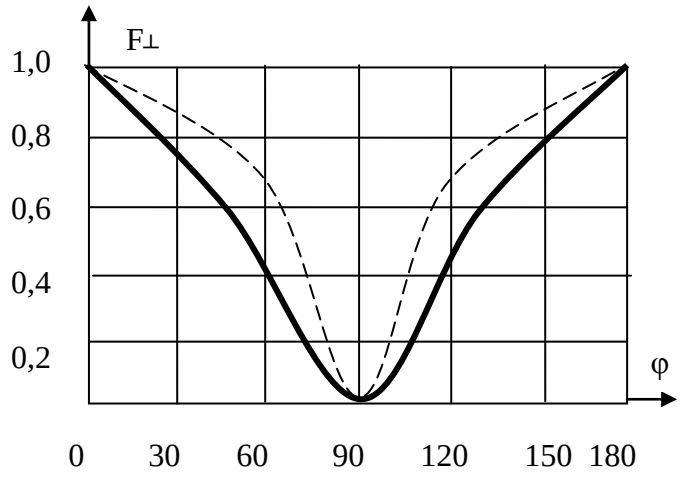


Figure 5.

In the main equatorial plane $\left(\theta = \pi/2\right)$ we have $F_{\perp}\left(\pi/2, \varphi\right) = 0$, while

$$F_{||}\left(\pi/2, \varphi\right) = q\sin(\gamma\cos\varphi),$$

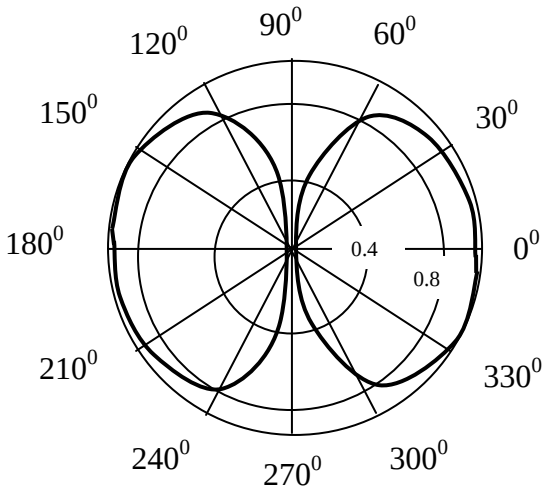


Figure 6.

and for the normalized characteristic we get

$$F_{Nor}(\varphi) = F_{||}\left(\pi/2\right)/q = \sin(\gamma\cos\varphi),$$

$$\left(\gamma = \pi\frac{a}{\lambda}\right). \qquad (13)$$

In the case, when the width of the loop a equals to half of λ wavelength, (13) transforms into

$$F_{Nor}(\varphi) = \sin\left(\frac{\pi}{2}\cos\varphi\right). \qquad (14)$$

In figure 5 the radiation pattern of the hinge is built up (firm line), at

$a = \lambda/2$ (taken from [1]), but it is unknown there, whether it is theoretical or experimental curve. As to the dotted line, it is constructed due to our theoretical expression (14). As it is seen there, these two curves, within the graphical error, are in good coincidence.

In figure 6 the same pattern, according to (14) formula, is presented in the polar coordinate system.

2.4. Electromagnetic compatibility (EMC) of the antenna

In this sectiont we consider the problem of estimation of the electromagnetic compatibility (EMC) of the square loop antenna, applying for this purpose so called compatibility function [2], given as follows:

$$C(\gamma) = \left| \int_{\Omega} F(\theta, \varphi, \gamma) d\Omega \right|, \tag{15}$$

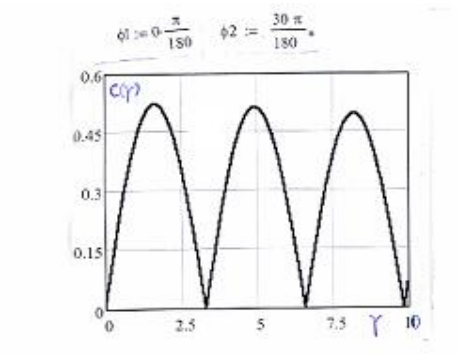
$F(\theta, \varphi, \gamma)$ being the radiation pattern of the antenna, $\Omega -$ the area of the side radiation, while $\gamma -$ the unit less, arbitrary, geometric characteristic of the antenna relative to the wavelength. It is known that, the side lobs of the radiation pattern resist the normal functioning of neighbor antennas.

The physical meaning of $C(\gamma)$ function consists in following: it is the integral modulus of the amplitude of the side radiation. Obviously, if γ_n $(n = 1,2,3,\ldots)$ are the roots of $C(\gamma)$ function, then in the radiation pattern of the antenna the side lobs should not exist at all, or their amplitude level should be as low, as possible

In the main equatorial plane $\left(\theta = \pi/2\right)$ the expression (15) transforms into

$$C(\gamma) = C_{\perp}(\gamma) = \left| \int_{\varphi_1}^{\varphi_2} F\left(\pi/2, \varphi, \gamma\right) d\varphi \right|, \tag{16}$$

while in the main vertical (meridian) plane $(\varphi = 0)$ it becomes



$$C(\gamma) = C_{\parallel}(\gamma) = \left| \int_{\theta_1}^{\theta_2} F(\theta, 0, \gamma) d\theta \right|. \tag{17}$$

In these expressions (φ_1, φ_2) and (θ_1, θ_2) intervals present the compatibility areas, where the side lobs are arranged.

Hence, at investigation of the electromagnetic compatibility of any antenna the decisive importance is awarded to the statement of the structure of compatibility functions as well, as to the determination of the values of their roots.

In our particular case $C_{\perp}(\gamma)$ function is determined from the following relation:

re 7a.

$$C_{\perp}(\gamma)=\left|\int_{\varphi_1}^{\varphi_2}\sin(\gamma\cos\varphi)d\varphi\right|. \qquad (18)$$

One may easily convince in that, in given case the side lobs appear in the radiation pattern only when the parameter γ exceeds π $(\gamma>\pi)$.

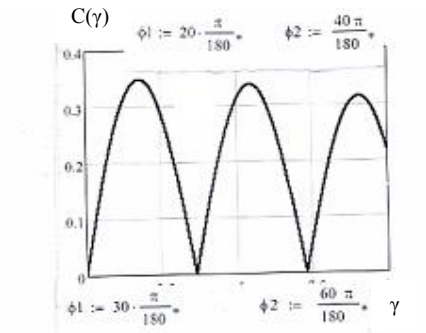
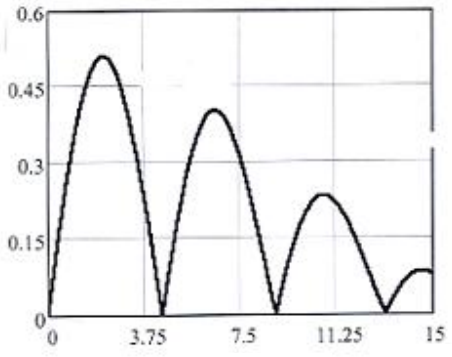


Figure 7b.

In figure 7 (a, b, c, d) the graphs of the compatibility function are presented in the upper semi-plane, within the following compatibility areas:

$$\Delta_a=(\varphi_2-\varphi_1)=\left(30^0-0^0\right),\ \Delta_b=\left(40^0-20^0\right),$$
$$\Delta_c=\left(60^0-30^0\right),\ \Delta_d=\left(90^0-60^0\right).$$

C(γ)



C(γ) Figure 7c.

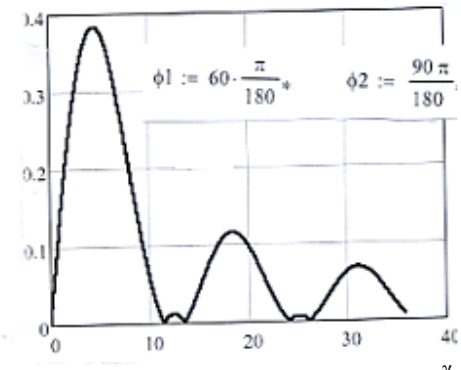


Figure 7d.

First three roots of the compatibility function presented in figure 7a are $\gamma_1=2.29$, $\gamma_2=6.57$, $\gamma_3=9.86$, . In the first case, when $\gamma_1<\pi$, the pattern has no side lobs. The corresponding pattern is presented in figure 8 (the pattern is built only in the upper semi-plane). The patterns, corresponding to $\gamma_2=6.57$ and $\gamma_3=9.86$ roots (figure 9(a,b)) within the compatibility area $\Delta_a=\left(30^0-0^0\right)$ (darkening sections), include the side lobs of very small energetic level, so that within the signed sector the EMC of the antenna is rather high.

First two roots of the compatibility function presented in figure 7b are $\gamma_1=3.64$, $\gamma_2=7.29$. In this case the compatibility area is

$\Delta_b = \left(40^0 - 20^0\right)$ (the compatibility areas in figure 10(a,b) are darkening as well), but now, different from the previous case, the patterns in these sectors have no side lobs at all, and the EMC of the antenna is 100%.

In the case of figure 7c, the roots of the compatibility function equal $\gamma_1 = 4.48$, $\gamma_2 = 8.94$, while the compatibility area is presented by $\Delta_c = \left(60^0 - 30^0\right)$, in the first case the compatibility sector (darkening) occupies the insignificant part of the sector (figure 11a), that does not cause the decrement in EMC, while for the pattern with $\gamma_2 = 8.94$ its wide area is covered by the significant level of side radiation – the EMC of the antenna will be rather low (figure 11b).

Figure 8.

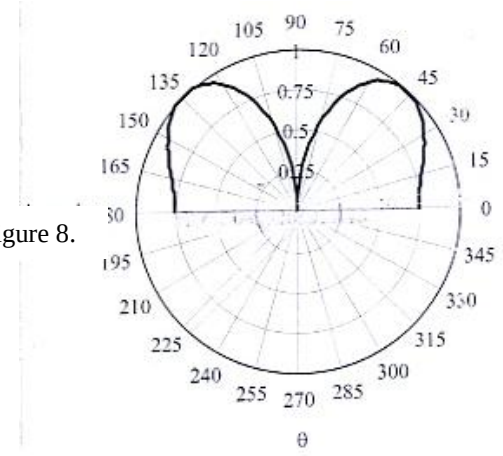
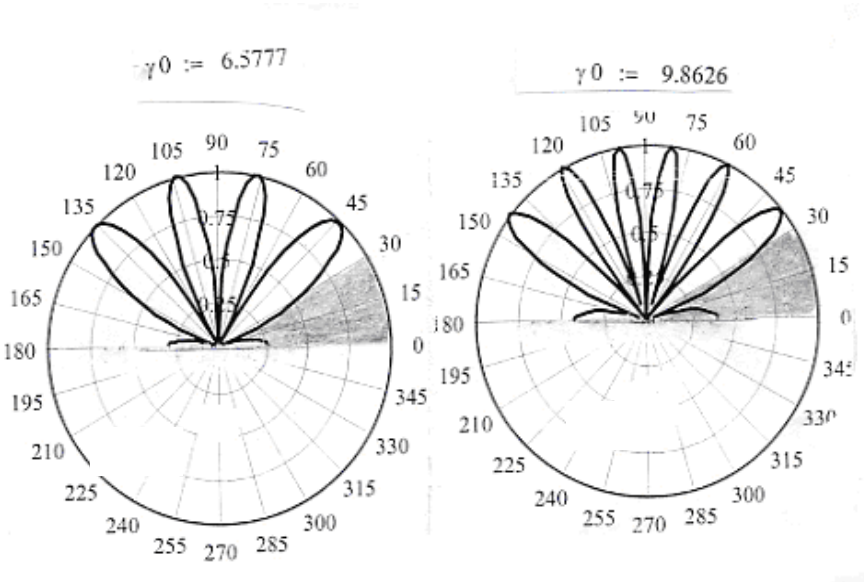


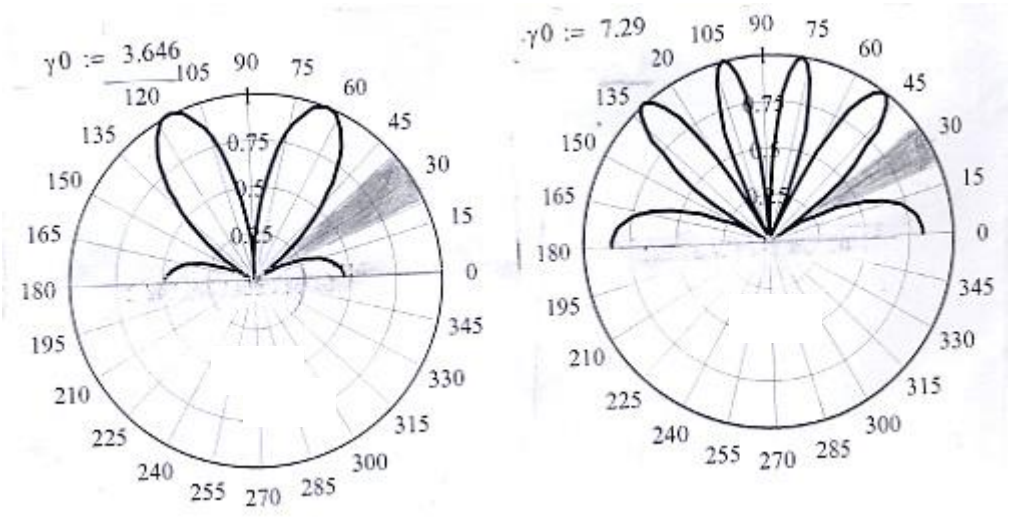
Figure 8.



