TRƯỜNG ĐẠI HỌC THỦ ĐÔ HÀ NỘI
Hanoi Metropolitan university

Tạp chí

KHOA HỌC

SCIENTIFIC JOURNAL OF HANOI METROPOLITAN UNIVERSITY

ISSN 2354-1504

Số 39 – KHOA HỌC TỰ NHIÊN VÀ CÔNG NGHỆ

Tháng 4 – 2020

MỤC LỤC

HIGGS SECTOR IN THE 3-3-1 MODEL WITH AXION DARK MATTER

Dang Van Soa, Hoang Ngoc Long, Pham Ngoc Thu, Vu Hoa Binh

Department of Physics, Hanoi Metropolitan University, Institute of Physics, VAST, Faculty of Natural Science and Technology, Tay Bac University, Department of Physics, Hanoi University of Education

Abstract: *The Higgs sector in the 3-3-1 model with axion dark matter is presented. The diagonal- ization of 4 × 4 square mass matrix for the CP-odd sector is exactly fulfilled. Our results show that the axion is mainly contained from the CP-odd part of the singlet φ, while the CP-even component of the later is the inflaton of the model. The positivity of the masses leads to constraints for some couplings of the Higgs sector. PACS numbers: 11.30.Fs, 12.15.Ff, 12.60.-i*

Keywords: *Higgs, axion, dark matter.*

Received: 10 March 2020

Accepted for publication: 20 April 2020

Email: soadangvan@gmail.com

1. INTRODUCTION

Being as Dark Matter (DM) candidate, nowadays, axion is very attracted subject in Particle Physics. The axion is a hypothetical CP-odd scalar, protected by a shift symmetry and derivatively coupled to Standard Model (SM) fields. It is predicted by the Peccei-Quinn solution to the strong CP problem [1, 2].

Among the beyond SM, the models based on $SU(3)_C \times SU(3)_L \times U(1)_X$ (3-3-1) gauge group [3] have some intriguing properties. It is emphasized that Peccei-Quinn symmetry is automatically satisfied in the 3-3-1 models [4]. That is why the 3-3-1 models is attractive for the axion puzzles.

In the framework of the 3-3-1 models, the axion has been studied in the papers [5–8]. In [5, 7], the axion is massless field and its mass is generated by quantum gravity effects. In addition, in diagonalization of square mass matrix for CP-odd scalars, the mixing matrix is not unitary leading to extra states such as PS_1 and PS_2. In this paper, these eros are corrected.

2. CONTENT

2.1. Brief review of the model

As usual, fermion content satisfying all the requirements is

$$\begin{gathered}\psi_{aL} = (\nu_a, e_a, \mathrm{N}_a)_L^T \sim (1,3,-1/3), \quad e_{aR} \sim (1,1,-1), \\ Q_{3L} = (u_3, d_3, U)_L^T \sim (3,3,1/3), \quad Q_{\alpha L} = (d_\alpha, -u_\alpha, D_\alpha)_L^T \sim (3,3^*,0), \\ u_{aR}, U_R \sim (3,1,1/3), \quad d_{aR}, D_{\alpha R} \sim (3,1,-1/3),\end{gathered} \tag{1}$$

where α = 2, 3 and a = 1, 2, 3 are family indices. The quantum numbers as given in parentheses are respectively based on $SU(3)_C$, $SU(3)_L$, U $(1)_X$ symmetries. The U and D are exotic quarks, while N_R are right-handed neutrinos. The above model is named by the 3-3-1 model with right-handed neutrinos.

The model with right-handed neutrinos requires three triplets:

$$\chi \sim (1,3,-1/3)\,, \eta \sim (1,3,-1/3)\,, \rho \sim\sim (1,3,2/3), \tag{2}$$

with expansions as follows

$$\chi = \begin{pmatrix} \frac{1}{\sqrt{2}}(R_\chi^1 + iI_\chi^1) \\ \chi^- \\ \frac{1}{\sqrt{2}}(\upsilon_\chi + R_\chi^3 + iI_\chi^3) \end{pmatrix},$$

$$\eta = \begin{pmatrix} \frac{1}{\sqrt{2}}(\upsilon_\eta + R_\eta^1 + iI_\eta^1) \\ \eta^- \\ \frac{1}{\sqrt{2}}(R_\eta^3 + iI_\eta^3) \end{pmatrix}, \rho = \frac{1}{\sqrt{2}} \begin{pmatrix} \rho_1^+ \\ \frac{1}{\sqrt{2}}(\upsilon_\rho + R_\rho + iI_\rho) \\ \rho_3^+ \end{pmatrix}. \tag{3}$$

In addition, ones introduce a singlet

$$\phi = \frac{1}{\sqrt{2}}\left(\upsilon_\phi + R_\phi + iI_\phi\right) \sim (1,1,0).$$

The full potential invariant under 3−3−1 gauge and $Z_{11} \otimes Z_2$ discrete symmetries is determined as [5]

$$\begin{aligned} V = {} & \mu_\phi^2 \phi^* \phi + \mu_\chi^2 \chi^\dagger \chi + \mu_\rho^2 \rho^\dagger \rho + \mu_\eta^2 \eta^\dagger \eta + \lambda_1 (\chi^\dagger \chi)^2 + \lambda_2 (\eta^\dagger \eta)^2 \\ & + \lambda_3 (\rho^\dagger \rho)^2 + \lambda_4 (\chi^\dagger \chi)(\eta^\dagger \eta) + \lambda_5 (\chi^\dagger \chi)(\rho^\dagger \rho) + \lambda_6 (\eta^\dagger \eta)(\rho^\dagger \rho) \\ & + \lambda_7 (\chi^\dagger \eta)(\eta^\dagger \chi) + \lambda_8 (\chi^\dagger \rho)(\rho^\dagger \chi) + \lambda_9 (\eta^\dagger)(\rho^\dagger \rho). \end{aligned} \tag{4}$$

Substitution of (3) into (4) leads to the following constraints at the tree level as follows

$$\begin{aligned}
\mu_\rho^2 + \lambda_3 v_\rho^2 + \frac{\lambda_5}{2} v_\chi^2 + \frac{\lambda_6}{2} v_\eta^2 + \frac{\lambda_6}{2} v_\phi^2 + \frac{L}{v_\rho^2} = 0, \\
\mu_\eta^2 + \lambda_2 v_\eta^2 + \frac{\lambda_4}{2} v_\chi^2 + \frac{\lambda_6}{2} v_\rho^2 + \frac{\lambda_{12}}{2} v_\phi^2 + \frac{L}{v_\eta^2} = 0, \\
\mu_\chi^2 + \lambda_1 v_\chi^2 + \frac{\lambda_4}{2} v_\eta^2 + \frac{\lambda_5}{2} v_\rho^2 + \frac{\lambda_{11}}{2} v_\phi^2 + \frac{L}{v_\chi^2} = 0, \\
\mu_\phi^2 + \lambda_{10} v_\phi^2 + \frac{\lambda_{11}}{2} v_\chi^2 + \frac{\lambda_{12}}{2} v_\rho^2 + \frac{\lambda_{13}}{2} v_\eta^2 + \frac{L}{v_\phi^2} = 0,
\end{aligned} \tag{5}$$

where $L \equiv \lambda_\Phi v_\Phi v_\chi v_\eta v_\square$.

2.2. Charged scalar sector

In this sector we have two square mass matrices. One of them is as follows: In the base (η_1^-, δ_1^-) ones get square mass matrix as

$$M_c = \begin{pmatrix} \frac{\lambda_9 v_\rho^2}{2} - \frac{L}{2v_\eta^2} & \frac{\lambda_9 v_\rho v_\eta}{2} - \frac{L}{2v_\rho v_\eta} \\ \frac{\lambda_9 v_\rho v_\eta}{2} - \frac{L}{2v_\rho v_\eta} & \frac{\lambda_9 v_\eta^2}{2} - \frac{L}{2v_\rho^2} \end{pmatrix} = -\frac{\left(L - \lambda_9 v_\rho^2 v_\eta^2\right)}{2} \begin{pmatrix} \frac{1}{v_\eta^2} & \frac{1}{v_\eta v_\rho} \\ \frac{1}{v_\eta v_\rho} & \frac{1}{v_\rho^2} \end{pmatrix}. \tag{6}$$

This matrix has one massless G_1^- and one massive H_1^- with mass equal to

$$m_{H_1^-}^2 = -\frac{\left(L - \lambda_9 v_\rho^2 v_\eta^2\right)}{2} \cdot \frac{\left(v_\rho^2 + v_\eta^2\right)}{v_\rho^2 v_\eta^2}. \tag{7}$$