(a) (b)

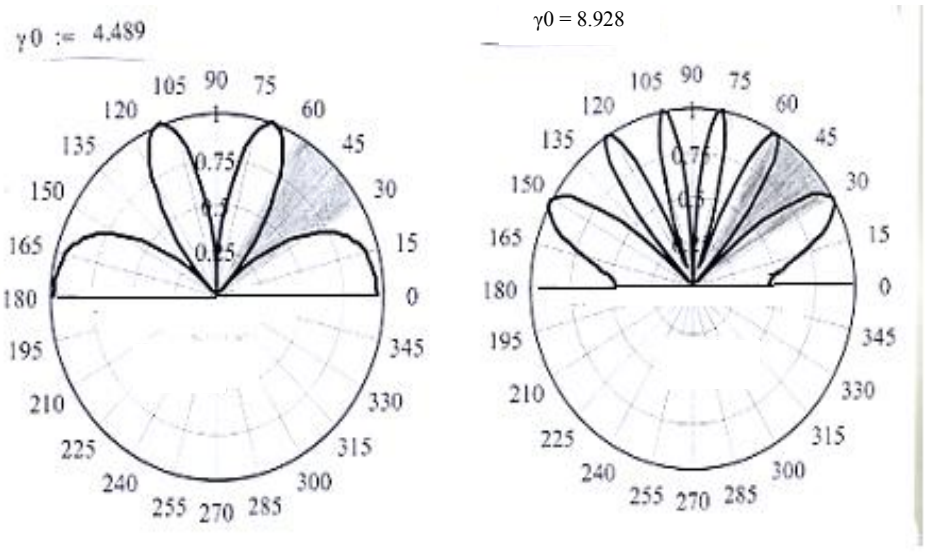
Figure 9.

Investigations show that, at high γ_n $(n=1,2,3,\ldots)$, in radiation patterns increases the number of side lobs, arranged very tight to each other. Due to it, the attempts of the choosing the compatibility areas become useless. To such situation corresponds, for example, the pattern given in figure 12 ($\Delta_d = \left(90^0 - 60^0\right)$, where two side lobs of high level are located.



(a) (b)

Figure 10.



(a) (b)

Figure 11.

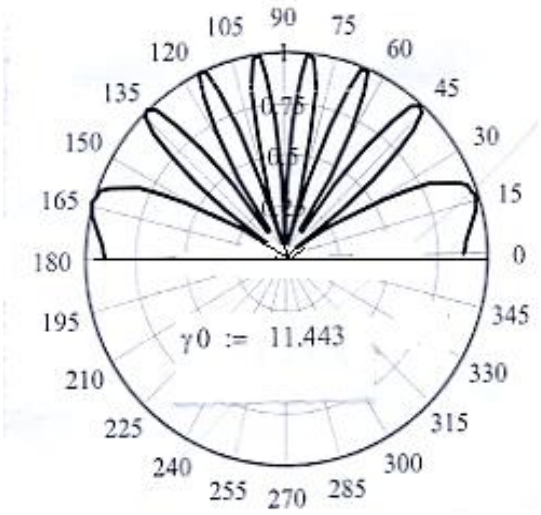


Figure 12.

3. CONCLUSION

In presented paper the electrodynamic theory of the square loop antenna formed of thin conductors is developed. The method for the estimation of the EMC of the square loop antenna is suggested. This method provides the optimal (needed) regime of functioning of the square loop antenna.

The present project has been fulfilled by financial support of Shota Rustaveli National Science Foundation (Grant №1, 1/51). Any idea in this publication is possessed by the authors and may not present the opinion of Shota Rustaveli National Science Foundation itself.

REFERENCES

[1]. K.Rothammel. Antennas, Moscow, "Energia", 1979, p. 16.

[2]. G.Sh. Kevanishvili, Z.I. Sikmashvili, I.G. Kevanishvili, K.V. Kotetishvili, A.V. Asanidze, and G.G. Chikhladze. About the Electromagnetic Compatibility of Dipole Antenna, IEEE, DIPED-08, Proceedings of XIIIth International Seminar Work/Shop, Tbilisi, 22-25, 9, 2008.

MANEV'S PROBLEM, 87 YEARS LATER

(Selected from International Conference on 'Electronics, Mathematics, Physics and Applied Electromagnetism', held at the National Military University 'Vasil Levski', V. Turnovo, Bulgaria, 25-3-2011)

Assen Kyuldjiev

Institute for Nuclear Research and Nuclear Energy,
Bulgarian Academy of Sciences; 72 Tzarigradsko chaussée, 1784 Sofia, Bulgaria
e-mail: kyuljiev@inrne.bas.bg

IN TRIBUTE TO PROF. GEORGI MANEV (1884–1965)

Abstract

We demonstrate here the existence of a local Darboux chart for the Manev model such that its dynamics becomes locally equivalent to Kepler model. This explains lot of similarities between these two models and especially why they share common invariants and symmetry algebras. We also discuss the role of a certain discrete symmetry appearing in the case of a bounded motion.

Keywords: Hamiltonian dynamics, superintegrability, symmetry algebras, Kepler problem

1. INTRODUCTION

Since time immemorial the circular motion was the archetype motion of the heavenly bodies, and circle was assumed to be embodiment of perfection. Since Kepler and Newton elliptical trajectories became the new archetype of the (bounded) planetary motion and the circular orbit is nowadays viewed upon rather as a degenerate ellipse than as an embodiment of perfection. The advent of Einstein's theory did not produce a new archetype of heavenly motions, apart from the exceptional case of a collapse into the (still hypothetical) black holes. Nevertheless, among the variety of relativistic effects the perihelion shift of inner planets is definitely the best recognizable effect in the Solar system. Maybe it is time to accept a new archetype of heavenly motions: *precessing ellipse* (or more generally, precessing conics). If precessing conics give us 'the typical' motion of planets it is tempting to ask which central force field produces them. Surprisingly or not, the answer [2] is: *the Manev model*.

Prof. Georgi Manev was born on 27 January 1884 in *Veliko Turnovo*. After graduating form Sofia University he worked as a teacher in Physics and Mathematics (mostly in his hometown, but also in Razgrad). He returned from the fronts of the two Balkan Wars

and the First World War decorated with a medal for courage and promoted to officer. After one year specialisation in Toulouse he started reading a 4-semester course on Theoretical Physics in Sofia University. Later on he became the first Holder of the Chair on Theoretical Physics and authored three University textbooks on Theoretical Physics. He served two terms as a Dean of the Physics Faculty, became Rector of the Sofia University and a Minister of Education in 1938. He is best known for his series of scientific papers on what is known now as Manev's model which attracted renewed attention in the last two decades and made him the best cited Holder of the Chair of Theoretical Physics in Sofia University.

In the last two decades Manev model had enjoyed an increased interest either as a very suitable approximation to Einstein's relativistic dynamics from astronomers' point of view or as a toy model for applying different techniques of the modern dynamics (see e.g. [4, 6, 8, 20, 21]). Prof. Manev was not the first one to study this model but he was the first to deduce it from 'first principles', namely as a *consequence* of Max Planck's (more general) action-reaction principle and for this reason his name is attached to the model. It is remarkable that such a simple model is capable to describe both the perihelion advance of the inner planets and the Moon's perigee motion. It was also argued in [8] that it is the natural classical analog of the Schwarzschild problem in a certain approximative regime.

By Manev model [19] we mean here the dynamics given by the Hamiltonian:

$$H = \frac{1}{2}(p_x^2 + p_y^2 + p_z^2) - \frac{A}{r} - \frac{B}{r^2} \tag{1}$$

where $r = \sqrt{x^2 + y^2 + z^2}$; A and B are assumed to be arbitrary real constants whose positive values correspond to attractive forces. Due to rotational invariance each component of the angular momentum is an obvious first integral and so, like the Kepler problem (and any central potential), the Manev model is *integrable*.

The motion is confined on a plane which we assume to be xOy and correspondingly the angular momentum $L \equiv L_z = yp_x - xp_y$ is in the z-direction. From now on we shall concentrate on this dynamics on the phase space $\mathcal{M} = T^*(\mathbb{R}^2 \backslash \{0\}) = T^*\mathbb{R}^+ \times T^*\mathbb{S}^1$ which is separable in radial coordinates r and $\theta = \arctan(y/x)$ as it is governed by:

$$\begin{aligned} H &= \frac{1}{2}\left(p_r^2 + \frac{L^2 - 2B}{r^2}\right) - \frac{A}{r} \\ \omega &= \mathrm{d}p_r \wedge \mathrm{d}r + \mathrm{d}L \wedge \mathrm{d}\theta. \end{aligned} \tag{2}$$

Qualitatively, its 'radial dynamics' behave like radial motion of Kepler dynamics with angular momentum squared $\ell^2 - 2B$; while the case $2B > \ell^2$ corresponds to overall centripetal effect. On the other hand, the angular equation of motion $\dot{\theta} = \ell/r^2$ is still governed by the 'authentic' angular momentum ℓ (and r is as just described). Consequently, the remarkable properties of Kepler dynamics that all negative energy orbits are closed and the frequencies of radial and angular motions coincide (for any initial conditions) are no more true. Thus we may have not only purely classical perihelion shifts but also if $2B \geq \ell^2 \neq 0$ we may have collapsing trajectories which are spirals; while in the Kepler dynamics the only allowed fall down is along straight lines. (Spiraling here has nothing to do with non-conservative forces but follows from the fact that in the Manev model collapse is possible for non-vanishing angular momentum as well.) For this reason the set of initial data leading to collision has a positive measure and this may offer an explanation why collisions in the Solar system are estimated to happen more often than Newton's theory predicts [7].

Easily one may get the suspicion that Manev problem actually offers a larger natural family of models with common properties among which Kepler model is a kind of degenerate case (just like the circle is a degenerate case among the conics). There are several types of arguments supporting this view.

• First, as we already said, it is a sensible generalization of Newton's gravitation law.

• Second, there are *stability* arguments as Kepler-type motion is generally not preserved by small perturbations and any sort of 'real world' interactions like Solar pressure, drag etc would destroy 'fixed ellipse' motion; while in the Manev model we have persistent KAM tori and cylinders for a large class of even non Hamiltonian perturbations [14].

• Third, Kepler problem is famous as one of archetypes of *superintegrable* systems – kind of property which is also assumed to be easily destroyed by small perturbations. Superintegrability is a kind of *global* property of a dynamical system which is not about the individual trajectories, but rather about how they are stacked together. It is defined by the presence of s smooth and functionally independent constants of motion where s is *greater* than the number of degrees of freedom n. Recently we reported [12] that Manev model also has an additional independent globally defined constant of motion, albeit not for all initial data.

• Fourth, even if one may expect that superintegrability could be connected with some hidden symmetry algebra, the actual finding of such a connection is not a trivial task.

(For example, we have the list of natural mechanical superintegrable models with integrals quadratic in momenta in [10] but still very little is known about their symmetry algebras, see e.g. [11].) In our case we were able to find that the symmetry algebras of the Kepler's model are present in the Manev's model as well.

• Fifth, Kepler and Manev problems share a separate common $\mathfrak{so}(2,1)$ algebra associated with the *radial motion only* and among the possible realizations of this algebra the Manev model forms a class of its own [3, 5].

We shall not dwell more on the equations of motion of the Manev model and the properties of their solutions as this is well known and more or less trivial stuff. In what follows we shall be rather concerned with the global and invariance features of the model.

2. REMINDER ON THE KEPLER PROBLEM INVARIANTS AND SYMMETRY ALGEBRAS

In the case of Kepler problem, the Hamiltonian is

$$H_K = \frac{1}{2}(p_x^2 + p_y^2 + p_z^2) - \frac{A}{r} \tag{3}$$

and we have more first integrals due to:

$$\{H_K, \vec{J}\} = 0 \quad with \quad \vec{J} = \vec{p} \wedge \vec{L} - A\vec{r}/r \tag{4}$$

being the Laplace-Runge-Lenz vector whose components are not independent as

$$J^2 = 2H_K L^2 + A^2.$$

Together with the Hamiltonian and angular momentum they close on an algebra with respect to the Poisson brackets and after redefining

$$\vec{E} = \vec{J}/\sqrt{|-2h_K|}$$

on each $H_K = h_K$ level set we get:

$$\{L, E_x\} = E_y\,, \qquad \{L, E_y\} = -E_x\,, \qquad \{E_x, E_y\} = -\mathrm{sign}(h_K)\, L \tag{5}$$

with Casimir invariant

$$E_x^2 + E_y^2 + \mathrm{sign}(-H_K)\, L^2 = \frac{A^2}{|2H_K|} \tag{6}$$

which makes obvious the fact that we have an $\mathfrak{so}(3)$ algebra for negative energies and $\mathfrak{so}(2,1)$ for positive ones. (In the case of the 3-dimensional Kepler problem the components of the angular momentum give us another copy of $\mathfrak{so}(3)$, so the combined symmetry algebra is $\mathfrak{so}(4)$ or $\mathfrak{so}(3,1)$ depending on the sign of h_K.) Actually, the first use of these first integrals (*even before they were combined into a vector*) was made by J. Hermann (= J. Ermanno) [9] in 1710 in the disguise of 'Ermanno-Bernoulli' constants:

$$J_{\pm} = J_x \pm \mathrm{i}J_y = \left(\frac{L^2}{r} - A \mp \mathrm{i}\, Lp_r\right)\mathrm{e}^{\pm \mathrm{i}\theta} \tag{7}$$

satisfying:

$$\{H_K, J_{\pm}\} = 0, \qquad \{L, J_{\pm}\} = \pm \mathrm{i}J_{\pm}, \qquad \{J_+, J_-\} = -4\mathrm{i}H_K L. \tag{8}$$

3. THE MANEV PROBLEM INVARIANTS AND SYMMETRY ALGEBRAS

Despite the fact that Manev's model has been known for so long, its additional invariants has not been found till very recently. It has been reported by us [12] (and independently, later on in [17]) the invariance of the expression:

$$\mathcal{J}_{\pm} = \nu L\left[p_r \pm \mathrm{i}\left(\nu p_{\perp} - \frac{A}{\nu L}\right)\right]\mathrm{e}^{\pm \mathrm{i}\nu\theta}\,, \quad \frac{d}{dt}\mathcal{J}_{\pm} = 0 \tag{9}$$

where

$$p_{\perp} = \frac{L}{r} \quad and \quad \nu^2 = \frac{L^2 - 2B}{L^2} \tag{10}$$

and $\mathcal{J}_+$ and $\mathcal{J}_-$ are not independent as

$$\mathcal{J}_+\mathcal{J}_- = 2\nu^2 HL^2 + A^2\,. \tag{11}$$

As $\mathcal{J}_+$ and $\mathcal{J}_-$ are not independent it would be better to distinguish $|\mathcal{J}_{\pm}|$ from $Arg(\mathcal{J}_+) = -Arg(\mathcal{J}_-)$ where the former depends only on H, L, A and the latter presenting a 'truly new' invariant expression. The existence of such an invariant which is specific for the Manev's model could be of help for some future analyses of astronomical data aiming at finding how well the motion of a heavenly body is described by the Manev's dynamical law. It may be computationally advantageous not to try to find the value of B for which we find the best fit between the observed motion and the one predicted by the Manev's

equations of motion but to try to find the value of ν for which

$$\kappa = Arg\left(\left[p_r + \mathrm{i}\left(\nu p_\perp - \frac{A}{\nu L}\right)\right]\mathrm{e}^{\mathrm{i}\nu\theta}\right) \tag{12}$$

would be (closest to a) constant along the trajectory.

Obviously in the Kepler case ν equals *one* and (up to a multiplication by i) we recover the 'Ermanno-Bernoulli' constants. The consequences of existence of the invariant expression (9) are to be analysed separately for the cases of compact and noncompact motion.

4. BOUNDED MOTION CASE

Trajectories always lie on the joint level sets of H and L which in the case when $0 \neq L^2 > 2B$ and $H < 0$ are 2-dimensional tori. The real valued ν is equal to the ratio of frequencies of the radial and angular oscillations. Trajectories fill densely these tori when ν is irrational, and hence there are no new (continuous functions on the phase space being) constants of motion in this situation.

In order to have closed trajectories and *globally* defined constants of motion we have to require that ν should be rational i.e.

$$\nu = \sqrt{L^2 - 2B} : L = m : k \tag{13}$$

with m and k coprime integers. Thus we have conditional constants of motion corresponding to (9) which exist only for disjoint but infinite set of values L, otherwise we would have invariant submanifolds of co-dimension two but not genuine constants of motion.

Let's remark that for any generic central potential we could have disjoint set of initial data corresponding to closed orbits but in our case *all* points on certain level sets of the angular momentum lie on closed orbits which are intersections with the level sets of the additional invariant. (To visualize the intersection of the different level sets we can fix the angular momentum L and use x, y and p_r as coordinates on this 3-dimensional level set. As both Hamiltonian and $Arg(\mathcal{J}_{\pm})$ level sets are represented as two-dimensional surfaces, their intersection gives the trajectories in the phase space.)

When ν is rational each trajectory in the *configuration space* $\{r, \theta\}$ looks like a

'rosette' with m petals and is invariant under rotations by angle $\frac{2\pi}{m}$. Trajectories in the *phase space* are invariant under the action of *another* cyclic group $\mathcal{C}_{km}$ generated by rotations by angle $\frac{2\pi k}{m}$:

$$\theta \to \theta + 2\pi\frac{k}{m}n \qquad n = 0,1,\ldots,m-1\,. \tag{14}$$

(while leaving the other variables intact: $r \to r$, $p_r \to p_r$, $L \to L$) and this is connected to the invarance of $\mathcal{J}_{\pm}$ (as well as H, L and ω) under the action of $\mathcal{C}_{km}$.

We shall now introduce another Darboux chart for our symplectic form (2) in the case when $L^2 - 2B > 0$ by defining new local coordinates $\vartheta = \nu(L)\theta$, and $\mathcal{L} = \nu(L)L$ which are canonically conjugate as

$$\mathrm{d}\mathcal{L}\wedge\mathrm{d}\vartheta = \mathrm{d}L\wedge\mathrm{d}\theta \qquad and\ hence \qquad \omega = \mathrm{d}p_r\wedge\mathrm{d}r + \mathrm{d}\mathcal{L}\wedge\mathrm{d}\vartheta. \tag{15}$$

(As the mapping is not one-to-one we do not have symplectomorphism.) When written in the new coordinates Manev's Hamiltonian takes the form of Kepler's:

$$H = \frac{1}{2}\left(p_r^2 + \frac{\mathcal{L}^2}{r^2}\right) - \frac{A}{r} \tag{16}$$

and this makes its dynamics locally equivalent to Kepler's. This link between the two models gives us a direct method to demonstrate that Manev's model possess exactly the same symmetry algebra as Kepler's. The fact that symplectic structure is the same in both charts thus means that for every pair of phase space functions we will have:

$$\{F(r,p_r,\theta,L),G(r,p_r,\theta,L)\} = \{F(r,p_r,\vartheta,\mathcal{L}),G(r,p_r,\vartheta,\mathcal{L})\} \tag{17}$$

and hence from any (Poisson brackets) algebra of the Kepler model we can immediately produce identical algebra of the Manev problem by just taking the same functions depending now on the variables ϑ and $\mathcal{L}$. This observation may be viewed as a minor extension of the following:

Newton's Revolving Orbits Theorem[1]. *Let* $r(\theta)$ *be an orbit generated by any central force* $F(r)$. *Then the revolving orbit* $r(\theta') = r(\alpha\theta)$ *is generated by a central force* $F'(r)$ *that*

[1] Comments on the original Newton's wording may be found in [18] and a standard exposition of this matter in [22].

differs from $F(r)$ *by an inverse-cube force, and conversely. In particular, if* L *and* L' *are the angular momenta corresponding to* $r(\theta)$ *and* $r(\theta')$*, respectively, then*

$$F'(r) = F(r) + \frac{L^2 - L'^2}{r^3} \quad and \quad \alpha = \frac{L'}{L}.$$

In our notations $\alpha = \nu$ and $2B = L^2 - L'^2$, and we may state that: all Poisson brackets (and hence algebras) for the *prime*-system are identical to the ones of the original system provided we replace the arguments of the phase space functions L and θ with $\mathcal{L} = L'$ and $\vartheta = \alpha\theta$.

Applying this to Kepler's invariants and defining:

$$K_1 = \frac{J_+ + J_-}{2\sqrt{|2H|}}, \qquad K_2 = \frac{\mathrm{i}(J_+ - J_-)}{2\sqrt{|2H|}}, \qquad K_3 = \mathcal{L} \tag{18}$$

we obtain $\mathfrak{so}(3)$ or $\mathfrak{so}(2,1)$ algebra:

$$\{K_1, K_2\} = \mathrm{sign}(-H) K_3\,, \quad \{K_2, K_3\} = K_1\,, \quad \{K_3, K_1\} = K_2 \tag{19}$$

with Casimir invariant:

$$C \equiv K_1^2 + K_2^2 + \mathrm{sign}(-H)\, K_3^2 = \frac{A^2}{|2H|} \tag{20}$$

and so, the space of first integrals for fixed value of the Hamiltonian is a sphere or hyperboloid (which degenerate to a point or cone if $A = 0$) i.e. exactly the same as in the Kepler model.

It may seem puzzling that ϑ is not even an angular type of coordinate (i.e. a coordinate ϕ which does not exist globally but '$\mathrm{d}\phi$' is still well defined closed 1-form). In our case even $\mathrm{d}\vartheta = \nu(L)\mathrm{d}\theta + \frac{\theta}{\nu(L)}\mathrm{d}L$ is not well defined globally due to the second term, but $\mathrm{d}\mathcal{L} \wedge \mathrm{d}\vartheta$ still makes sense. On the other hand ϑ (together with r) are the natural coordinates for our configuration space factored by the action of $\mathcal{C}_{km}$: $(\mathbb{R}^2 \backslash \{0\})/\mathcal{C}_{km}$. $\mathcal{L}$ is the canonically conjugate of ϑ and so $\{r,\ \vartheta,\ p_r,\ \mathcal{L}\}$ are the natural coordinates for our reduced phase space $\mathcal{M}_{red} \equiv \mathcal{M}/\mathcal{C}_{km} = T^*\big((\mathbb{R}^2 \backslash \{0\})/\mathcal{C}_{km}\big)$.

Thus we have common invariants and symmetry algebras on both $\mathcal{M}$ and $\mathcal{M}_{red}$. This bears some resemblance to the situation with the anisotropic two-dimensional harmonic

oscillator having rational proportion between the frequences [1]. There we also have a discrete group acting on the phase space, common invariants and symmetry algebras on both the phase space and the reduced one, and equations of motion for the natural coordinates of the reduced phase space having the form of the equations for *isotropic* harmonic oscillator. What is different is that frequence ratio is not fixed for all initial data (except $\{0\}$) but changes from one L-level set to another.

5. UNBOUNDED MOTION CASES

Here the additional invariants are *always* globally defined. For completeness we shall list them below together with their corresponding symmetry algebras.

1. When $0 \neq L^2 > 2B$ and $H \geq 0$ the additional invariants are having the form and symmetry algebras just described.

2. When $L^2 = 2B$ we have the first integral:

$$j = Lp_r + A\theta \tag{21}$$

satisfying $\{H, j\} = 0, \{L, j\} = A$.

3. When $0 \neq L^2 < 2B$ we may denote $\upsilon = \mathrm{i}\nu$ with υ real and

$$\mathcal{J}_{\pm} = \upsilon L \left[p_r \pm \left(\upsilon p_{\perp} + \frac{A}{\upsilon L} \right) \right] \mathrm{e}^{\pm \upsilon \theta} \tag{22}$$

will be first integrals for *any* L. In this case (which has no direct analogue in the Kepler mechanics) we can again introduce new Darboux chart denoting $\mathcal{L} = \upsilon(L)L$ and $\vartheta = \upsilon(L)\,\theta$. When written in the new coordinates Manev's Hamiltonian takes the form:

$$H = \frac{1}{2}\left(p_r^2 - \frac{\mathcal{L}^2}{r^2} \right) - \frac{A}{r} \tag{23}$$

and we can define:

$$K_1 = \frac{\mathcal{J}_+ + \mathcal{J}_-}{2\sqrt{|2H|}}, \qquad K_2 = \frac{\mathcal{J}_+ - \mathcal{J}_-}{2\sqrt{|2H|}}, \qquad K_3 = \mathcal{L} \tag{24}$$

to obtain $\mathfrak{so}(2,1)$ algebra:

$$\{K_1, K_2\} = \text{sign}(-H) K_3\,, \quad \{K_2, K_3\} = K_1\,, \quad \{K_3, K_1\} = -K_2 \tag{25}$$

for both choices of the sign of H. Its Casimir invariant is:

$$C \equiv K_1^2 - K_2^2 + \text{sign}(-H)\, K_3^2 = -\frac{A^2}{|2H|} \tag{26}$$

and thus the space of invariants is 1- or 2-sheet hyperboloid.

6. CONCLUSIONS

We have demonstrated the existence of a *local* Darboux chart for the Manev model such that its dynamics becomes locally equivalent to the Kepler model when $L^2 > 2B$. This explains why we observe so many similarities between these two models and especially why they have common symmetry algebras. The existence of such local chart is connected with the presence of a certain discrete symmetry in the case of a bounded motion and the factoring out its action on the phase space.

To summarise, Manev model is still an active field of research with new developments coming up. And it is still more popular abroad than in Bulgaria …

ACKNOWLEDGMENT

The present talk is mostly based on joint works with Professors V. Gerdjikov, G. Marmo and G. Vilasi.

REFERENCES

[1] Amiet J.-P., Weigert S., *Commensurate Harmonic Oscillators: Classic Symmetries,* J. Math. Phys. **43** (2002) 4110–4126.

[2] Blaga C., *Precessing Orbits, Central Forces and Manev Potential,* In: Prof. G. Manev's Legacy in Contemporary Astronomy, Theoretical and Gravitational Physics, V. Gerdjikov, M. Tsvetkov (Eds), Heron Press, Sofia 2005, pp. 134–139.

[3] Brajamani S, and Singh C., *Single-variable Realization of the* $SU(1,1)$ *Spectrum*

Generating Albebra and Discrete Eigenvalue Spectra of a Class of Potentials, J. Phys. A **23** (1990) 3421–3430.

[4] Caballero J. and Elipe A., *Universal Solution for Motions in a Central Force Field,* Astron. Astrophys. Transact. **19** (2001) 869–874.

[5] Cordero P. and Ghirardi G., *Search for Quantum Systems with a Given Spectrum-Generating Algebra: Detailed Study of the Case of* $SO_{2,1}$, Nuovo Cim. **A2** (1971) 217–236.

[6] Delgado J., Diacu F., Lacomba E., Mingarelli A., Mioc V., Perez E. and Stoica C., *The Global Flow of the Manev Problem,* J. Math. Phys. **37** (1996) 2748–2761.

[7] Diacu F., Mingarelli A., Mioc V. and Stoica C., *The Manev Two-body Problem: Quantitative and Qualitative Theory,* In: Dynamical Systems and Applications, World Sci. Ser. Appl. Anal. **4**, World Sci. Publ., River Edge NJ 1995, pp. 213–227.

[8] Diacu F., Mioc V. and Stoica C., *Phase-space Structure and Regularisation of Manev-type Problems,* Nonlinear Analysis **41** (2000) 1029–1055.

[9] Ermanno, G., *Metodo d'investigare l'Orbite de Pianeti, nell'ipotesi che le forze centrali o pure le gravità degli stessi Pianeti sono in ragione reciproca de quadrati delle distanze, che i medesimi tengono dal Centro, a cui si dirigono le forze stesse,* G. Lett. Ital. **2** (1710) 447–467.

[10] Evans N., *Superintegrability in Classical Mechanics,* Phys. Rev. A **41** (1990) 5666–5676.

[11] Evans N., *Group Theory of the Smorodinsky-Winternitz System,* J. Math. Phys. **32** (1991) 3369–3375.

[12] Kyuldjiev A., Gerdjikov V., Marmo G. and Vilasi G., *On Superintegrability of the Manev Problem and its Real Form Dynamics,* In: Proceedings of Sixth International Conference on Geometry, Integrability and Quantization, I. Mladenov and A. Hirshfeld (Eds), SOFTEX, Sofia 2005, pp. 262–275;

Gerdjikov V., Kyuldjiev A., Marmo G. and Vilasi G., *Superintegrability in the Manev Problem and its Real Form Dynamics,* In: Prof. G. Manev's Legacy in Contemporary Astronomy, Theoretical and Gravitational Physics, V. Gerdjikov and M. Tsvetkov (Eds), Heron Press, Sofia 2005, pp. 155–166.