From (7) it follows

$$\lambda_9 > \lambda_\phi \frac{v_\phi v_\chi}{v_\eta v_\eta}. \tag{8}$$

The physical fields are given as

$$\begin{pmatrix} G_1^- \\ H_1^- \end{pmatrix} = \begin{pmatrix} \cos\theta_\alpha & \sin\theta_\alpha \\ \sin\theta_\alpha & \cos\theta_\alpha \end{pmatrix} \begin{pmatrix} \rho_1^- \\ \eta^- \end{pmatrix}, \tag{9}$$

where

$$\tan\theta_\alpha = \frac{v_\eta}{v_\rho} \tag{10}$$

In the limit $v_\rho \gg v_\eta$ we have

$$G_1^- = \rho_1^- \simeq G_{w^-}. \tag{11}$$

In the base (χ^-, ρ_3^-) ones get square mass matrix as

$$M_{c2} = \begin{pmatrix} \frac{\lambda_8 v_\rho^2}{2} - \frac{L}{2v_\chi^2} & \frac{\lambda_8 v_\rho v_\chi}{2} - \frac{L}{2v_\rho v_\eta} \\ \frac{\lambda_8 v_\rho v_\chi}{2} - \frac{L}{2v_\rho v_\chi} & \frac{\lambda_8 v_\chi^2}{2} - \frac{L}{2v_\rho^2} \end{pmatrix} = -\frac{\left(L - \lambda_8 v_\rho^2 v_\chi^2\right)}{2} \begin{pmatrix} \frac{1}{v_\chi^2} & \frac{1}{v_\chi v_\rho} \\ \frac{1}{v_\chi v_\rho} & \frac{1}{v_\rho^2} \end{pmatrix}. \quad (12)$$

This matrix has one massless G_2^- = and one massive H_2^- with mass equal to

$$m_{H_2^-}^2 = -\frac{\left(L - \lambda_8 v_\rho^2 v_\chi^2\right)}{2} . \frac{\left(v_\rho^2 + v_\chi^2\right)}{v_\rho^2 v_\chi^2} \quad (13)$$

From (13) it follows

$$\lambda_8 > \lambda_\phi \frac{v_\phi v_\eta}{v_\chi v_\rho}. \quad (14)$$

The physical fields are given as

$$\begin{pmatrix} G_2^- \\ H_2^- \end{pmatrix} = \begin{pmatrix} \cos\theta_\beta & \sin\theta_\beta \\ \sin\theta_\beta & \cos\theta_\beta \end{pmatrix} \begin{pmatrix} \chi^- \\ \rho_3^- \end{pmatrix}, \quad (15)$$

where

$$\tan\theta_\beta = \frac{v_\rho}{v_\chi}. \quad (16)$$

In the limit $v_\rho \gg v_\eta$ we have

$$G_2^- = X_1^- \simeq G_y^-. \quad (17)$$

2.3. CP-ODD sector

For CP-odd scalars, in the base (I_1^x, I_1^3) ones get square mass matrix as

$$M_A\left(I_\chi^1, I_\eta^3\right) = \begin{pmatrix} \frac{\lambda_7}{4} v_\eta^2 - \frac{L}{2v_\chi^2} & -\frac{\lambda_7}{4} v_\chi v_\eta + \frac{L}{2v_\chi v_\eta} \\ -\frac{\lambda_7}{4} v_\chi v_\eta + \frac{L}{2v_\chi v_\eta} & \frac{\lambda_7}{4} v_\chi^2 - \frac{L}{2v_\eta^2} \end{pmatrix}. \quad (18)$$

Diagonalization of matrix in (18) yields one massless scalar G_1 and one massive field A_1 with mass as follows

$$m_{A_1}^2 = -\frac{\left(L - \lambda_7 v_\eta^2 v_\chi^2\right)}{2} . \frac{\left(v_\eta^2 + v_\chi^2\right)}{v_\eta^2 v_\chi^2}. \quad (19)$$

From (19) it follows

$$\lambda_7 > \lambda_\phi \frac{v_\phi v_\rho}{v_\chi v_\eta}. \tag{20}$$

The physical fields are

$$\begin{pmatrix} G_1 \\ A_1 \end{pmatrix} = \begin{pmatrix} \sin\beta & \cos\beta \\ \cos\beta & -\sin\beta \end{pmatrix} \begin{pmatrix} I_x^1 \\ I_\eta^3 \end{pmatrix} \tag{21}$$

.

Next, in the base ($I_\chi{}^3$, $I_\eta{}^1$, I_ρ, I_φ) we have square mass matrix

$$M_{4odd} = \frac{L}{2} \begin{pmatrix} \frac{1}{u_c^2} & \frac{1}{u_c u_h} & \frac{1}{u_c u_r} & \frac{1}{u_c u_f} \\ & \frac{1}{u_h^2} & \frac{1}{u_r u_h} & \frac{1}{u_f u_h} \\ & & \frac{1}{u_r^2} & \frac{1}{u_r u_f} \\ & & & \frac{1}{u_f^2} \end{pmatrix}. \tag{22}$$

Let us diagonalize the matrix in (22). For this aim, we denote

$$N = \frac{1}{v_\chi}, \quad B = \frac{1}{v_\eta}, \quad C = \frac{1}{v_\rho}, \quad D = \frac{1}{v_\phi}. \tag{23}$$

The the matrix in (22) is rewritten as

$$M_{4odd} = \frac{L}{2} \begin{pmatrix} N^2 & NB & NC & ND \\ NB & B^2 & BC & BD \\ NC & BC & C^2 & CD \\ ND & BD & CD & D^2 \end{pmatrix}. \tag{24}$$

The above matrix has three massless states and one massive with the following eigenvectors

$$U = \begin{pmatrix} -\frac{D}{N}I_c^3 & -\frac{C}{N}I_h^1 & -\frac{B}{N}I_r & \frac{C}{N}I_f \\ 0 & 0 & I_r & \frac{B}{D}I_f \\ 0 & I_h^1 & 0 & \frac{C}{D}I_f \\ I_c^3 & & & I_f \end{pmatrix}. \quad (25)$$

Taking eigenvector in the first column of (25) and write rotation matrix

$$C_{43} = \begin{pmatrix} \frac{D}{C_2} & 0 & 0 & -\frac{N}{C_2} \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ \frac{N}{C_2} & 0 & 0 & \frac{D}{C_2} \end{pmatrix}, \quad (26)$$

where we denote

$$C_2 = \sqrt{N^2 + D^2} \Rightarrow \frac{D}{C_2} = \sin\theta_1 \frac{N}{C_2} = \cos\theta_1, \tan\theta_1 = \frac{D}{N} = \frac{v_\chi}{v_\phi}. \quad (27)$$

For the limit $v_\varphi \gg v_\chi$ we have

$$C_2 \approx 1/v_x$$

We can check out that

$$M_{4diag} = C43 \times M_{4odd} \times C43^\dagger = \frac{L}{2}\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & B^2 & BC & BC_2 \\ 0 & BC & C^2 & CC_2 \\ 0 & BC_2 & CC_2 & C_2^2 \end{pmatrix}. \quad (28)$$

Here we get one massless state a which is identified to axion and one massive state A_2

$$\begin{aligned} a &= \cos\theta_1 I_\phi - \sin\theta_1 I_\chi^3, \\ A_2 &= \sin\theta_1 I_\phi + \cos\theta_1 I_\chi^3. \end{aligned} \quad (29)$$

From (29), it follows that in the limit $v_\varphi \gg v_\chi$

$$a = I_\varphi. \quad (30)$$

Summarising the first step

$$\begin{pmatrix} I_c^3 \\ I_h^1 \\ I_r \\ I_f \end{pmatrix} = C43 \begin{pmatrix} a \\ I_\eta^1 \\ I_\rho \\ A_2 \end{pmatrix}, \tag{31}$$

In the basis $a,\ I_\eta^1, A_2$, we have square mass matrix given in (28). The 3×3 matrix in right-bottom here has two massless states and one massive as follows

$$\left\{\left\{-\frac{C_2}{B}, 0, 1\right\}, \left\{-\frac{C}{B}, 1, 0\right\}, \left\{\frac{B}{C_2}, \frac{C}{C_2}, 1\right\}\right\} \tag{32}$$

Using the second solution, we have rotation matrix

$$C32 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{C}{C_3} & -\frac{B}{C_3} & 0 \\ 0 & \frac{B}{C_3} & \frac{C}{C_3} & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{33}$$

where

$$C_3 = \sqrt{C^2 + B^2} \sim \mathcal{O}\left(\frac{1}{v_\eta}, \frac{1}{v_\rho}\right)$$

. Therefore:

$$\tan\theta_2 = \frac{C}{B} \sim \frac{v_\eta}{v_\rho} \Rightarrow \sin\theta_2 = \frac{C_3}{B}, \cos\theta_2 = \frac{C_3}{C}. \tag{34}$$

Then

$$M_{3diag} = C32 \times M_{4diag} \times C32^\dagger = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & C_3^2 & C_2C_3 \\ 0 & 0 & C_2C_3 & C_2^2 \end{pmatrix} \tag{35}$$

Here we have one massless G_2 and one massive fields A_3

$$G_2 = \sin_{\theta_2 I_\rho} - \cos_{\theta_2 I_\eta^1,}$$
$$A_3 = cos_{\theta_2 I_\rho} + \sin_{\theta_2 I_\eta^1}. \quad (36)$$

Here, G_2 is Goldstone boson for the Z boson. We have

$$\begin{pmatrix} I_c^3 \\ I_h^1 \\ I_r \\ I_f \end{pmatrix} = C43.C32 \begin{pmatrix} a \\ G_Z \\ A_3 \\ A_2 \end{pmatrix}, \quad (37)$$

The 2 × 2 matrix in right-bottom of (35) is easily diagonalized. Let define

$$\tan\theta_3 = \frac{C_2}{C_3} \sim \frac{\upsilon_\eta}{\upsilon_\chi}. \quad (38)$$