[13] Kyuldjiev A., Gerdjikov V., Marmo G. and Vilasi G., *On the Symmetries of the Manev*

Problem and its Real Hamiltonian Form, In: Proceedings of Eighth International Conference on Geometry, Integrability and Quantization, I. Mladenov and Manuel de Leon (Eds), SOFTEX, Sofia 2007, pp. 221–233.

[14] Lacomba E., Llibre J. and Nunes A., *Invariant Tori and Cylinders for a Class of Perturbed Hamiltonian Systems,* In: The Geometry of Hamiltonian Systems (Math. Sci. Res. Inst. Publ. vol 22), Springer, New York 1991, pp. 373–385;

Llibre J., Teruel A., Valls C. and de la Fuente A., *Phase Portraits of the Two-body Problem with Manev Potential,* J. Phys. A **34** (2001) 1919–1934.

[15] Leach P., Andriopoulos K. and Nucci M., *The Ermanno-Bernoulli Constants and Representations of the Complete Symmetry Group of the Kepler Problem,* J. Math. Phys. **44** (2003) 4090–4106.

[16] Leach P. and Flessas G., *Generalisations of the Laplace-Runge-Lenz Vector,* J. Nonlinear Math. Phys. **10** (2003) 340–423.

[17] Lynden-Bell D., *Hamilton's Eccentricity Vector Generalized to Newton Wonders,* Observatory **126** (2006) 176–182.

[18] Lynden-Bell D. and Lynden-Bell R., *On the Shapes of Newton's Revolving Orbits,* Notes Rec. R. Soc. Lond. **51** (1997) 195–198.

[19] Maneff G., *La gravitation et le principe de l'égalité de l'action et de la réaction,* C. R. Acad. Sci. Paris **178** (1924) 2159–2161;

Die Gravitation und das Prinzip von Wirkung und Gegenwirkung, Zeit. Phys. **31** (1925) 786–802;

Le principe de la moindre action et la gravitation, C. R. Acad. Sci. Paris **190** (1930) 963–965;

La gravitation et l'énergie au zéro, C. R. Acad. Sci. Paris **190** (1930) 1374–1377.

[20] Mioc V. and Stoica C., *Discussion et résolution complète du problème des deux corps dans le champ gravitationnel de Maneff I, II,* C. R. Acad. Sci. Paris **320**, Série II (1995) 645–648; ibid. **321**, Série I (1995) 961–964.

[21] Mioc V. and Stoica C., *On the Manev-type Two-body Problem,* Baltic Astr. **6** (1997) 637–650.

[22] Whittaker E., *A Treatise on the Analytical Dynamics of Particles and Rigid Bodies,* 4^{th} edition, Cambridge Univ. Press, 1988.

ON THE OPTIMAL DESIGN OF NON-UNIFORM CIRCULAR ANTENNA ARRAYS

Nihad Dib and Ashraf Sharaqa

Electrical Eng. Dept., Jordan Univ. of Science & Technology, P. O. Box 3030,
Irbid 22110, Jordan.
e-mail: nihad@just.edu.jo , elengashraf@yahoo.com

Abstract
In this paper, the design of non-uniform circular antenna arrays of isotropic radiators with optimum side lobe level reduction is investigated. Two global evolutionary optimization methods (namely; the biogeography based optimization and the self-adaptive differential evolution) are used to determine an optimum set of weights and positions that provide a radiation pattern with optimum side lobe level reduction with the constraint of a fixed major lobe beamwidth. The results obtained from these two evolutionary are compared with those obtained using the Matlab function Fmincon which uses a sequential quadratic programming (SQP) method. The comparison shows that the design of non-uniform circular antenna arrays using SQP method provides a side lobe level reduction that is comparable to that obtained using global stochastic optimization methods.

Keywords: Antenna arrays, Circular arrays, Optimization methods, Biogeography based optimization.

1. INTRODUCTION

Among the different types of antenna arrays, recently, circular antenna arrays have become more popular in mobile and wireless communications [1-4]. In contrast to linear antenna arrays, the radiation pattern of circular arrays inherently covers the entire space; the main lobe could be oriented in any desired direction. For the design of circular arrays, one has to adequately choose the number of antennas in the array, their positions along the circle, the circle's radius, and the feeding currents (amplitudes and phases) of the antenna elements. In general, the circular array optimization problem is more complicated than the linear array optimization. To provide a very directive pattern, it is necessary that the fields from the array elements add constructively in some desired directions and add destructively in other directions. This is important to reduce interference from the side lobes of the antenna. Thus, the design of circular antenna arrays with minimum side lobes levels has been a subject of very much interest in the literature. To accomplish this, different well-known global evolutionary optimization techniques; particle swarm optimization (PSO), genetic algorithm (GA), invasive weed optimization (IWO), and differential evolution (DE); have been used in the synthesis of non-uniform circular antenna arrays [5-10].

In this paper, two different global optimization methods, the biogeography based optimization (BBO) and the self-adaptive differential evolution (SADE), are used to determine an optimum set of weights and antenna element separations for non-uniform circular antenna arrays that provide a radiation pattern with minimum side lobe level for a fixed major lobe beamwidth. Moreover, the Matlab function *Fmincon*, which is based on the sequential quadratic programming (SQP) method, is used to perform the same design. It is shown that the results obtained using the SQP method are comparable to (and sometimes better than) those obtained using the BBO and SADE.

Biogeography-based optimization (BBO) is a new method to solve optimization problems [11-13]. BBO is based on the science of biogeography which is the nature's way of species distribution. It is modeled after the immigration and emigration of species between islands in search of more friendly habitats. A habitat is any island (area) that is geographically isolated from other islands. Islands that are well suited as residences for biological species are said to have a high habitat suitability index. The variables that characterize habitability are called suitability index variables, which are considered as the independent variables of the habitat. The habitat suitability index can be calculated using these variables. BBO has already proven itself as a valuable optimization technique compared to other already developed techniques [11-13].

Recently, the BBO has been successfully applied in optimal power flow problems [14-16]. In the electromagnetics area, BBO has been applied to the optimal design of Yagi-Uda antenna [17], the calculation of the resonant frequencies of rectangular and circular microstrip patch antennas [18, 19], and linear antenna array synthesis [20]. Here, the BBO is applied to design non-uniform circular antenna arrays with minimum side lobe levels. Moreover, the differential evolution (DE) with competitive control-parameter setting technique (*debr18.m*) [21] is used to perform the optimization for the same design problems. This is a self-adaptive DE (SADE) in which the setting of the control parameters is made adaptive through the implementation of a competition into the DE algorithm. Very recently an application of the BBO for the design of non-uniform circular arrays appeared in [22], in which the BBO results were compared to GA [7] and PSO [8] results. In that Reference [22], BBO was successful in obtaining circular arrays with lower side lobe levels (SLL) and narrower beam widths than the circular arrays obtained using the GA and the PSO methods. However, this was performed for sizes (circumferences) of the designed arrays larger (by a factor of 1.5-2) than the PSO and GA designed ones. In this paper, the BBO designed circular

arrays have almost the same size (and thus the same major lobe beamwidth) as the PSO and GA designed ones.

The objectives of this paper are twofold: first; it is shown that the newly proposed BBO method gives results that are as good as other evolutionary well-developed techniques. Second, it is shown that the results of the SQP method are also comparable to those obtained using global optimization methods; which indicates that global optimization methods are not really needed in this class of problems [23]. This paper is divided as follows: in section 2**,** the geometry and the array factor for the non-uniform circular antenna array are presented. In section 3, the fitness (or cost) function is given. In section 4, the BBO algorithm is briefly described; the reader can consult the references cited above for the full details of the BBO algorithm, and [24] to obtain the basic BBO Matlab codes. Finally, several design examples are presented in section 5.

2. GEOMETRY AND ARRAY FACTOR

Figure 1 shows the geometry of a circular antenna array (CAA) with *N* isotropic antenna elements placed non-uniformly on a ring (of radius *a*) lying in the *x-y* plane. Since isotropic elements are assumed, the radiation pattern of this array can be described by its array factor. In the *x-y* plane, the array factor for this CAA is given as follows [1]:

$$AF(\phi)=\sum_{n=1}^{N} I_n \exp\{j[ka\cos(\phi-\phi_n)+\alpha_n]\} \tag{1}$$

where

$$ka=\frac{2\pi}{\lambda}a=\sum_{i=1}^{N} d_i \tag{2}$$

$$\phi_n=\frac{2\pi\sum_{i=1}^{n} d_i}{ka} \tag{3}$$

$$\alpha_n=-ka\cos(\phi_0-\phi_n) \tag{4}$$

In the above equations, I_n and α_n represent the excitation amplitude and phase of the n^{th} element. Moreover, d_n represents the arc separation (in terms of wavelength) between element n and element $n-1$ (d_1 being the arc distance between the first (n=1) and last (n=N) elements), ϕ_n is the angular position of the n^{th} element in the x-y plane, ϕ is the azimuth angle measured from the positive x-axis, and ϕ_0 is the direction of the main beam. In our design problems, ϕ_0 is chosen to be 0, *i. e.*, the peak of the main beam is directed along the positive x axis.

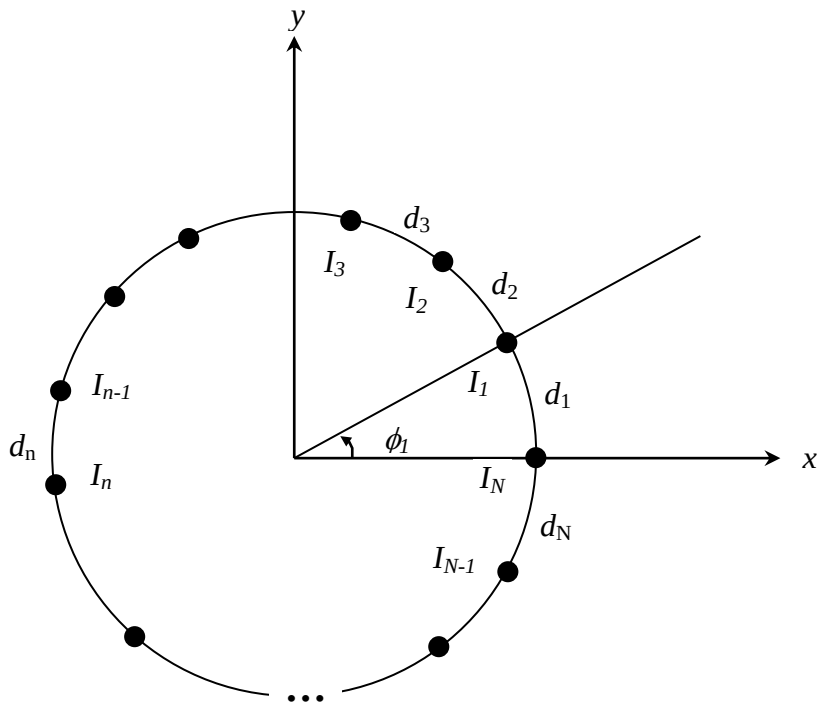


Figure 1. Geometry of a non-uniform circular antenna array with N isotropic antennas.

3. FITNESS FUNCTION

In antenna array problems, there are many parameters that can be used to evaluate the fitness (or cost) function such as gain, side lobe level, radiation pattern, and size. Here, the goal is to design a CAA with minimum side lobes levels for a specific first null beamwidth (FNBW). Thus, the following fitness function is used:

$$Fitness = (W_1 F_1 + W_2 F_2) / |AF_{\max}|^2 \tag{5}$$

$$F_1 = \left|AF(\phi_{nu1})\right|^2 + \left|AF(\phi_{nu2})\right|^2 \tag{6}$$

$$F_2 = Max\{ \left|AF(\phi_{ms1})\right|^2 , \left|AF(\phi_{ms2})\right|^2 \} \tag{7}$$

where

ϕ_{nu} is the angle at a null. Here, the array factor is minimized at the two angles ϕ_{nu1} and ϕ_{nu2} defining the major lobe, i.e., the first null beam width (FNBW) = $\phi_{nu2} - \phi_{nu1} = 2\ \phi_{nu2}$

ϕ_{ms1} and ϕ_{ms2} are the angles where the maximum side lobe level is attained during the optimization process in the lower band (from -180^o to ϕ_{nu1}) and the upper band (from ϕ_{nu2} to 180^o), respectively. An increment of 1^o is used in the optimization process. Thus, the function F_2 minimizes the maximum side lobe level around the major lobe.

Moreover, $AF_{\max}$ is the maximum value of the array factor, i.e., its value at ϕ_0. W_1 and W_2 are weighting factors; both are chosen here to be unity. Thus, the optimization problem is to search for the current amplitudes (I_n's) and the arc distances between the elements (d_n's) that minimize the fitness function in equation (5).

4. BIOGEOGRAPHY BASED OPTIMIZATION

Biogeography-Based Optimization (BBO) is based on the science of biogeography which is the nature's way of species distribution. It is modeled after the immigration and emigration of species between islands in search of more friendly habitats. A habitat is any island (area) that is geographically isolated from other islands.

Suppose that we are presented with a global optimization problem and some candidate solutions. The candidate solutions of a problem are represented by an array of integers as $\text{Habitat} = [\text{SIV}_1, \text{SIV}_2, \text{SIV}_3, \cdots\cdots, \text{SIV}_\text{N}]$. The variables in the array that characterize habitability are called suitability index variables (SIVs), which are considered as the independent variables of the habitat.

The value of the fitness function in BBO is called habitat suitability index (HSI) which is found by evaluating the fitness function:

$$fitness(Habitat) = HSI = f(SIV_1, SIV_2, SIV_3, \cdots\cdots, SIV_N) \tag{8}$$

A good solution and a weak solution are similar to an island with a high habitat suitability index (HSI) and a low HSI, respectively. A good solution shows high resistance more than a weak solution to accept change. Weak solutions accept many features from good solutions which tend to share them with it. Shared features with low HSI do not mean that, they will disappear from high HSI; they remain and appear as new features in low HSI. This is similar to species migrating to a new habitat, while some species still remain in their original habitat. The BBO algorithm consists of three steps: creating a set of solutions to the problem, where they are randomly selected, and then applying migration and mutation steps to reach the optimal solution.