Then we have one massless G_3 which is identified as $G_{Z'}$ and one massive A_4 fields

$$G_3 = \sin\theta_3 A_3 - \cos\theta_3 A_2, \quad (39)$$
$$A_4 = \cos\theta_3 A_3 + \sin\theta_2 A_2, \quad (40)$$

where mass of A_4 is given as

$$m_{A_4}^2 = L \times \frac{C_3^2}{cos^2\theta_3}. \quad (41)$$

Let us write

$$C21 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \cos\theta_3 & -\sin\theta_3 \\ 0 & 0 & \sin\theta_3 & \cos\theta_3 \end{pmatrix} \quad (42)$$

then

$$\begin{pmatrix} \mathrm{I}_c^3 \\ \mathrm{I}_h^1 \\ \mathrm{I}_r \\ \mathrm{I}_f \end{pmatrix} = C43.C32.C21 = \begin{pmatrix} \sin\theta_1 & 0 & -\cos\theta_1\cos\theta_3 & -\cos\theta_1\sin\theta_3 \\ 0 & \sin\theta_2 & -\cos\theta_2\sin\theta_3 & \cos\theta_2\cos\theta_3 \\ 0 & \cos\theta_2 & \sin\theta_2\sin\theta_3 & -\cos\theta_3\sin\theta_2 \\ \cos\theta_1 & 0 & \cos\theta_3\sin\theta_2 & \cos\theta_1\sin\theta_3 \end{pmatrix} \begin{pmatrix} a \\ \mathrm{G}_{\mathrm{Z}} \\ \mathrm{G}_{\mathrm{Z}'} \\ \mathrm{A}_4 \end{pmatrix}. \quad (43)$$

For practical analysis

$$\begin{pmatrix} a \\ \mathrm{G}_{\mathrm{Z}} \\ \mathrm{G}_{\mathrm{Z}'} \\ \mathrm{A}_4 \end{pmatrix} = \begin{pmatrix} \sin q_1 & 0 & 0 & \cos q_1 \\ 0 & \sin q_2 & \cos q_2 & 0 \\ -\cos q_1\cos q_3 & -\cos q_2\sin q_3 & \sin q_2\sin q_3 & \cos q_3\sin q_2 \\ -\cos q_1\sin q_3 & \cos q_2\cos q_3 & -\cos q_3\sin q_2 & \cos q_1\sin q_3 \end{pmatrix} \begin{pmatrix} \mathrm{I}_c^3 \\ \mathrm{I}_h^1 \\ \mathrm{I}_r \\ \mathrm{I}_f \end{pmatrix}. \quad (44)$$

Note that here we do not have massive states PS_1 and PS_2 as in Ref. [?].

From (44), it follows that in the limit $v_\varphi \gg v_\chi \gg v_\rho \gg v_\eta$

$$\begin{aligned} a &\backsimeq I_\phi, \\ G_{\mathrm{Z}} &\backsimeq I_\rho, \\ G_{\mathrm{Z}'} &\backsimeq I_\chi^3, \\ A_4 &\backsimeq I_\eta^1 . \end{aligned} \quad (45)$$

Substituting related values into (41) yields

$$m_{A_4}^2 = L\left(\frac{1}{v_\phi^2}+\frac{1}{v_\chi^2}+\frac{1}{v_\rho^2}+\frac{1}{v_\eta^2}\right) \quad (46)$$

$$\approx \lambda_\phi v_\phi v_\chi \left(\frac{v_\rho}{v_\eta}\right). \quad (47)$$

From (46) it follows

$$\lambda_\varphi > 0. \quad (48)$$

Hence, if $\lambda_\varphi \sim O(1)$ then A_4 is very heavy with mass in the range of v_φ.

Summary: In the CP-odd sector we have 6 fields: two Goldstone bosons for Z and Z′, one axion a, one massless field G_1 and two massive pseudoscalars A_1 and A_4.

2.4. CP-EVEN sector

There are six CP-even scalars and they separate into two square mass matrices.

Within the constraint conditions in (5), ones get square mass matrix of CP-even scalars written in the basis of (R_χ^1, R_η^3) as

$$M_R\left(R_\chi^1, R_\eta^3\right) = \begin{pmatrix} \frac{\lambda_7}{4}v_\eta^2 - \frac{L}{2v_\chi^2} & \frac{\lambda_7}{4}v_\chi v_\eta - \frac{L}{2v_\chi v_\eta} \\ \frac{\lambda_7}{4}v_\chi v_\eta - \frac{L}{2v_\chi v_\eta} & \frac{\lambda_7}{4}v_\chi^2 - \frac{L}{2v_\eta^2} \end{pmatrix} \quad (49)$$

Diagonalization of matrix in (5) yields one massless scalar G_4 and one massive field H_1 with masse as follows

$$m_{H_1}^2 = -\frac{\left(L - \lambda_7 v_\eta^2 v_\chi^2\right)}{2} \cdot \frac{\left(v_\eta^2 + v_\chi^2\right)}{v_\eta^2 v_\chi^2}. \quad (50)$$

The physical fields are

$$\begin{pmatrix} G_4 \\ H_1 \end{pmatrix} = \begin{pmatrix} -\sin\beta & \cos\beta \\ \cos\beta & \sin\beta \end{pmatrix} \begin{pmatrix} R_\chi^1 \\ R_\eta^3 \end{pmatrix}. \quad (51)$$

In the limit $v_\eta \ll v_\chi$, we have $R_\eta{}^3 = G_4, R_\chi{}^1 = H_1, G_1 = I_\eta{}^3, I_\eta{}^1 = A_1$, hence

$$\eta_3^0 \equiv G_{X^0}. \quad (52)$$

Here

$$G_{X^0} = \frac{1}{\sqrt{2}}(G_4 + iG_1) = \frac{1}{\sqrt{2}}\left(R_\eta^3 + iI_\eta^3\right),$$

is the Goldstone boson for the X^0 boson.

Looking at Eqs (41) and (50) we realize that A_1 and H_1 have the same mass and they are component of χ_1^0. Hence we can compose them to new massive complex scalar ϕ^0

$$\varphi^0 = \frac{1}{\sqrt{2}}\left(R_\chi^1 + iI_\chi^1\right),$$

with mass given in (50).

In the limit $v_\phi \gg v_\chi \gg v_\rho \gg v_\eta$, one has

$$\chi \simeq \frac{1}{\sqrt{2}}\begin{pmatrix} \varphi^0 \\ G_{Y^-} \\ \frac{1}{\sqrt{2}}(v_\chi + R_\chi^3 + iG_{Z'}) \end{pmatrix}, \eta \simeq \begin{pmatrix} \frac{1}{\sqrt{2}}(u + R_\eta^1 + iA_4) \\ H_1^- \\ G_{X^0} \end{pmatrix}, \rho \simeq \frac{1}{\sqrt{2}}\begin{pmatrix} G_{W^+} \\ \frac{1}{\sqrt{2}}(v + R_\rho + iG_Z) \\ H_2^+ \end{pmatrix}. \quad (53)$$

Thus, at this step we have already determined Goldstone bosons for Z, Z′ and neutral bilepton X^0 and one massive complex scalar φ^0.

Next let us consider the second part of CP-even scalars. In the basis $(R_\chi^3, R_\eta^1, R_\rho, R_\phi)$, one has

$$\begin{pmatrix} 2\lambda_1 v_\chi^2 - \frac{L}{2v_\chi^2} & \frac{\lambda_4 v_\chi v_\eta}{2} + \frac{L}{2v_\eta v_\chi} & \frac{\lambda_5 v_\chi v_\rho}{2} + \frac{L}{2v_\rho v_\chi} & \frac{L}{2v_\phi v_\chi} \\ \frac{\lambda_4 v_\chi v_\eta}{2} + \frac{L}{2v_\eta v_\chi} & 2\lambda_2 v_\eta^2 - \frac{L}{2v_\eta^2} & \frac{\lambda_6 v_\eta v_\rho}{2} + \frac{L}{2v_\rho v_\eta} & \frac{L}{2v_\eta v_\phi} \\ \frac{\lambda_5 v_\chi v_\rho}{2} + \frac{L}{2v_\rho v_\chi} & \frac{\lambda_6 v_\eta v_\rho}{2} + \frac{L}{2v_\rho v_\eta} & 2\lambda_3 v_\rho^2 - \frac{L}{2v_\rho^2} & \frac{L}{2v_\rho v_\phi} \\ \frac{L}{2v_\phi v_\chi} & \frac{L}{2v_\eta v_\phi} & \frac{L}{2v_\rho v_\phi} & 2\lambda_{10} v_\phi^2 - \frac{L}{2v_\phi^2} \end{pmatrix} \quad (54)$$

In the case of $v_\phi \gg v_\rho, v_\eta, v_\chi$, we have that R_φ decouples and its mass is predicted to be

$$m_{R_\phi}^2 \simeq 2\lambda_{10} v_\phi^2. \quad (55)$$

From (55) it follows

$$\lambda_{10} > 0. \quad (56)$$

For the future studies, we will identify R_ϕ to inflaton.