In the migration step, equations (9) and (10) are used to evaluate the immigration rate (λ) and the emigration rate (μ) of each solution, respectively, which are used to probabilistically share information between habitats with probability P_{mod} (P_{mod} known as the habitat modification probability).

$$\lambda = I\left(1 - \frac{S}{S_{max}}\right) \tag{9}$$

$$\mu = E\left(\frac{S}{S_{max}}\right) \tag{10}$$

The following flow chart summarizes the migration process:

```
For i = 1 to n  (where n is the number of islands)
  Select H_i with probability ∝ λ_i
        If H_i is selected
                For j = 1 to n
  Select H_j with probability ∝ μ_j
                        If H_j is selected
                                Randomly select an SIV σ from H_j
                                Replace a random SIV in H_i with σ
                        End
                End
        End
End
```

The mutation step tends to increase the diversity among the population and gives the solutions the chance to improve their selves to the best. Performing mutation on a solution is done by replacing it with a new solution that is randomly generated. The following flow chart summarizes the mutation process:

```
For i = 1 to n
        For j = 1 to N (where N is the number of variables)
          Use λ_i and μ_i to compute the probability P_i
          Select SIV H_i(j) with probability ∝ P_i
            If H_i(j) is selected
                    Replace H_i(j) with a randomly generated SIV
            End
        End
End
```

5. RESULTS

Several examples with different number of antenna elements (N=8, 10, 12) are optimized using the BBO, SADE, and SQP methods. The description of the implementation of *Fmincon* function can be found in the corresponding MATLAB help file. It should be noted that the SQP method is not a stochastic method. The results of the SQP method depend mainly on the initial estimate. In our design examples, we set the initial estimate to be a random vector using the *rand* function in Matlab. For comparison purposes, the same design examples presented in [7, 8] are used here. In [7], genetic algorithm was used for the design of CAAs, while the PSO was used in [8].

In the BBO implementation, for N=8 and N=10 cases, the following parameters are used: population size = 50, number of generations = 300, habitat modification probability = 1, mutation probability = 0.01 and elitism parameter = 2. The same parameters are used for N=12 case except for using a population size = 150 and number of generations = 500. The minimum and maximum allowable values for the variables (i.e., the weights and the inter-element arc distances) are set to 0.1 and 1, respectively. The design examples are performed for a specific FNBW, which corresponds to a uniformly-fed circular array with a uniform $\lambda/2$ element-spacing and the same number of elements. BBO, SADE and SQP codes are run for 10 independent times. Tables 1-3 show the best results obtained using BBO, SADE and SQP, respectively. "*Best results*" are defined as the ones that give the smallest value of the fitness function. For the BBO results in Table 1, the values of the fitness function were 0.0674, 0.0572, and 0.0405 for N=8, 10, and 12, respectively. On the other hand, for the SADE results in Table 2, the values of the fitness function were 0.0586, 0.0503, and 0.0501 for N=8, 10, and 12, respectively. Finally, for the SQP results in Table 3, the values of the fitness function were 0.0547, 0.055, and 0.049 for N=8, 10, and 12, respectively. The current amplitudes for the array elements are normalized such that $max(I)$=1. As mentioned above, the same examples were considered in [7] and [8] using GA and PSO, respectively.

Table 1. Examples of non-uniform circular antenna array optimized using BBO.

N	ϕ_{nu2} (deg)	[dm_1, dm_2, dm_3, …, dm_N] in λ's [I_1, I_2, I_3, …, I_N]
8	34	[0.3406, 0.7682, 0.2988, 0.5756, 0.6627, 0.8805, 0.6337, 0.4214] ➔ ∑ = 4.5814 [0.7637, 0.6075, 0.1090, 1.0000, 0.8722, 0.5396, 0.7177, 0.4858]
10	27	[0.387, 0.9088, 0.3232, 0.2549, 0.8932, 0.5083, 0.8781, 0.6733, 0.88, 0.3498] ➔ ∑ = 6.0565 [0.8848, 0.5265, 0.3690, 0.3744, 1.0000, 1.0000, 0.6374, 0.5803, 0.8792, 0.5606]
12	23	[0.4083, 0.6416, 0.7554, 0.7185, 0.6943, 0.3818, 0.3284, 0.8152, 0.9981, 0.3097, 0.7983, 0.3701] ➔ ∑ = 7.2196 [0.6567, 0.3879, 0.6960, 0.4596, 0.5627, 0.9600, 0.4168, 0.5890, 0.5368, 0.6230, 0.6910, 1.0000]

Table 2. Examples of non-uniform circular antenna array optimized using SADE.

N	ϕ_{nu2} (deg)	[dm_1, dm_2, dm_3, …, dm_N] in λ's [I_1, I_2, I_3, …, I_N]
8	34	[0.3438, 0.6668, 0.2059, 0.7951, 0.6272, 0.8437, 0.8295, 0.3383] ➔ ∑ = 4.6505 [0.8749, 0.2302, 0.4633, 0.9542, 1.0000, 0.6442, 0.9099, 0.1844]
10	27	[0.2775, 0.9516, 0.5141, 0.9865, 0.6166, 0.9703, 0.2755, 0.2648, 0.8826, 0.3137] ➔ ∑ = 6.0532 [0.9333, 0.5834, 0.4528, 1.0000, 0.9620, 0.3544, 0.2959, 0.4202, 0.8792, 0.1412]
12	23	[0.4272, 0.6798, 0.6380, 0.6954, 0.9017, 0.5223, 0.7686, 0.6235, 0.2582, 0.5151, 0.8151, 0.3343] ➔ ∑ = 7.1793 [0.3617, 0.3740, 0.3498, 0.6514, 1.0000, 0.8604, 0.4864, 0.3960, 0.3696, 0.3390, 0.5058, 0.8387]

Table 3. Examples of non-uniform circular antenna array optimized using SQP.

N	ϕ_{nu2} (deg)	[dm_1, dm_2, dm_3, …, dm_N] in λ's [I_1, I_2, I_3, …, I_N]
8	34	[0.3192, 0.3867, 0.4809, 0.8277, 0.6450, 0.8066, 0.8573, 0.3287] ➔ ∑ = 4.6521 [0.8849, 0.1438, 0.5516, 1.0000, 0.9998, 0.6233, 0.9158, 0.1053]
10	27	[0.3311, 0.4761, 0.5888, 0.3355, 1.0000, 0.5818, 0.9346, 0.7570, 0.7405, 0.2865] ➔ ∑ = 6.0320 [0.9288, 0.1237, 0.3838, 0.5450, 1.0000, 0.8450, 0.7054, 0.4424, 0.8559, 0.1589]
12	23	[0.4177, 0.5963, 0.7442, 0.7173, 0.7994, 0.4433 , 0.8958 , 0.7129 , 0.7622 , 0.4557, 0.1607, 0.3786] ➔ ∑ = 7.0840 [0.4685, 0.4221, 0.6701, 0.3644, 0.6449, 0.8077, 0.4698, 0.4986, 0.5102, 0.1341, 0.2746, 1.0000]

Figure 2 shows the array factor obtained using the results in Tables 1-3 for N = 8. The maximum side lobe level obtained using the BBO, SADE, and SQP are -12.18 dB, -12.7 dB, and -13.16 dB respectively. It should be noted that these values are better than those obtained using GA [7] and PSO [8]. Specifically, the maximum side lobe levels obtained using GA and PSO were -9.8 dB and -10.8 dB, respectively. It is worth mentioning that a uniform circular array with the same number of elements and λ/2 element-to-element spacing has a maximum side lobe level of -4.17 dB.

Similarly, Figure 3 shows the array factor for N = 10 CAA. Again, SQP results are as good as those obtained using the global optimization BBO and SADE methods. Moreover, the results shown in Figure 3 are, in general, better than GA and PSO results presented in [7] and [8]. A uniform circular array with the same number of elements and λ/2 element-to-element spacing has a maximum side lobe level of -3.6 dB. Lastly, Figure 4 shows a comparison between the array factors obtained using the different optimization methods for N = 12.

These figures clearly show the effectiveness of the newly-proposed BBO in solving circular antenna array problems. Moreover, they show that the Matlab function *Fmincon* which implements the SQP method is as powerful as the stochastic global optimization techniques in designing non-uniform CAAs with optimum side lobe level.

It should be mentioned here that for the *N*=8 case, the PSO-optimized CAA had a circumference of 4.4931λ [8]; while that obtained using the GA had a circumference of 4.4094λ [7]. The BBO, SADE and SQP-optimized CAAs, shown in Tables 1-3, have slightly larger circumference. Moreover, the circumferences of the PSO and GA-designed *N*=10 CAAs were 5.9029λ and 6.0886λ, respectively, while the PSO and GA-designed *N*=12 CAAs had circumferences of 7.1501λ and 7.77λ, respectively [7, 8]. The circumferences obtained in Tables 1-3, using BBO, SADE and SQP, are very close to these values.

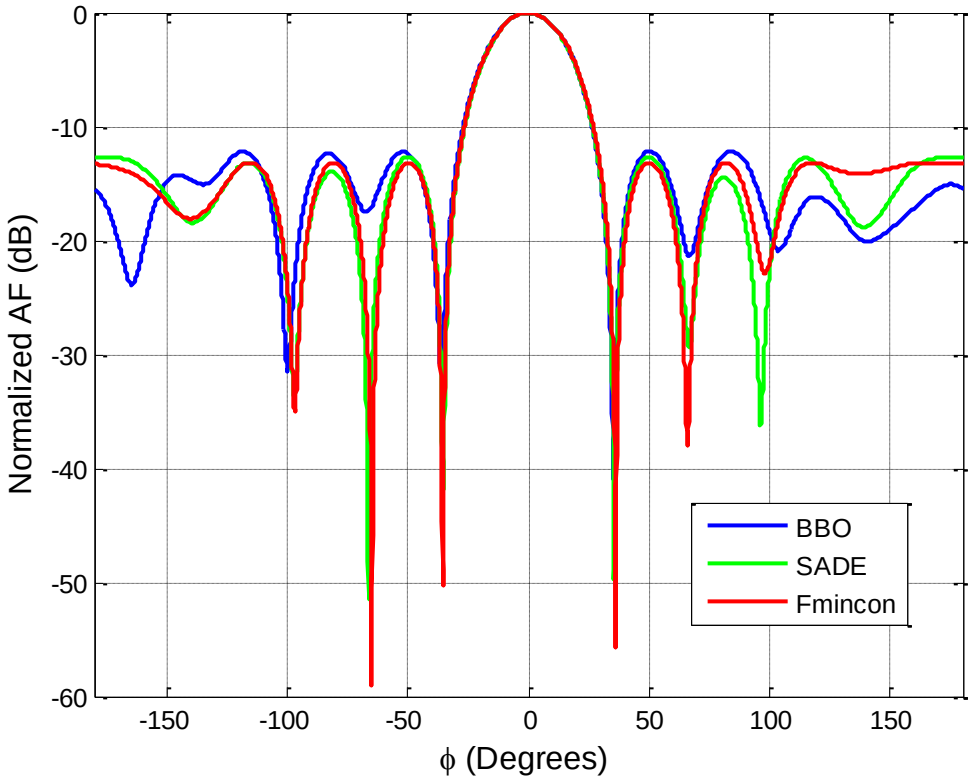


Figure 2. Radiation pattern for *N*=8 using the BBO, SADE and SQP results in Tables 1-3.

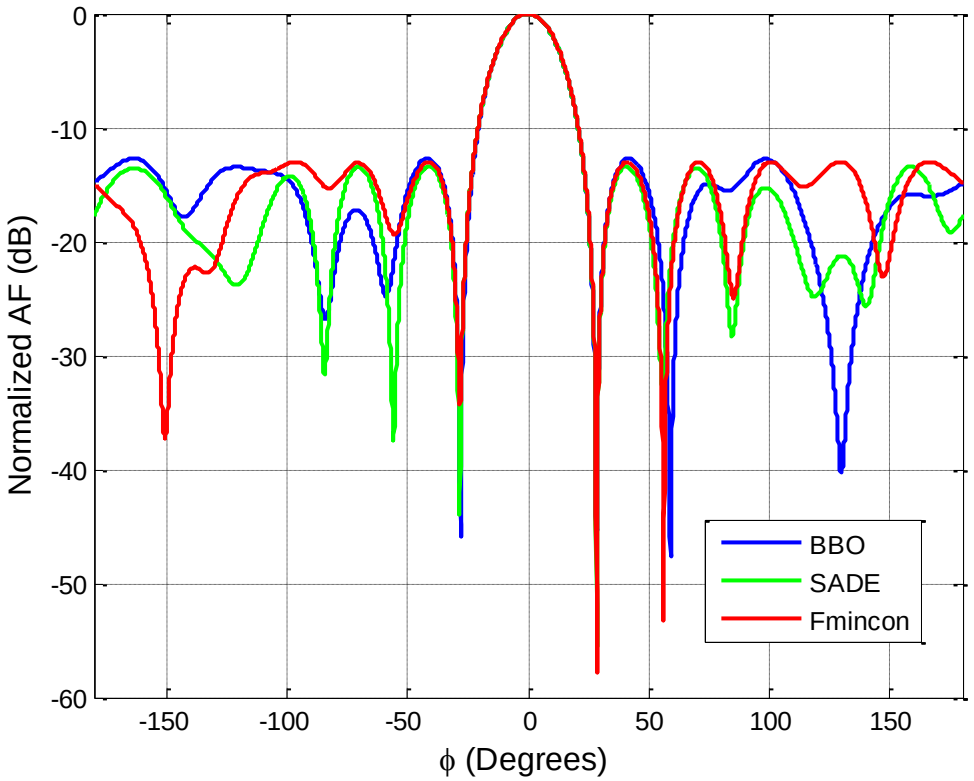


Figure 3. Radiation pattern for N=10 using the BBO, SADE and SQP results in Tables 1-3.

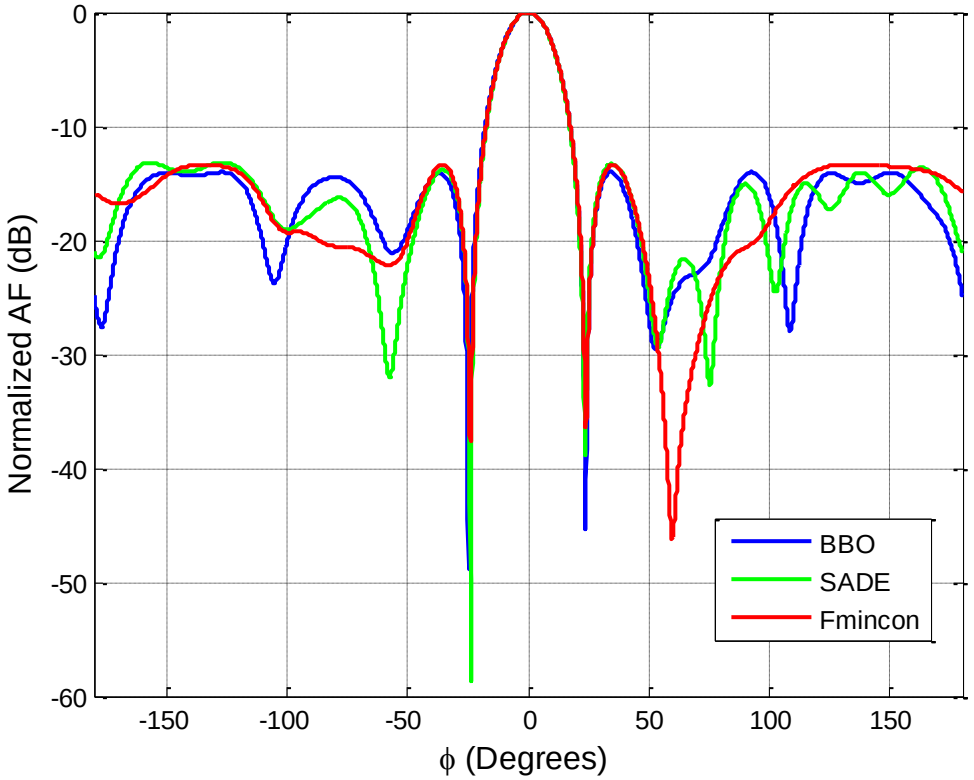


Figure 4. Radiation pattern for N=12 using the BBO, SADE and SQP results in Tables 1-3.

As a last more challenging example, Figure 5 shows the array factor for an optimized CAA with *N*=20. Here, a population size = 150, and number of generations = 500 are used for the BBO design, while the other parameters are as those mentioned above. BBO, SADE and SQP optimized CAAs along with the array factor of a conventional CAA are shown in the figure. Each algorithm was run for 10 independent times, and the results with the smallest fitness function value are used to plot the array factor. The conventional array consists of 20 uniformly fed elements with λ/2 element-to-element spacing. The maximum side lobe level for the conventional array is -6.08 dB, while the maximum SLL for the BBO and SADE optimized CAAs is -13.84 dB (an improvement of almost 7.7 dB). On the other hand, the maximum side lobe level obtained using the SQP method is -14.87 dB. Again, BBO, though still a new optimization algorithm, proves to be a powerful optimization technique compared to other already developed techniques, like the DE. Moreover, the SQP method gives slightly better maximum side lobe level compared to BBO and SADE. Table 4 shows the obtained weights and inter-element arc distances for the optimized *N*=20 CAAs.

Table 4. Weights and spacings for the optimized *N*=20 CAA.

	$[dm_1, dm_2, dm_3, \ldots, dm_N]$ in λ's $[I_1, I_2, I_3, \ldots, I_N]$
BBO	[0.4196 0.4588 1.0000 0.4406 0.6314 0.3635 0.8939 0.3215 1.0000 0.4786 0.4856 0.5848 0.4761 0.5695 0.8245 0.9013 0.7483 0.8833 0.4748 0.4417] ➔ ∑ = 12.3979 [0.8227 0.9057 0.3545 0.1653 0.7815 0.6918 0.6865 0.7171 1.0000 1.0000 0.9981 0.7308 1.0000 0.6543 0.9493 0.1000 0.5944 0.6473 0.5730 1.0000]
SADE	[0.4825 0.1795 0.1793 0.6164 0.9753 0.3628 0.6046 0.9890 0.2913 0.9223 0.5750 0.7937 0.9161 0.8519 0.9921 0.5791 0.2997 0.7233 0.5166 0.3347] ➔ ∑ = 12.1851 [0.9072 0.4465 0.1364 0.7688 0.6309 0.5683 0.5877 0.2080 0.3060 1.0000 0.9771 0.5235 0.7805 0.3250 0.7083 0.6721 0.6070 0.8726 0.6762 0.8076]
SQP	[0.5682 0.6833 0.4381 0.4709 0.7318 0.8736 0.6329 0.2492 0.9944 0.6006 0.5653 1.0000 0.8636 0.8893 0.3727 0.3865 0.4148 0.5996 0.5457 0.4453] ➔ ∑ = 12.3259 [0.8708 0.2126 0.6157 0.5070 0.2995 0.4988 0.3782 0.3621 0.8598 1.0000 0.6950 0.5762 0.7123 0.3460 0.1815 0.6140 0.3407 0.3590 0.8188 0.9923]
Conv.	[0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5 0.5] ➔ ∑ = 10 [1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1 1]

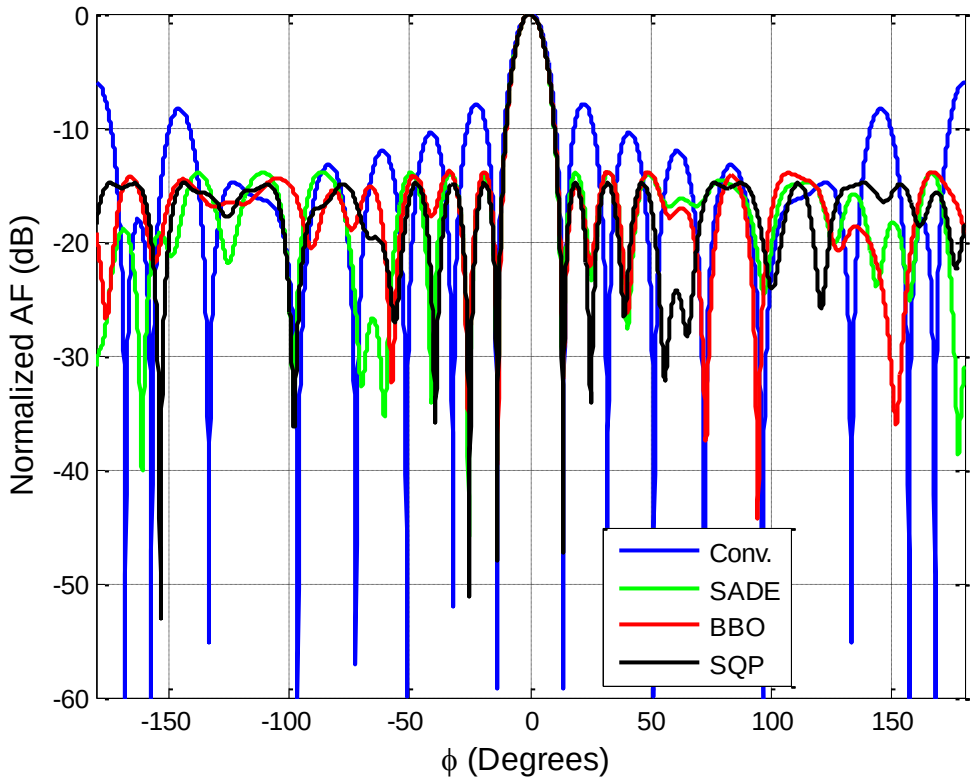


Figure 5. Radiation pattern for *N*=20 using the results in Table 4 compared to conventional CAA.

From the above Tables, it can be noticed that the optimized arrays always have an aperture (i.e., circumference) larger than that of a uniform array with half-wavelength spacing between the elements. Now, to make the comparison between a conventional uniform CAA and an optimized one more fair, one has to force the optimized array to have an aperture as close as possible to that of a uniform array. To accomplish this, in the BBO and SADE optimizations, the fitness function in equation (5) is modified as follows:

$$Fitness = (W_1 F_1 + W_2 F_2) / |AF_{\max}|^2 + F_3 \tag{10}$$

where

$$F_3 = Abs\left(\left\{\sum_{i=1}^{N} d_i\right\} - (desired\ circumference)\right) \tag{11}$$

where the *desired circumference* is that of a uniform CAA, or any other desired circumference. It should be noted that the fitness function in equation (5) is still used with the *Fmincon* optimization since the *Fmincon* function implements a constrained version of the SQP. The constraint of having a specific aperture is enforced as a constraint in a separate *m*-file (the reader can consult the help file for *Fmincon* function). As an example, the N=20 CAA is again optimized with the fitness function (8), with a desired circumference of 10λ. The obtained results are shown in Table 5, while Figure 6 shows the corresponding array factor. The conventional uniform CAA has a maximum SLL of -6.08 dB, while the BBO, SADE and SQP optimized CAAs have maximum SLLs of -10.7 dB, -11.3 dB, and -11.81 dB, respectively. The optimized CAAs have exactly the same aperture as the conventional CAA. As expected, forcing the aperture of the optimized CAA to be the same as that of a conventional CAA gives a maximum SLL less than the one obtained in Figure 5.

Table 5. Weights and spacings for the optimized N=20 CAA using the fitness function (8).

	$[dm_1, dm_2, dm_3, \ldots, dm_N]$ in λ's $[I_1, I_2, I_3, \ldots, I_N]$
BBO	[0.3992 0.9735 0.2235 0.3108 1.0000 0.2110 0.4899 0.1000 0.3585 0.4811 0.2717 0.5740 0.7261 1.0000 0.3988 0.2548 0.8606 0.5380 0.5212 0.3077] ➔ ∑ = 10 [0.5576 0.1000 0.9253 0.6877 1.0000 0.3850 0.1000 0.1000 0.3295 0.1000 1.0000 0.8999 0.5202 0.6851 0.6877 1.0000 0.4435 0.2389 0.6996 1.0000]
SADE	[0.2188 0.2808 0.7180 0.3800 0.8207 0.2932 0.9706 0.6226 0.5565 0.1364 0.4640 0.5064 0.8742 0.1964 0.8028 0.2789 0.4280 0.6616 0.4842 0.3062] ➔ ∑ = 10 [0.7398 0.4071 0.3937 0.7197 1.0000 0.9281 0.7220 0.3150 0.8843 0.7853 0.5352 0.3133 0.6459 0.9790 0.9843 0.7044 0.1857 0.3126 0.8326 0.6337]
SQP	[0.1817 0.1738 0.8299 0.7268 0.2955 0.8317 0.7392 0.8753 0.2633 0.1484 0.2256 0.9999 0.7963 0.3393 0.3556 0.4066 0.9978 0.5057 0.2006 0.1070] ➔ ∑ = 10 [0.1278 0.4208 0.1790 0.7576 1.0000 0.8536 0.3456 0.1000 0.9955 0.5949 0.3683 0.3726 0.5160 0.9991 0.5923 0.6132 0.1000 0.3723 0.3600 0.7845]

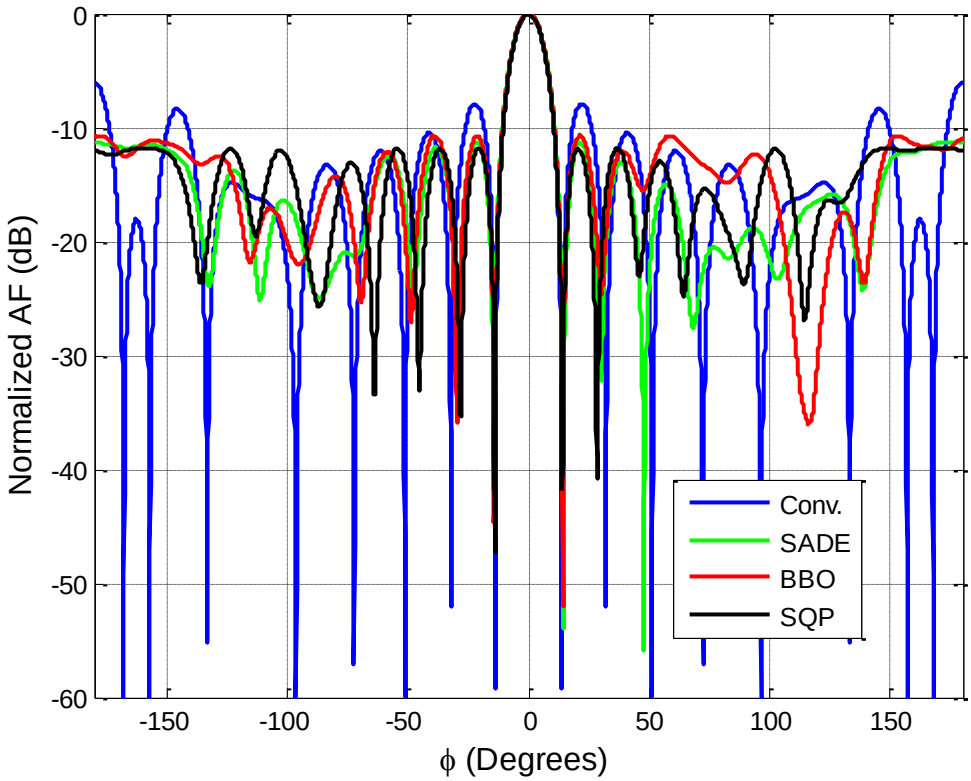


Figure 6. Radiation pattern for *N*=20 using the results in Table 5 compared to conventional CAA.