Keeping the next term of order $v_\varphi v_\chi$ yields

$$M_{4R} = \begin{pmatrix} 2\lambda_1 u_c^2 & 0 & 0 & 0 \\ 0 & -\frac{L}{2u_h^2} & \frac{L}{2u_r u_h} & 0 \\ 0 & \frac{L}{2u_r u_h} & -\frac{L}{2u_r^2} & 0 \\ 0 & 0 & 0 & 2\lambda_{10} u_f^2 \end{pmatrix} \quad (57)$$

Hence at this step one has one massive state R_χ^3 with mass

$$m^2_{R^3_\chi} \simeq \lambda_1 v^2_\chi. \tag{58}$$

From (58) it follows

$$\lambda_1 > 0. \tag{59}$$

We will identify $R^3_\chi \equiv H_5$. This is heavy scalar.

It is easily diagonalize matrix in (57) and there are two solutions: one massless G_5 and one massive H_2

$$\begin{aligned} G_5 &= -\cos\theta_4 R_\eta - \sin\theta_4 R_\rho, \\ H_2 &= +\sin\theta_4 R_\eta - \cos\theta_4 R_\rho, \end{aligned} \tag{60}$$

where $\tan\theta_4 = v_\eta v_\rho$ and the H_2 mass is given by

$$m^2_{H_2} = \frac{L v^2_\rho}{2v^2_\eta cos^2\theta_4} = \frac{L(v^2_\rho + v^2_\eta)}{2v^2_\eta} \tag{61}$$

To avoid massless state G_5, let us diagonalize 2 × 2 matrix in central part of (54), e.g.,

$$M_{R2} = \begin{pmatrix} 2\lambda_2 v^2_\eta - \dfrac{L}{2v^2_\eta} & \dfrac{\lambda_6 v_\eta v_\rho}{2} + \dfrac{L}{2v_\rho v_\eta} \\ \dfrac{\lambda_6 v_\eta v_\rho}{2} + \dfrac{L}{2v_\rho v_\eta} & 2\lambda_3 v^2_\rho - \dfrac{L}{2v^2_\rho} \end{pmatrix} \equiv -\begin{pmatrix} \text{b} & \text{-d} \\ \text{-d} & \text{c} \end{pmatrix} \tag{62}$$

□

Then we have two massive states H_3 and H_4

$$\begin{aligned} H_3 &= -\sin\theta_5\, R^1_\eta + \cos\theta_5\, R_\rho, \\ H_4 &= -\cos\theta_5\, R^1_\eta - \sin\theta_5\, R_\rho, \end{aligned} \tag{63}$$

where

$$\tan 2\theta_5 = \frac{2d}{c-b}. \tag{64}$$

with masses given by

$$2m^2_{H_3} = (b+c) - \sqrt{(c-b)^2 + 4d^2}, \tag{65}$$

$$2m^2_{H_4} = (b+c) - \sqrt{(c-b)^2 + 4d^2}. \tag{66}$$

We can identify H_3 as the SM-like Higgs boson h.

Let us consider the limit $v_\varphi \gg v_\chi \gg v_\rho \gg v_\eta$, then

$$b = -2\lambda_2 v_\eta^2 + \frac{L}{2v_\eta^2} \approx \lambda_\phi v_\phi v_\chi \frac{v_\rho}{v_\eta},$$
$$c = -2\lambda_3 v_\rho^2 + \frac{L}{2v_\rho^2} \approx -2\lambda_3 v_\rho^2 + \lambda_\phi v_\phi v_\chi \frac{v_\rho}{v_\eta}, \quad (67)$$
$$d = \frac{\lambda_6 v_\eta v_\rho}{2} + \frac{L}{2v_\rho v_\eta} \approx \frac{\lambda_6 v_\eta v_\rho}{2} + \frac{1}{2}\lambda_\phi v_\phi v_\chi.$$

Hence

$$c - b \approx -2\lambda_3 v_\rho^2 - \lambda_\phi v_\phi v_\chi \frac{v_\rho}{v_\eta},$$
$$(c + b) \approx -2\lambda_3 v_\rho^2 + \lambda_\phi v_\phi v_\chi \frac{v_\rho}{v_\eta}, \quad (68)$$
$$\tan 2\theta_5 = \frac{2d}{c-b} \approx \frac{v_\eta}{v_\rho} \Rightarrow \tan\theta_5 \approx \frac{v_\eta}{2v_\rho}.$$
$$\Delta = \left(2\lambda_3 v_\rho^2 + \lambda_\phi v_\phi v_\chi \frac{v_\rho}{v_\eta}\right)^2 + 4\left(\frac{\lambda_6 v_\eta v_\rho}{2} + \frac{1}{2}\lambda_\phi v_\phi v_\chi\right)^2$$
$$\approx \left(\lambda_\phi v_\phi v_\chi \frac{v_\rho}{v_\eta}\right)^2 \Rightarrow \sqrt{\Delta} \approx \lambda_\phi v_\phi v_\chi \frac{v_\rho}{v_\eta} \quad (69)$$

Substituting (69) into (65) and (66) yields

$$m_{H_3}^2 = -\lambda_3 v_\rho^2, \quad (70)$$
$$m_{H_4}^2 = \lambda_\phi v_\phi v_\chi \frac{v_\rho}{v_\eta} - \lambda_3 v_\rho^2. \quad (71)$$

From (70) it follows

$$\lambda_3 < 0\,. \quad (72)$$

From (70) it follows that $\lambda_3 < 0$ and H_3 can be identified to SM Higgs boson h, while H_4 is heavy scalar and $\lambda_\varphi > 0$. Note that

$$h \approx R_\rho, \qquad H_4 \approx -R_\eta^1. \quad (73)$$

Hence

$$\chi \simeq \begin{pmatrix} \varphi^0 \\ G_{Y^-} \\ \frac{1}{\sqrt{2}}(v_\chi + H_5 + iG_{Z'}) \end{pmatrix}, \eta \simeq \begin{pmatrix} \frac{1}{\sqrt{2}}(u - H_4 + iA_4 \\ H_1^- \\ G_{\chi^0} \end{pmatrix}, \rho \simeq \begin{pmatrix} G_{W^+} \\ \frac{1}{\sqrt{2}}(v_\chi + H_5 + iG_{Z'} \\ H_2^- \end{pmatrix}, \quad (74)$$
$$\phi \simeq \frac{1}{\sqrt{2}}(v_\phi + \Phi + ia).$$

Note that inflaton Φ has mass in the range of 10^{11} GeV, the SM Higgs with mass ~ 126 GeV, and other fields A_4, H_4, H_5 and φ^0 with masses in TeV scale.

3. CONCLUSION

In this paper we have studied the scalar sector of the 3-3-1 model containing axion as DM candidate. The diagonalization of 4 × 4 square mass matrix for the CP-odd sector is exactly fulfilled. Our results show that the axion is mainly contained from the CP-odd part of the singlet φ, while the CP-even component of the later is the inflaton of the model. The positivity of the masses leads to constraints for some couplings of the Higgs sector: only λ_3 is negative, while coupling constants λ_φ and λ_i,i = 1,7,8,9,10 are positive.

Note that the axion in the model under consideration is still massless. Our next task is to make it massive by introducing soft breaking mass term.

Acknowledgments

This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 103.01-2017.356.

REFERENCES

1. R. D. Peccei and H. Quinn (1977), Phys. Rev. Lett.38, 1440.
2. R. D. Peccei and H.Quinn (1977), Phys. Rev. D 16, 1791.
3. J. W. F. Valle and M. Singer (1983), Phys. Rev. D 28, 540; F. Pisano and V. Pleitez, Phys. Rev. D 46, 410 (1992) [hep-ph/9206242]; R. Foot, O. F. Hernandez, F. Pisano and V. Pleitez, Phys. Rev. D 47, 4158 (1993) [hep-ph/9207264]; P. H. Frampton, Phys. Rev. Lett. 69, 2889 (1992). R. Foot, H. N. Long and T. A. Tran, Phys. Rev. D 50, no. 1, R34 (1994) [hep-ph/9402243]; H. N. Long, Phys.
4. Rev. D 54, 4691 (1996) [hep-ph/9607439]; H. N. Long, Phys. Rev. D 53, 437 (1996) [hep-ph/9504274].
5. P. B. Pal (1995), Phys. Rev. D 52 1659 [hep-ph/9411406].
6. A. G. Dias, C. A. de S. Pires, P. S. Rodrigues da Silva (2003), Phys. Rev. D 68 115009, arXiv:hep-ph/0309058
7. A. G. Dias, V. Pleitez (2004), Phys. Rev. D 69,077702.
8. J. G. Ferreira Jr., C. A. de S. Pires, J. G. Rodrigues, P. S. Rodrigues da Silva (2017), Phys. Lett. B 771, 199.
9. J. C . Montero, A. R. R. Castellanos, and B. L. Sanchez- Vega (2018), Phys. Rev. D 97,063015

BỘ THẾ HIGGS TRONG MÔ HÌNH 3-3-1 VẬT CHẤT TỐI AXION

Tóm tắt: *Bộ thế Higgs trong mô hình 3-3-1 với vật chất axion được mô tả. Ma trận khối lượng 4x4 cho phần lẻ CP được chéo hóa chính xác. Kết quả chỉ ra rằng axion xuất hiện ở phần lẻ CP trong khi phần chẵn CP chứa lạm phát. Tính dương của khối lượng dẫn tới chúng ta giới hạn được một số hằng số tương tác trong bộ thế Higgs.*

PACS: 11.30.Fs, 12.15.Ff, 12.60.-i

Từ khoá: *Bộ thế Higgs, axion, vật chất tối.*

INFLUENCE OF THE GRAPHENE SIZE ON THE ABSORPTION COEFFICIENT OF WEAK ELECTROMAGNETIC WAVE IN TWO-DIMENSIONAL GRAPHENE

Nguyen Vu Nhan, Tran Anh Tuan

Hanoi Metropolitan University, Hanoi University of Science

Abstract: *On the basis of the quantum kinetic equation has received analysing expressions for current density and absorption coefficient of electromagnetic wave in two-dimenssional graphene (2D graphene) with electron- optical phonon scattering is dominant. The dependence of the absorption coefficient into the parameters characteristic for the external wave field and the size parameters of graphene is very complex and nonlinear. The results are calculated, graphed proved well for theoretical results. The results are comparable to the case in the normal semiconductors indicating the difference, and results are new.*

Keywords: *Absorption coefficient, Quantum kinetic equation, 2D Graphene.*

Received: 15 March 2020

Accepted for publication: 20 April 2020

Email: nvnhan@daihocthudo.edu.vn

I. INTRODUCTION

Two-dimensional (2D) graphene was first studied by the scientists Andrei Geim and Konstantin Sergeevich Novoselov on the transport of electrons and holes. They were awarded the 2010 Nobel Prize in Physics [1]. Graphene is a flat sheet about the thickness of an atomic layer of carbon atoms, which forms a honeycomb shaped array. The length of carbon-carbon bond (C-C) in graphene is about 0.142 nm. 2D graphene is important in making graphene transistors in electronic chips and can replace existing silicon transistors.

Graphene has the ability to fill its electron holes with other electrons almost instantly, meaning it transfers a large amount of charge in an extremely short amount of time. Graphene is a separate carbon atom sheet bundled into a two-dimensional (2D) honeycomb network, which is the basis of graphite-like materials of different dimensions. It can be encapsulated into non-dimensional fluleren (0D); rolled into one-dimensional nanocarbon tubes (1D); or folded into 3D graphite (three dimensions) [2]. Put simply, graphene is a sheet

of graphite separated at atomic size.

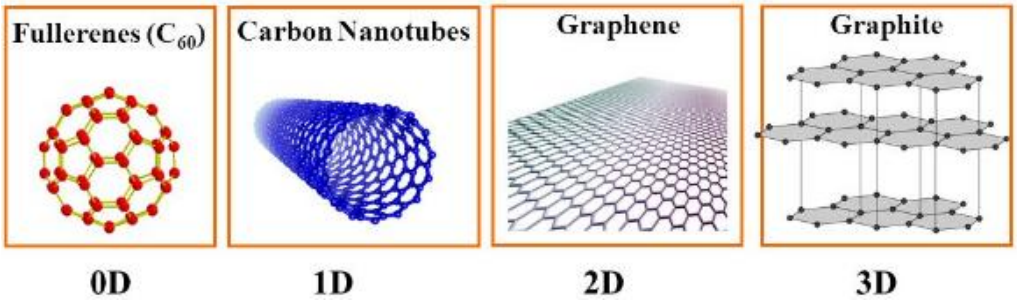
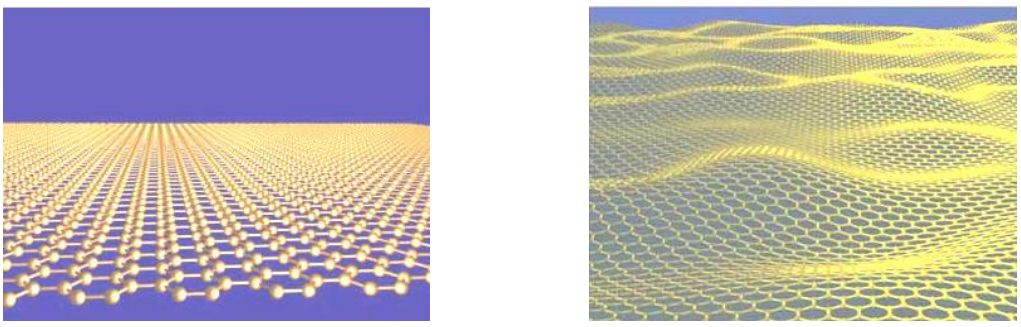


***Fig.1**: Low-dimensional graphene: zero-Dimensional (0D), one-dimensional (1D), bidirectional (2D) and three-dimensional (3D)*

Twenty years ago, the term "nanotechnology" was little known, but now it has become a familiar term at every stage of modern society. Nanometer is the abbreviation for nanometer (symbol nm, 1 nm = 10^{-9}m) is a unit of measurement in the atomic system. Nanotechnology involves taking advantage of new physical phenomena (quantum-stimulating effects) in nanometer-sized materials to create materials with special functions.

Although the theory does not accept an all-crystal lattice in 2-dimensional space, that is, on an absolute plane, but it does not prohibit a 2-dimensional lattice that relies on a 3-dimensional cube. This is true for the observations of Geim, Novoselov et al. in [3]. Under the microscope they were able to observe suspended graphene plaques in a state of non-flat free space, which were as protruding as the microscopic wavefront in three-dimensional space (Fig. 2).



(a) (b)

***Fig. 2:** Graphene does not exist in an absolute plane (a),but existing with the convex surface of the 3D space (b).*

Graphene is considered to be a super-thin, electrically conductive, heat-resistant material that is 100 times stronger than steel. With such superior structure, the study of properties of graphene systems attracts the attention of scientists. Many physical phenomena

are of interest and research in low dimensional physics due to the quantum size effect of materials. The effect of electromagnetic absorption is one of those research directions. This effect has been studied relatively sufficiently in normal semiconductor (3D) materials as well as other low-dimensional semiconductors (0D, 1D, 2D) both theoretically and empirically (for example, in [4]). In the movement of free particles in low-dimensional materials, the interaction between the carrier and the phonon is the most significant contribution. This means that when the electron moves in the crystal lattice of the materials, it will be affected by electromagnetic waves in the direction of acceleration, and also affected by the oscillation of the crystal lattice in the direction of motion impedance. With the assumption that the strongly variable electric field is a flat electromagnetic wave and that the propagation of this electromagnetic wave along the Oz axis of the material has a decreasing electromagnetic wave intensity. The characteristic quantity for reducing the intensity of the electromagnetic wave when going deep into the semiconductor is called the electromagnetic absorption coefficient. Investigating the strong electromagnetic absorption coefficient caused by the electron-phonon interaction is a classic but very important problem because this is the effect that occurs in semiconductor materials present in electronic accessories. The theory of electromagnetic nonlinear absorption in cubic semiconductors has been studied and published by V. Pavlovich and E. M. Epshtein since 1977 [5] on the basis of the method of quantum dynamic equations for electronics. With graphene materials, in the world in recent years, there are also a number of groups of authors studying the magnetic - phonon resonance effect [6] with the interaction process between electron - phonon in graphene. Good results and consistent with experiment. However, the coefficient of electromagnetic wave absorption in Graphene 2D has not been studied. Therefore, we are interested in studying this physics problem. In [7], we formulated quantum theory for the absorption of weak electromagnetic wave in 2D graphene and received analytical expression for the absorption coefficient. Numerical calculations show the nonlinear dependence of the absorption coefficient on the frequency of electromagnetic wave, wave intensity and temperature of the system. In this paper, we continue to study the effect of graphene size on absorption coefficient. Results are calculated, graphed and discussed.

2. CONTENT

2.1. The 2d graphene lattice structure

Graphene is a lattice layer of two-dimensional carbon atoms separated from graphite. Graphene has a hexagonal structure like a honeycomb (Fig. 3). This special nano structure has excellent properties such as good electrical conductivity due to high carrier mobility, very mechanical and thermal stability. Therefore, each carbon atom is linked to the three nearest carbon atoms by the bonds created by the overlap of s-p orbitals, corresponding to the sp2 hybridization state.

The $2p_z$ trajectory perpendicular to the graphene sheet will not participate in the hybridization process but will overlap and form liên bonds, which are not localized, will

form conductive regions π and create irregular electrical properties. graphene (Fig. 4 and Fig. 5). Although graphene has a high symmetry in the structure, the hexagonal cell in graphene was not chosen as the unit cell because the adjacent carbon atoms are not equivalent. However, we can consider graphene lattice as a combination of sub-lattices including carbon atoms at A and carbon at B, so the surrounding carbon atoms are completely equivalent each in terms of structure and properties. The graphene lattice structure can be described by the unit vectors of these subnets. The elemental vectors $\vec{a}_1, \vec{a}_2$ are valued as [8]:

$\vec{a}_1 = \frac{a}{2}(3, \sqrt{3})$; $\vec{a}_2 = \frac{a}{2}(3, \sqrt{3})$; $a_{\min} = 0{,}142\, nm$, with a is the lattice constant.