6. CONCLUSIONS

In this paper, the newly proposed BBO method was used to adjust the positions and the excitations of the antenna elements in a circular array to obtain an optimum side lobe level. The obtained optimized array factor was compared to that obtained using other well-known optimization techniques (SADE, GA, and PSO). Array factor patterns obtained from BBO results are generally as good as those presented in the literature. Moreover, the Matlab function *Fmincon*, which uses the SQP method, has been used to design the same arrays and has shown to give results that are as good as, or even better than, those obtained using global stochastic optimization methods. This indicates that for the problem under consideration (i.e., the design of CAA with optimum side lobe level), stochastic global optimization methods are not really needed [23].

ACKNOWLEDGEMENT

This work was supported by the Deanship of Research at Jordan University of Science and Technology (JUST).

REFERENCES

[1] C. A. Balanis, Antenna Theory: Analysis and Design, John Wiley & Sons, New York, 1997.

[2] S. Zainud-Deen, E. Mady, K. Awadalla, and H. Harsher, "Controlled radiation pattern of circular antenna array," *IEEE Antennas and Propagation Symp*., 2006, pp. 3399-3402.

[3] L. Pesik, D. Paul, C. Railton, G. Hilton and M. Beach, "FDTD technique for modeling eight element circular antenna array," *Electronics Letters,* Vol. 42, No. 14, pp. 787-788, July 2006.

[4] W. Mahler, and F. Landstorfer, "Design and optimization of an antenna array for WiMAX base stations," *IEEE/ACES International Conference on Wireless Communications and Applied Computational Electromagnetics*, 2005, pp. 1006 – 1009.

[5] M. A. Panduro, C. A. Brizuela, L. I. Balderas and D. A. Acosta, "A Comparison of Genetic Algorithms, Particle Swarm Optimization and the Differential Evolution Method for the Design of Scannable Circular Antenna Arrays," *Progress In Electromagnetics Research B*, Vol. 13, pp. 171-186, 2009.

[6] Panayiotis Ioannides, and Constantine A. Balanis, "Uniform Circular and Rectangular Arrays for Adaptive Beamforming Applications," *IEEE Antennas and Wireless Propagation Letters*, Vol. 4, pp. 351-354, 2005.

[7] M. Panduro, A. Mendez, R. Dominguez, and G. Romero, "Design of non-uniform circular antenna arrays for side lobe reduction using the method of genetic algorithms," *Int. J. Electron. Commun. (AEU),* Vol. 60, pp. 713-717, 2006.

[8] M. Shihab, Y. Najjar, N. Dib, and M. Khodier, "Design of NonUniform Circular Antenna Arrays Using Particle Swarm Optimization," *Journal of Electrical Engineering*, Vol. 59, No. 4, pp. 216-220, 2008.

[9] Korany R. Mahmoud, Mohamed I. Eladawy, Rajeev Bansal, Saber H. Zainud-Deen, and Sabry M. M. Ibrahem, "Analysis of Uniform Circular Arrays for Adaptive Beamforming Applications Using Particle Swarm Optimization Algorithm*," International Journal of RF and Microwave Computer-aided Engineering*, Vol. 18, No. 1, 42-52, Jan. 2008.

[10] Gourab Ghosh Roy, Swagatam Das, Prithwish Chakraborty, and Ponnuthurai N. Suganthan, "Design of Non-Uniform Circular Antenna Arrays Using a Modified Invasive Weed Optimization Algorithm," *IEEE Trans. on Antennas and Propagation,* Vol. 59, No. 1, pp. 110-118, Jan. 2011.

[11] Simon, D., "Biogeography-based optimization," *IEEE Trans. Evolut. Computat*., Vol. 12, No. 6, pp. 702–713, December 2008.

[12] D. Simon, M. Ergezer, D. Du, and R. Rarick, "Markov Models for biogeography-based optimization," *IEEE Transactions on Systems, Man, and Cybernetics - Part B: Cybernetics,* vol. 41, no. 1, pp. 299-306, January 2011.

[13] D. Simon, R. Rarick, M. Ergezer, and D. Du, "Analytical and numerical comparisons of biogeography-based optimization and genetic algorithms," *Information Sciences*, vol. 181, no. 7, pp. 1224-1248, April 2011.

[14] P. Roy, S. Ghoshal, and S. Thakur, "Biogeography based optimization for multi-constraint optimal power flow with emission and non-smooth cost function," *Expert Systems with Applications*, vol. 37, no. 12, pp. 8221-8228, December 2010.

[15] A. Bhattacharya and P. Chattopadhyay, "Hybrid differential evolution with biogeography based optimization for solution of economic load dispatch," *IEEE Transactions on Power Systems*, vol. 25, no. 4, pp. 1955-1964, November 2010.

[16] A. Bhattacharya and P. Chattopadhyay, "Solution of optimal reactive power flow using biogeography-based optimization," *International Journal of Energy and Power Engineering*, vol. 3, no. 4, pp. 269-277, November 2010.

[17] U. Singh, H. Singla, and T. Kamal, "Design of Yagi-Uda antenna using biogeography based optimization," *IEEE Transactions on Antennas and Propagation*, vol. 58, no. 10, pp. 3375-3379, October 2010.

[18] M. Lohokare, S. Pattnaik, S. Devi, B. Panigrahi, K. Bakwad, and J. Joshi, "Modified BBO and calculation of resonant frequency of circular microstrip antenna," *World Congress on Nature & Biologically Inspired Computing*, pp. 487-492, Coimbatore, India, December 2009.

[19] M. Lohokare, S. Pattnaik, S. Devi, K. Bakwad, and J. Joshi, "Parameter calculation of rectangular microstrip antenna using biogeography-based optimization," *Applied Electromagnetics Conference*, Coimbatore, India, December 2009.

[20] U. Singh, H. Kumar, and T. S. Kamal, "Linear array synthesis using biogeography based optimization," *Progress In Electromagnetics Research M*, Vol. 11, pp. 25-36, 2010.

[21] Tvrd´ık, J., V. Pavliska, and H. Habiballa, "Stochastic self-adaptive algorithms for global optimization — MATLAB and C++ library," http://albert.osu.cz/oukip/optimization/.

[22] U. Singh, and T. S. Kamal, "Design of Non-Uniform Circular Antenna Arrays Using Biogeography-Based Optimization," *IET Microwaves, Antennas and Propagation*, Vol. 5, No. 11, pp. 1365-1370, 2011.

[23] O. Bucci, M. D'Urso, and T. Isernia, "Some Facts and Challenges in Array Antenna Synthesis Problems," *AUTOMATICA- Journal for Control, Measurement, Electronics, Computing and Communications*, Vol. 49, pp. 13-20, 2008.

[24] http://embeddedlab.csuohio.edu/BBO/

DIFFERENTIAL METHODS FOR CONTOUR IMAGE SEGMENTATION

Dimitar I. Petrov[1] , Miroslav D. Petrov[2], Plamenka I. Borovska[3]

[1] Department of Mathematical Analysis and Applications, Faculty of Mathematics and Informatics, University of Veliko Tarnovo "St. St. Cyril and Methodius"
2, T.Tarnovski Str., BG-5000, Veliko Tarnovo, Bulgaria
e-mail : d.petrov@uni-vt.bg

[2] Department of Computer Systems and Technologies, Faculty of Mathematics and Informatics, University of Veliko Tarnovo "St. St. Cyril and Methodius"
2, T.Tarnovski Str., BG-5000, Veliko Tarnovo, Bulgaria
e-mail : m.petrov@uni-vt.bg

[3] Department of Computer Systems, Faculty of Computer Systems and Control, , Technical University of Sofia, 8, Kliment Ohridski Blvd., BG-1000, Sofia, Bulgaria
e-mail :pborovska@tu-sofia.bg

Abstract

The present study provides a brief overview of the differential methods for contour image segmentation. The starting criteria are the precision in defining the contour of an object in the image and the noise resistance, determined by the ratio between the number of the wrongly detected contours and the number and size of the contour discontinuities.

1. INTRODUCTION

The main problem in the image processing is the choice of proper criteria, suitable for the assigned task. When analysing, segmenting and detecting objects in a given image, most information is obtained by their outlines, called contours and characterized by a sharp change in their intensity [1,2]. The transition to processing image contours reduces considerably the amount of the information used.

The contour segmentation methods are divided into two main groups: differential and correlation-extremal. This study focuses on some segmentation methods, belonging to the first group. The principle of the differential methods is first to highlight the sharp changes in the image through differentiation and then to detect the contour by a thresholding. The result is a binary image as big as the input one. It is re-processed in order to get a pixel-thin contour. These methods are simple to use and very quick but they are not very noise-prooved.

2. CANNY EDGE DETECTION ALGORITHM

One of the most popular differential methods for contour segmentation is the Canny´s [3]. The algorithm he suggested detects the points of sharp changes in the intensity of the image $f\left(\vec{r}\right)$, $\vec{r}=(x,y)$ by calculating the vector module of the gradient $\vec{\nabla}f=\left(\frac{\partial f}{\partial x},\frac{\partial f}{\partial y}\right)$ т.е. $\left|\vec{\nabla}f\right|=\sqrt{\left(\frac{\partial f}{\partial x}\right)^2+\left(\frac{\partial f}{\partial y}\right)^2}$ and its direction $arctg\left(\left(\frac{\partial f}{\partial y}\right)\times\left(\frac{\partial f}{\partial x}\right)^{-1}\right)$. It is well known that the gradient $\vec{\nabla}f$ shows the direction of the maximal change of the function $f\left(\vec{r}\right)$. The point $\vec{r}_0\in R^2$ is determined as a point from the image edge $f\left(\vec{r}\right)$, if $\left|\vec{\nabla}f\left(\vec{r}\right)\right|$ has a local maximum at $\vec{r}=\vec{r}_0$, when $\vec{r}=\vec{r}_0+\varepsilon\vec{\nabla}f\left(\vec{r}_0\right)$ for small enough values of ε. This means that partial derivatives of $f\left(\vec{r}\right)$ have a local extremum at the point $\vec{r}_0$, when $\vec{r}$ changes in a one-dimensional area of $\vec{r}_0$ towards the fastest variation of f at $\vec{r}_0$.

The main stages in Canny´s contour segmentation method are:

- Smoothing the intensity values of the image, using a Gaussian filter in order to reduce the additive noise components;
- Estimating the gradient of the filtered image in order to emphasize the image edges;
- Thresholding the estimated image gradient in order to trace the real contour points. This includes non-maximum suppression of the intensity changes. The real contour points are detected by eliminating the ones that are not local maximum points of the gradient module. To do this the value of the intensity change in a contour edge is set to be zero, if it is not greater than the change of two of its neighbouring points in the direction of the image gradient;
- Processing the image contour morphologically (a hysteresis method), at which two threshold values of the image intensity are set. Only the neighbouring points at which the intensity values are higher than the lower threshold, are connected to the points, having values greater than the higher threshold.

3. MULTISCALE EDGE DETECTION

Irrespective of the application of the hysteresis method, the numerical differentiation leads to a reduction of the accuracy of detecting object contours in the image. The noise-resistance of Canny´s method also depends on the correlation between the threshold value and the line width of the smoothing filter. The multiscale methods of the wavelet analysis help to solve this problem. They process the image, using filters of different lengths.

The logic of this approach is based on the fact that image processing often has a locally non-homogeneous texture. That is why the detection, for example, has to be done at different detailing levels, according to the assigned task. In some cases it is enough to use the outer contour only (the outline of the object), whereas in others – its smaller details contain the information. Therefore, the analysed images have to undergo transformations that have good spatial and frequency localization.

The multiscale version of the Canny´s detector is implemented by smoothing the image with the convolution kernel $\Theta(x)$, most often dilated with the help of the dyadic line $\{2^k\}_{k\in Z}$ [4-6]. This kernel is computed with wavelets which are its partial derivatives: $\psi^1(x,y)=\dfrac{\partial\Theta(x,y)}{\partial x}$ and $\psi^2(x,y)=\dfrac{\partial\Theta(x,y)}{\partial y}$. In fact, the scales used change, according to the dyadic line $\{2^k\}_{k\in Z}$. Thus, the wavelet transform of the image $f(x)$ can be written in the following way:

$$\begin{pmatrix} W^1 f(x,y,2^k) \\ W^2 f(x,y,2^k) \end{pmatrix} = 2^k \begin{pmatrix} \dfrac{\partial}{\partial x}\left(f * \overline{\Theta}_{2^k}\right)(x,y) \\ \dfrac{\partial}{\partial y}\left(f * \overline{\Theta}_{2^k}\right)(x,y) \end{pmatrix} = 2^k \vec{\nabla}\left(f * \overline{\Theta}_{2^k}\right)(x,y),$$

where $\overline{\Theta}_{2^k}(x,y)=2^{-k}\Theta\left(-2^k x,-2^k y\right)$.

Therefore, the modulus of the gradient vector is proportional to the modulus of the wavelet transform $Mf(x,y,2^k)=\left(\left|W^1 f\left(x,y,2^k\right)^2\right|+\left|W^2 f\left(x,y,2^k\right)^2\right|\right)^{\frac{1}{2}}$ and it's angle this transform in the (x,y) plane is defined by the expression $arctg\dfrac{W^2 f\left(x,y,2^k\right)}{W^1 f\left(x,y,2^k\right)}$.

The point $(x_0,y_0)\in R^2$ is called *edge point* in the image $f(x,y)$ at a scale 2^k, if $Mf\left(x,y,2^k\right)$ has a local maximum at the one $(x,y)=(x_0,y_0)$, when (x,y) changes

within a small one-dimensional area of the point (x_0, y_0), set by the gradient direction. It`s also called wavelet transform *modulus maximum.* This is well-known that analysing wavelet transform modulus maxima of $f(x, y)$ at fine scales leads to detecting all irregularities in the image.

The image $f(x, y)$ is termed regular at the point (x_0, y_0) with an indicator $\alpha \in [0,1]$, if there is a constant $K > 0$, such that for every point $(x, y) \in R^2$ the inequality $|f(x, y) - f(x_0, y_0)| \le K\left(|x - x_0|^2 + |y - y_0|^2\right)^{\frac{\alpha}{2}}$ is true.