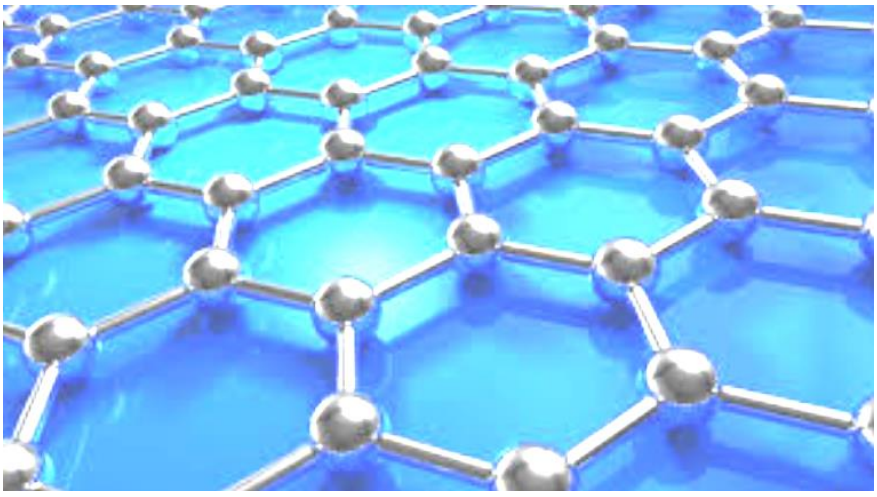


Fig. 3: *Graphene crystal structure* (*hexagonal structure, also known as honeycomb structure*)

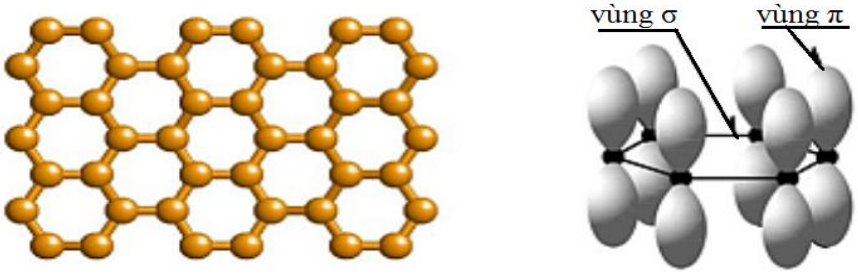


Fig. 4: *Bonds of C atoms in graphene lattice and illustrate the regions σ and the region π*

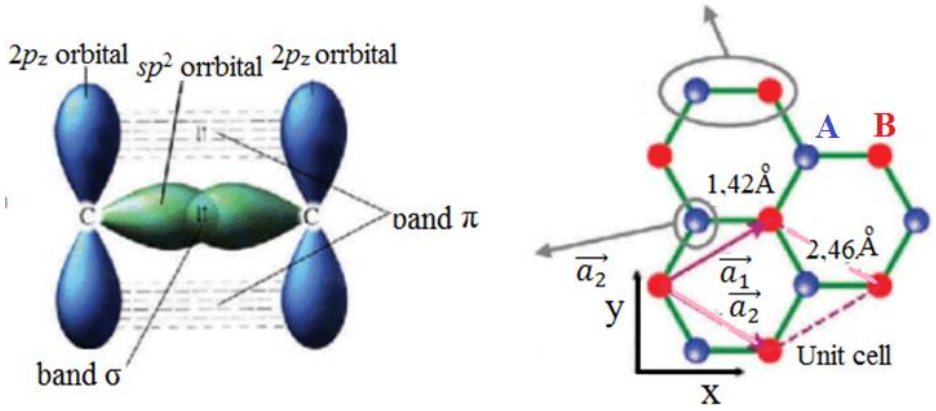


Fig. 5: *2D graphene lattice structure*

2.2. The carrier current density and absorption coefficient of a weak electromagnetic wave in 2d graphene

A high-frequency electromagnetic wave is applied to the system in the z direction with electric field vector $\vec{E} = \vec{E}_0 \sin \Omega t$ with $\vec{E}_0$ and Ω are the amplitude and the frequency of the electromagnetic wave, respectively. Assume 2D graphene lattice is in the oxy plane in the magnetic field $\vec{B} = (0, 0, B)$, then the wave function and energy spectrum the electron in graphene [7] have the following form (1) and (2):

$$\vec{F}_{n,X}(\vec{r}) = \frac{C_n}{\sqrt{L}} \exp\left(-i\frac{Xy}{l^2}\right) \begin{bmatrix} sng(n) h_{|n|-1}(x - X) \\ \\ h_{|n|-1}(x - X) \end{bmatrix} \quad (1)$$

$$\varepsilon_n = S_n \hbar \omega_B |n|^{1/2} \ , \quad (2)$$

where:

$$C_n = \begin{cases} 1 & n = 0 \\ \dfrac{1}{\sqrt{2}} & n \neq 0 \end{cases} \ ; \ sng(n) \equiv S_n = \begin{cases} 1 & n > 0 \\ 0 & n = 0 \\ -1 & n < 0 \end{cases} \ ; \ h_{|n|-1}(x)$$

$$= \frac{i^{|n|}}{\sqrt{2^{|n|}|n|!\sqrt{\pi} l}} \exp\left[-\frac{1}{2}\left(\frac{x}{l}\right)^2\right] H_{|n|}\left(\frac{x}{l}\right) \ ;$$

L - the linear dimension of the system; X - a center coordinate, $H_n(t)$ – the Hermite polynomial; $l = (c\hbar / eB)^{1/2}$, $n = 0, \pm 1 \ldots$; $\hbar\omega_B = \sqrt{2}\gamma / l$ is the effective magnetic energy.

Procced from the hamiltonian of the electron - optical phonon systein 2D Graphene in the second quantization presentation can be written as [9] :

$$H = \sum_{n,\vec{k}_\perp} \varepsilon_n \left[\vec{k}_\perp - \frac{e}{\hbar c}\vec{A}(t)\right] a^+_{n,\vec{k}_\perp} a_{n,\vec{k}_\perp} + \sum_{\vec{q}} \hbar\omega_{\vec{q}} \left(b^+_{\vec{q}} b_{\vec{q}} + \frac{1}{2}\right) + \sum_{n,n',\vec{q},\vec{k}_\perp} M_{n,n'}(\vec{q})\, a^+_{n',\vec{k}_\perp+\vec{q}} a_{n,\vec{k}_\perp} \left(b^+_{-\vec{q}} + b_{\vec{q}}\right) \quad (3)$$

where: $\vec{A}(t)$- the vector potential of an external electromagnetic wave; *n* denotes the quantization of the energy spectrum in the z direction (n = 1, 2, 3, ...); $\vec{k}, \vec{q}$ respectively are wave vectors of electron, phonon, $(n, \vec{k}_\perp), (n', \vec{k}_\perp + \vec{q}_\perp)$are electron states before and after scattering, respectively, $\vec{k}_\perp$ is in plane (x,y) wave vector of the electron. $a^+_{n,\vec{k}_\perp}, a_{n,\vec{k}_\perp}, b^+_{\vec{q}}, b_{\vec{q}}$ are the creation and annihilation operators of electron,phonon, respectively, $\vec{q} = (\vec{q}_\perp, q_z)$; $M_{n,n'}(\vec{q})$ is the matrix factor of electron in the formula

$$\left|M_{n,n'}(\vec{q})\right|^2 = \hbar D^2_{op}(2\rho L^2\omega_{\vec{q}})^{-1} C^2_n C^2_{n'} \frac{m!}{(m+j)!} e^{-u} u^j \left[L^j_m(u) + S_n S_{n'} \sqrt{\frac{m+j}{m}} L^j_{m-1}(u)\right], \quad (4)$$

with: ρ is mass density of 2D Graphene, D_{op} is deformed potential of optical phonon; $L^m_j(u)$ is the associated Laguerre polynomial, $u = \ell^2 q^2/2$, $q^2 = q_x^2 + q_y^2, m = \min(|n|, |n'|), j = \big||n| - |n'|\big|$.We calculated and received quantum dynamic equation for electronics in the 2D graphene (5) and the nonlinear absorption coefficient (6) in 2D graphene [7], written as follows:

- *The carrier current density expression in 2D graphene*:

$$\vec{J}_\perp(t) = -\frac{e}{m^* c}\sum_{n,\vec{k}_\perp} \vec{A}(t) n_{n,\vec{k}_\perp}(t) + \sum_{\ell=1}^{\infty} \sin(\ell\Omega t) \frac{2\pi e\hbar^2}{m^* l\Omega} \sum_{n,n',\vec{k}_\perp,\vec{q}} \left|M_{n,n'}(\vec{q})\right|^2 \sum_{k=-\infty}^{\infty} \vec{q}_\perp J_k\left(\frac{e\vec{q}_\perp \vec{E}_o}{m^*\Omega^2}\right) \times \left[J_{k+l}\left(\frac{e\vec{q}_\perp \vec{E}_o}{m^*\Omega^2}\right) + J_{k-l}\left(\frac{e\vec{q}_\perp \vec{E}_o}{m^*\Omega^2}\right)\right] N_{\vec{q}} \left(\overline{n_{n,\vec{k}_\perp}} - \overline{n_{n',\vec{k}_\perp+\vec{q}}}\right) \delta\left(\varepsilon_{n',\vec{k}_\perp+\vec{q}} - \varepsilon_{n,\vec{k}_\perp} + \hbar\omega_o - k\hbar\Omega\right) \quad (5)$$

where: $\overline{n_{n,\vec{k}_\perp}} \equiv f_{n,\vec{k}_\perp}$ is the time - independent component of the electron distribution function; $\overline{N}_q$ is the time-independent component of the phonon distribution function; $\delta(x)$

is the Dirac delta function; $J_k(x)$ is the Bessel function [9] and the quantity δ is infinitesimal and appears due to the assumption of an adiabatic interaction of the electromagnetic wave. Note that, the motion of electrons is confined along the z direction in a 2D Graphene, so we only consider the in plane (x, y) current density vector of electrons $\vec{J}_{\perp}(t)$.

- *The nonlinear absorption coefficient of the electromagnetic wave in 2D graphene:*

$$\alpha = \alpha_0 \frac{L_x L_y}{l^6 \Omega^3}\left[\exp\left(\frac{\hbar\omega_0}{k_B T}-1\right)\right]^{-1}\sum_{n,n',N} f_{n,N} C_n^2 C_{n'}^2 \times\left[(2m+j+1)-2S_n S_{n'}\sqrt{m(m+j)}+S_n^2 S_{n'}^2(2m+j-1)\right]\Gamma_o(\varepsilon^2+\Gamma_o^2)^{-1} \quad (6)$$

with:

$$\alpha_0 = 2\pi\hbar e^2 D_{op}^2 S_o c^{-1} L^{-2}(\chi_\infty \rho\omega_0 m^{*2})^{-1/2} \quad ; \quad (7)$$

$$\Gamma_o = \gamma\ \hbar\omega_R D_{op}(8\hbar\pi\rho\omega_o)^{-1/2} \quad ; \quad (8)$$

$$f_{n,N} = \frac{n_o^2}{\pi^2}\left[\hbar\Omega_B(\mathrm{N}+\tfrac{1}{2})+\frac{\pi^2 n^2\hbar^2}{2m^* L^2}\right]^{-2}\left(1-\exp\left[-\frac{m^*\Omega}{\hbar\Omega_B(\mathrm{N}+\frac{1}{2})+\frac{\pi^2 n^2\hbar^2}{2m^* L^2}}\right]\right)^2 \quad ; \quad (9)$$

$$\varepsilon = \varepsilon_n - \varepsilon_{n'} + \hbar\omega_0 - \hbar\Omega \ , \quad (10)$$

where: η_0 is the equilibrium distribution function of the electron [10]; S_o is unit acreage of 2D graphene lattice χ_∞ is the high-frequency dielectric constants; $N = 0, \pm 1, \ldots$; $m = \min(n, n')$; $j = \big||n| - |n'|\big|$; γ is the conductivity coefficient.

2.3. Numerical results and discussion

In [7], the nonlinear absorption coefficient α is calculated and plotted when considered as a function depending on the parameters: temperature T, frequency of electromagnetic wave Ω, magnetic field intensity B. Below, we continue to calculate the number and plot the dependence of the nonlinear absorption coefficient on the size of the 2D graphene sheet. The parameters used in computational calculations are as follows:

$$\rho = 7,7.10^{-8} g.cm^{-2};\ D_{op} = 1,4.10^{-9} eV.cm^{-1};\ k_B = 1,3807.10^{-23} J.K^{-1};$$
$$\hbar\omega_o = 6$$
$$\hbar = 1,05459.10^{-34} J.s;\ \chi_\infty = 10,9;\ \gamma = 6,46\ eV.\overset{0}{A}.$$

The graph in Fig.6 shows clearly the dependence of the nonlinear absorption coefficient on the size of the Graphene sheet. When the graphene sheet has a longer length, the absorption of electric waves in the graphene will be stronger. This dependency is the basis for implementing standards for two-dimensional graphene fabrication technology. That gives us more complete technology of making Graphene sheet to apply it in high-end, super small, versatile and intelligent electronic components.

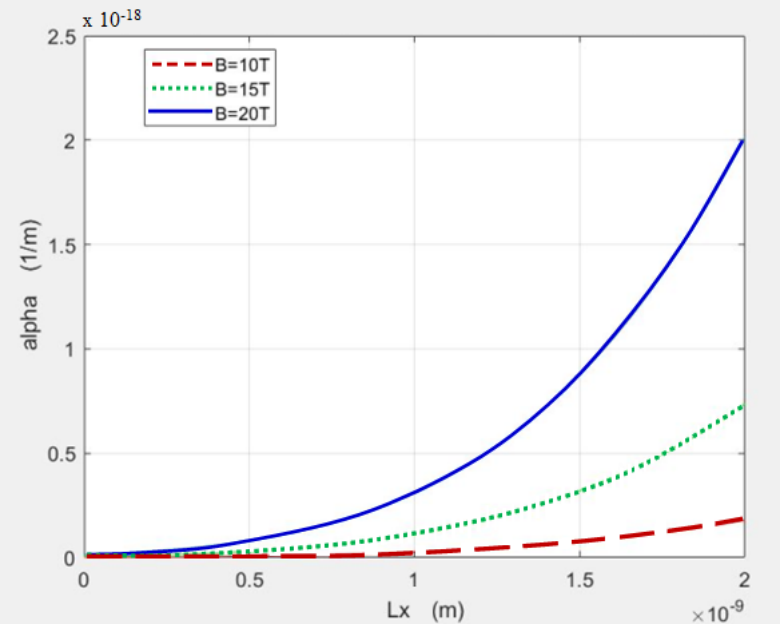


Fig. 6. *The dependence of coefficient α (alpha) on graphene lattice thickness*

The results obtained above are the basis for new research directions in theory and experiment in 2D graphene. The physical problem could expand for electron- acoustic phonon scattering, or the presence of electromagnetic waves have a strong variable amplitude.

3. CONCLUSION

In this paper, on the basic of quantum kinetic equation, We obtained the analytic expressions of current density and nonlinear absorption coefficient of a electromagnetic wave in 2D graphene for the presence of an external magnetic field. The electromagnetic wave is assumed high-frequency with weak amplitude. Results are numerically calculated, graphed to clarify the dependence of absorption coefficient on the size of 2D graphene. In [7], we have shown the magnetic resonance region. At each of the different values of the magnetic field, the absorption coefficient reaches the maximum value with different amplitudes of oscillation.

Note that, because amplitude of electromagnetic wave is weak should have influence of electromagnetic wave intensity on the absorption coefficient is negligible. So, in this paper we used the quadratic approximation of the Bessel function to simplify the physic problem.

Graphene interacts with electric waves from the infrared to the ultraviolet region. We know that silicon and other semiconductors also interact with electromagnetic waves but only in the infrared region. Graphene induce with electromagnetic waves stretches from the microwave zone (cm wavelength) to the ultraviolet region (nm wavelength). Therefore, the intensity of graphene's interaction with electromagnetic waves is superior to the conducting polymers and nanotubes. To date, no theory has ever been able to explain why photons can affect organic materials that conduct electricity in a wide range of waves from centimeter to nanometer [11]. Therefore, physicists and electronics engineers are still actively researching and searching for useful applications of graphene in optoelectronics [12].

In particular, the intensity of the interaction of graphene is very strong with terahertz waves (Terahertz waves are located between microwaves and infrared waves). Terahertz waves can see through fabric, plastic but are absorbed by metals and inorganic compounds, so this is the wave used to detect weapons, explosives hidden in people or in luggage. The interaction of graphene in region of terahertz wave shows the potential of using graphene in future anti-terrorist terahertz sensors.

Acknowledge

To complete this work, we would like to thank the theoretical research team of the Department of Theoretical Physics (VNU - University of Natural Sciences) and special thanks to the Basic Research Support Program for the topic C2019-44.

REFERENCES

1. K. Geim (2010), "*Random walk to graphene*" , Nobel Lecture

2. K. M. Borysenko, J. T. Mullen, E. A. Barry, S. Paul, Semenov, and K. W. Kim (2010), "First-principles analysis of electron-phonon interactions in graphene", *Phy. Rev B.*, Vol. 81, No. 12, pp. 3-6.

3. J. C. Meyer, A. K. Geim, M. I. Katsnelson, K. S. Novoselov, T. J. Booth and S. Roth (2007), Nature 446, p. 60.

4. N. Q. Bau, N. V. Nhan, and T. C. Phong (2002), "Calculation of the absorption coef- ficient of a weak electromagnetic wave by free carriers in doped superlattices by using the kubo - mori method", *J. Korean Phys.*, Vol. 41, p. 149, 2002.

5. E. M. Epstein (1975), Sov. Communicacattion of HEE of USSR, p. 785

6. Mori and T. Ando (2011), Magnetophonon resonance in monolayer graphene, *Journal of the Physical Society of Japan*, Vol. 80, No. 4, pp. 044 706.

7. Nguyen Vu Nhan, Tran Anh Tuan, Do Tuan Long, Nguyen Quang Bau (2019), "Quatum theory of absorption of electromagnetic wave in two-dimenssional graphene". Sci. Joun.HNMU, Science and Technology, No.31, pp.107-116.

8. B. Ha, N. V. Hieu, and N. Hoai (2005), "*Lectures on electron structures of graphene and graphene bilayer*," Advanced Center of Physics Vietnam Academy of Science and Technology, pp. 1–4.
9. T. Ando (2007), "Magnetic oscillation of optical phonon in graphene", *J. Phys. Soc. Jpn*, Vol. 76, No. 2, p. 024712.
10. A. K. Geim and K. S. Novoselov (2007), "The rise of graphene, Nature Materials", Vol. 6, No. 3, pp. 183–191.
11. V.T. Truong, P. J. McMahon and A. R. Wilson (2012), J. Polym. Sci. Part B: Polym. Phys., Vol.50, p. 624.
12. P. Avouris and F. Xia (2012), MRS Bulletin, Vol. 27, p.1225.