When solving the contour segmentation problem, by means of the wavelet transformation, the possible points of this contour are the ones at which the image has a regularity indicator $\alpha < 1$. At the edge points of the image it is estimated with the help of the inequality $\left|Mf\left(x, y, 2^k\right)\right| \le const * 2^{k(\alpha+1)}$, by measuring the slope of $\log_2 \left|Mf\left(x, y, 2^k\right)\right|$ as a function of k.

The multiscale edge detector defines them as a set of sudden image intensity change points. For contour segmentation these edges must specify closed curves, outlining approximately the boundaries of the respective areas. The presence of noise or light variations are the most often causes local detectors to make holes in the contours, whose filling requires additional processing or prior information on the image edges.

Polyakova and Lyubchenko [7] suggest contour segmentation methods, based on stationary wavelet transform. In order to make the Canny´s method more noise-prooved, that is in order to avoid wrongly detected contours, a nonlinear contrast method is used. It employs the coefficients of stationary wavelet transform. To this aim, image intesity values in rows and columns are interpolated with a certain function $f(t)$. The regularity indicator α of the function $f(t)$ is expressed by the coefficients of its discrete stationary wavelet transform d_{jk}, where j is the decomposition level, and k – the position of the wavelet coefficient. In Ref. [4] we can see that, if the function $f(t)$ is regular with an indicator α at the point $\boldsymbol{t}$, then $d_{jk} = O\left(2^{-\alpha\left(j+\frac{1}{2}\right)}\right)$, where k is the closest to the t position of the wavelet coefficient. Nason and B. Silverman [8] offer to define the singularity of function $f(t)$ by the

sum $\varepsilon(t)=\sum_{j=1}^{J}2^{-j+1}d_{jk}{}^{2}$: if $f(t)$ is regular with an indicator $\alpha<1$ at point $\boldsymbol{t}$, then $\varepsilon(t)=+\infty$; if $f(t)$ is regular with an indicator $\alpha\geq 1$ at point $\boldsymbol{t}$, then $\varepsilon(t)<+\infty$. Thus, the points at which $f(t)$ is regular with an indicator $\alpha<1$, are defined, calculating $\varepsilon(t)$ and comparing the threshold value set priorly. The function $\varepsilon(t)$, as an entropic characteristic, combines the results of the processing of the image at different scales. The choice of the discrete stationary wavelet transform (DSWT) is determined by its invariance with respect to its translation and to the keeping the size of the input image after its transformation.

The main stages of the contour segmentation method, described in Polyakova and Lyubchenko´s article [7], are:

- Choosing of upper and lower threshold value, having in mind the selected parametres;
- DSWT transforming every row of the image up to a certain level J and at every point k of this row the sum $\varepsilon(k)=\sum_{j=1}^{J}d_{jk}{}^{2}2^{j-1}$ is counted which yields a matrix R_1, having the same size as the input image;
- DSWT transforming every column of the image up to a certain level J and at every point k of this column the sum $\varepsilon(k)=\sum_{j=1}^{J}d_{jk}{}^{2}2^{j-1}$ is calculated which gives a matrix R_2, possessing the same size as the input image;
- Calculating the matrix $R=R_1+R_2$;
- Finding the local maxima of the gradient of the approximate matrix R for each of the four directions (*Fig. 1*). For example, the point (x,y) is taken to be a local maximum one in direction 2, if $R(x,y)>R_A=\left(d\sqrt{R(x+1,y+1)}+(1-d)\sqrt{R(x,y+1)}\right)^2$ and $R(x,y)>R_B=\left(d\sqrt{R(x-1,y-1)}+(1-d)\sqrt{R(x,y-1)}\right)^2$, where $d=\sqrt{R_1(x,y)R_2^{-1}(x,y)}$ (*Fig. 2*);
- Entering the local maxima found that surpass the low threshold constant in a weak contour set;

- Entering the local maxima found that surpass the upper threshold constant in a strong contour set;
- Processing by means of the hysteresis method – the weak contours are added to the neighbouring strong ones.

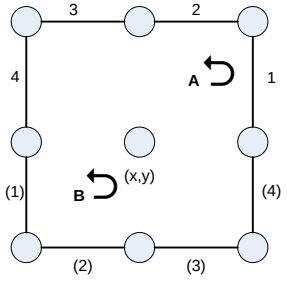


Figure 1. Neighbourhood of the point (x,y) in the image with marked directions ***A*** and ***B***

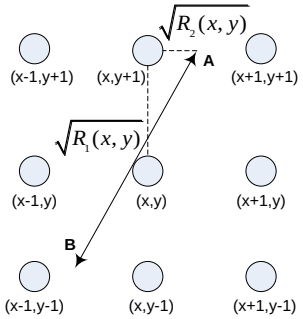


Figure 2. Finding the local maximum in direction 2

4. CONTOUR SEGMENTATION, USING HYPERBOLIC WAVELET TRANSFORM

A number of studies [9-13] shows the usefulness of Hilbert transformation (it can be interpreted as a convolution of the signal with the hyperbola $\frac{1}{\pi x}$) and the hyperbolic wavelet transformation, based on it for effective highlighting of the image contours. The Hilbert transformation is more resistant to noise, than the differentiation, when highligthing the contours. Besides, in the Hilbert transformation space the ideal and the extended image edge have more similar shapes than in the original one.

The hyperbolic wavelet transform combines the characteristics of both the Hilbert transformation and wavelet transform. It is defined by a convolution of the analysed signal $f(t)$ and the double-parameter basic function $\psi_{s,b}(t)$:

$$HWT(s,b) = \frac{1}{\sqrt{s}} \int_{-\infty}^{+\infty} f(t)\, \psi^* \left(\frac{t-b}{s} \right) dt.$$

The basic function is obtained by the mother wavelet $\psi_0(t)$ as a result of scaling with a parameter $s \in R^+$ and translating with a parameter $b \in R$.

In Ref. [14] Klich , Antoshchuk and Nikolenko obtained adaptive wavelet functions, based on the mother wavelet $\psi_0(t) = \frac{1}{\pi \alpha t} G(t)$, where $G(t)$ is an adapting

function which satisfies the conditions: $G(-t)=G(t), G(t)=1(t-\varepsilon)-1(t-\gamma)$, where $1(t)$ is the Heaviside step function, and $\alpha>0$ is a scaling factor, ε and γ are certain positive parameters – (cf. *Fig.3*).

The discrete hyperbolic wavelet transformation (*HWT*) of the line $\{f_{n'}\}$ is expressed by the formula: $HWT(s,n)=\sum_{n'=0}^{N-1} f_{n'}\psi^*\left(\frac{(n'-n)\Delta t}{s}\right)$, where Δt is the discretization step.

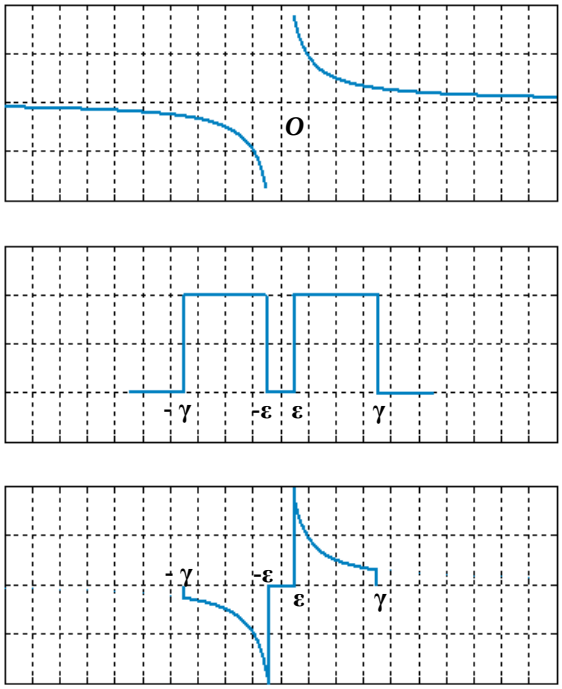


Figure 3. Design on the basic function

In Ref. [11] Antoshchuk, Babilunga and Nikolenko focus on a differential contour segmentation method, using a hyperbolic wavelet transform twice. During the first stage contours are highlighted with the help of *HWT* and during the second – the highlighted image is analysed using *HWT* in terms of local extrema, in order to detect the contour. The discrete variant of the algorithm for finding local extrema is $x[k]=x[k-1]-\beta.HWT(x[k-1])$ and the problem is reduced to finding such a $x[k]$, for which $HWT(x[k-1])$ is almost equal to zero. This ensures the estimate $|x[k]-x[k-1]|\le\delta$, where δ is a given accuracy for finding the extremum. The coefficient β controls the similarity speed and the stability of the iteration process. Hence, the contour indication is the reversal of the sign of the double hyperbolic wavelet transform. The method algorithm is shown in *Fig. 4*.

The *HWT* scale and the minimal threshold constant of the image edge to be found are given in the prior information block. When changing the sign of the double-hyperbolic wavelet transform, a zero is input in the binary image. The efficiency and the advantages of the method are tested on $64\times64-$ sized images, whose brightness varies within the $0-127$ range and the signal/noise ratio is from $0-100$.

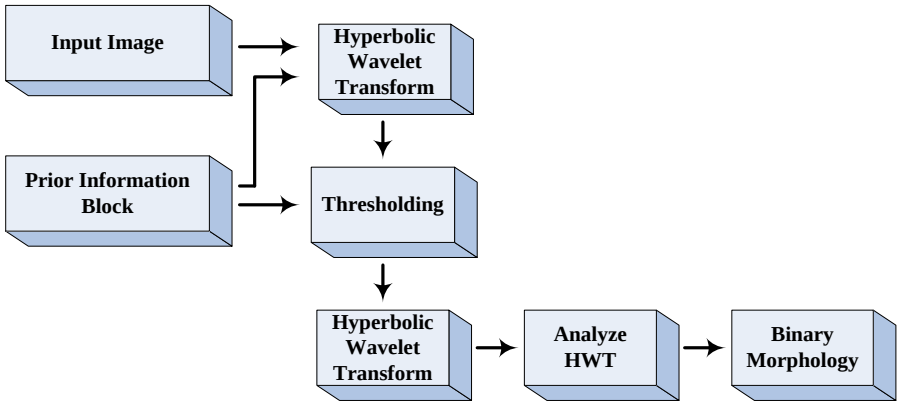


Figure 4. Scheme of the double Hyperbolic Wavelet Transform

5. CONCLUSION

The abovesaid outlines the trend for the development of the contour segmentation methods in order to detect a more and more accurate localization of the singularities of the signals analysed. This is confirmed in Ref. [15] in which Polyakova and Krylov have presented a morphological method of contour image segmentation based on the so-called *repagulum wavelet transform*. It uses a family of functions localized at a single point and for this reason allows a much more accurate detection of the place of the object contour in the image. Unlike the standard wavelet transform, instead a scaling parameter, one that characterizes the function regularity is used. Consequently, the repagulum wavelet transform is a convolution of the analysed signal with functions of different regularity. Combining this transform with the mathematics morphology eables the authors to build a contour image segmentation method that guarantees a great accuracy in the contour detection and an adjustable object detailization in the image.

REFERENCES

[1] U. Prett. "Digital image processing". Mir, Moscow, 1982, (Russian translation).
[2] V. Soyfer. "Methods of computer processing of images". Fizmatgiz, Moscow, 2003, (in Russian).

[3] J. Canny. "A computational approach to edge detection". IEEE Trans. Patt. Anal. and Mach. Intell, vol.36, pp.961-1005, Sept.1986.
[4] S. Mallat and S. Zhong. "Characterization of signals from multiscale edges". IEEE Trans. Patt. Anal. And Mach. Intell., vol.14(7), pp.710-732, July 1992.
[5] S. Mallat. " A wavelet tour of signal processing". Moscow, 2005, (Russian translation).
[6] S. Mallat. "Wavelets for Vision". Proceedings of the IEEE, vol.84(4), April 1996.
[7] M. Polyakova and V. Lyubchenko. "Contour segmentation of images, using wavelet-analysis". Proceedings of the Odessa Polytechnical University,vol.1(21) pp. 121-125, 2004, (in Russian).
[8] G. Nason and B. Silverman, "The stationary wavelet transform and some statistical applications. Lecture notes in statistics: wavelets and statistics". New York: Springer-Verlag, pp.281-299, 1995.
[9] S. Antoshchuk and V. Krylov. " Processing of images in the domain of the hyperbolic wavelet transform". AAEKS, Modeling of objects and systems for control, vol. 2(12). 2003, (in Russian).
[10] S. Antoshchuk and V. Krylov. "Hyperbolic wavelet domain image processing". Proceedings, Int. Sci. Conf., Lvov, Ukraine, 2004. Publishing House of the National University "Lvov Polytechnica", pp. 219-220, 2004, (in Russian).
[11] S. Antoshchuk, O. Babilunga and A. Nikolenko. " Separation of contours of objects, using the method of double-hyperbolic wavelet transform". Electromachine building and electrocontrol, vol. 65, pp. 65-69, 2005, (in Russian).
[12] S. Antoshchuk, O. Babilunga and A. Nikolenko. "Method of the morphological processing of semi-tone images on the basis of the hyperbolic wavelet transform". Electromachine building and electrocontrol, vol. 67, pp. 119-123, 2006, (in Russian).
[13] V.Krylov, M. Polyakova and N. Valkova. "Contour segmentation in the domain of hyperbolic wavelet transform, using mathematical morphology". AAEKS Information-measurement systems, vol.2(18). 2006, (in Russian).
[14] Yu. Klich , S. Antoshchuk and Nikolenko. Adaptive basis functions for wavelet transform. Proceedings of the Odessa Polytechnical University, vol.2(22) pp. 121-125, 2004, (in Russian).
[15] M. Polyakova and. V.Krylov. "The morphological method of contour image segmentation based on repagular wavelet transform". Proceedings of the Odessa Polytechnical University, vol.1(25), pp. 98-103, 2006, (in Russian).

Acknowledgments

This paper is supported by the Operational Program “Human Resources Development” 2007 – 2013, financed by the European Union (EU) through European Social Fund (ESF) under grant BG 051 PO 001-3.3.04/13