ẢNH HƯỞNG CỦA KÍCH THƯỚC GRAPHENE LÊN HỆ SỐ HẤP THỤ SÓNG ĐIỆN TỪ YẾU TRONG GRAPHENE HAI CHIỀU

Tóm tắt: *Trên cơ sở phương trình động học lượng tử đã nhận được biểu thức giải tích cho mật độ dòng điện và hệ số hấp thụ sóng điện từ trong graphene hai chiều (2D graphene) với giả thiết tán xạ electron - phonon quang là trội. Sự phụ thuộc của hệ số hấp thụ vào các tham số đặc trưng cho trường ngoài và các tham số kích thước của graphene là rất phức tạp và phi tuyến. Các kết quả được tính toán và vẽ đồ thị đã minh chứng tốt cho các kết quả lý thuyết. Các kết quả được so sánh với trường hợp trong chất bán dẫn khối cho thấy sự khác biệt và tính mới của kết quả.*

Từ khóa: *Hệ số hấp thụ, phương trình động lượng tử, 2D Graphene.*

ORIGINAL ARTICLE CHARACTERISTIC INVESTIGATIONS OF A COMMERCIAL CYLINDRICAL-TYPE LITHIUM-ION BATTERY

Nguyen Van Nghia, Nguyen Van Ky, Luong Trung Son, Nguyen Thi Thu Hoa, Hoang Manh Ha, Nguyen Si Hieu, Ta Anh Tan, Dang Tran Chien

Hanoi Architectural University, Hanoi Architectural University, Le Quy Don Technical University, Vietnam Academy of Science and Technology, Hanoi Metroplitan University, Hanoi University of Natural Resources and Environment

Abstract: *A commercial 4000 mAh cylindrical 26650-type lithium-ion battery was disassembled for the purpose of studying the electrode materials that defines the electrochemical performance of the battery. The cathode material is found to be a mixture of LiM_2O_4 and $LiMO_2$ oxides (M = Mn, Co, Ni, Al or its combinations) and the anode material is dominated by graphite. The cathode materials composes of the homogeneous particles with mean diameter of 1 to 3 µm meanwhile the mean diameter of the graphite particles is about 10 µm. The first discharge capacity of the battery is 3820 mAh, which accounted for approximately 95.5% of the nominated capacity. The discharge capacity is gradually decreases during cycling. The average discharge voltage plateau is approximately 3.7 V, which is identical with the rated working voltage battery.*

Keywords: *Lithium-ion battery, commercial battery, anode materials, cathode materials, lithium compounds.*

Received: 25 March 2020

Accepted for publication: 20 April 2020

Email: dtchien@hunre.edu.vn

1. INTRODUCTION

Lithium-ion batteries (LIBs) have attracted much attention in the past few decades. Due to high energy density, reduced pollution, stable performance and long-life cycle, the commercial lithium-ion batteries are widely used to supply power for electric vehicles (EVs), electronic devices and battery energy storage (BES) for renewable energy [1]. Lithium-ion batteries have been widely used in portable electronic products and transportation vehicles. The lithium-ion battery has multiple sub-types such as the aqueous lithium-ion battery, nickel manganese cobalt oxide battery, and lithium-ion phosphate battery [2].

According to a research from Frost & Sullivan, the global lithium-ion battery market is expected to be worth $22.5 billion in 2016. North America and China hold more than half the global revenues for LiBs. The global lithium-ion market size was around $35 billion to $40 billion as of 2018, and the demand will be incrementing at a positive Compound Annual Growth Rate (CAGR) of 13% to 18% during the forecast period of 2019 to 2025 [2]. In general, battery applications are predicted to have a turnover of $90 billion by 2025. Investments of five companies such as Siemens, Total, Daimler, Dyson, AshiKasei, Catl, Tesla/Panasonic are over 13.7 billion USD that have been spent or planned in 2016 and 2017. In there, Contemporary Amperex Technology Co Ltd. (CATL) accounted for $2.9 billion [3]. There are many lithium ion battery manufacturers in China. Those factories are separated in two kinds. One is the packaging factory which buys the battery cell and assembles the battery packs including mobile battery, laptop battery, storage battery, RC battery. Some of these kinds are Desay Battery Technology, Sunwoda Electronic, Eve Energy manufactures that do battery packaging for Huawei Technologies and Apple Inc.

According to a study of Jeff Desjiardins [4], global lithium-ion battery production capacity will increase by 521% between 2016 and 2020. The capacity in 2016 was 28 GWh, and it will reach 174 GWh in 2020. By 2020, mass production of lithium batteries will still be dominated by 4 countries including USA, Poland, Korea and China. The lithium-ion battery capacity of China will rise to 62% (108 GWh) by 2020.

Although lithium batteries were commercialized in 1991 by Sony [5,6], scientists have been constantly researching to improve the performance of batteries and reduce the product's costs. The studies were mainly focused on the materials used to fabricate the cathodes, anodes and electrolytes of the batteries. In addition, numerous investigations have been paid on the overcharge/over-discharge failure [1], cycle aging at different temperatures [7], potential dependence of gas evolution in 18650 cylindrical cells [8], the impedance characteristics at various conditions [9], working at very high temperature for capacity fading and understanding of aging mechanisms [10], analyzing the rechargeable LiBs after prolonged cycling [11], interfacing investigations of a commercial lithium ion battery graphite anode material [12], etc. In this work, a commercial lithium-ion battery was investigated.

2. CONTENT

2.1. Materials and Methods

A 26650 cylindrical Li-ion battery was dismantled in a box filled with argon to separate the cathode tape and anode tape. The anode and cathode materials were easily collected from the electrode tapes. The crystal structure of the electrode materials was identified by X-ray diffraction (XRD, D8 Advance Bruker, Cu Kα radiation, $\lambda = 1.5406$ Å). The XRD data were collected in the diffraction angle range (2θ) between 10º and 70º. The morphologies of the materials were observed using scanning electron microscopy (SEM, Jeol 6490 JED 2300,

Japan). The charge/discharge measurements were taken using an Auto-Lab Potentiostat PGS-30. All cells were cycled between 3.0 - 4.25 V at current density of 0.2 C.

2.2. Results and Discussion

A conventional LiB usually consists of two electrodes (a negative and a positive electrode), an electrolyte, and a porous polymer separator. The positive electrode was made of a transition metal oxide or a phosphate, while the negative was made of graphite. The electrolyte most often composed of an organic solvent containing a lithium salt. The porous polymer separator separated the two electrodes and prevents them from short circuit. The electrolyte easily penetrates into the porous polymer separator. During charging, the electrons are made to be released at the cathode and move externally to the anode. At the same time, the lithium ions move in the same direction, but internally, from cathode to anode via the electrolyte. When the battery is discharged, electrons travel from anode to the cathode through the external load. Concurrently, lithium ions travel from anode to the cathode in the electrolyte.

In order to study on the cathode and anode materials, it is necessary to disassemble the battery and analyze the relevant battery components. The cutting positions are about 2 mm near the positive or negative connector. The procedure of the disassemble process is illustrated in Fig. 1.

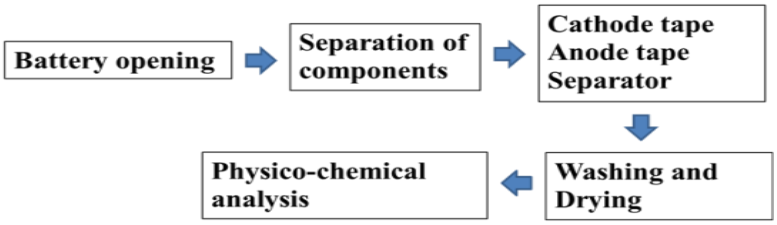


Fig.1. *Chart for disassembling of the Li-ion battery.*

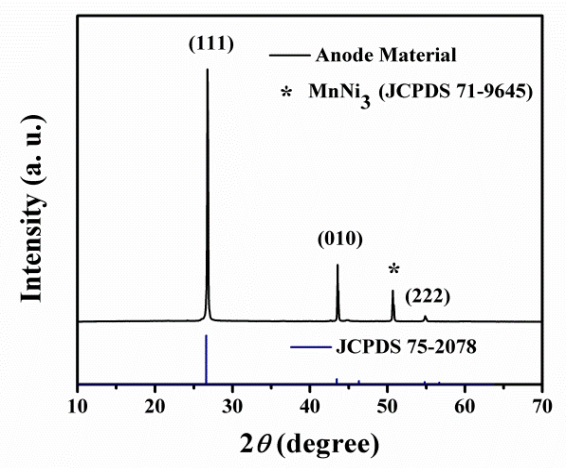


Fig. 2. *XRD pattern of the anode of the battery*

The XRD pattern of the anode material is shown in Fig. 2. Looking at the pattern, one can see that graphite dominated the field of anode materials of the battery. This phenomenon is due to graphite's competent capability of insertion and extraction of Li-ions. The diffraction peaks at the 2θ values of approximately 26.59° 43.58◦, 54.80◦ correspond to the (111), (010), (222) planes, respectively (JCPDS card No. 75-2078). One diffraction peak presented at the 2θ value of 50.70° is attributed to the $MnNi_3$ phase according to JCPDS card No. 71-9645. The $MnNi_3$ phase possibly formed due to the dissolve of Mn and Ni from the cathode material into the electrolyte then deposited onto the anode surface.

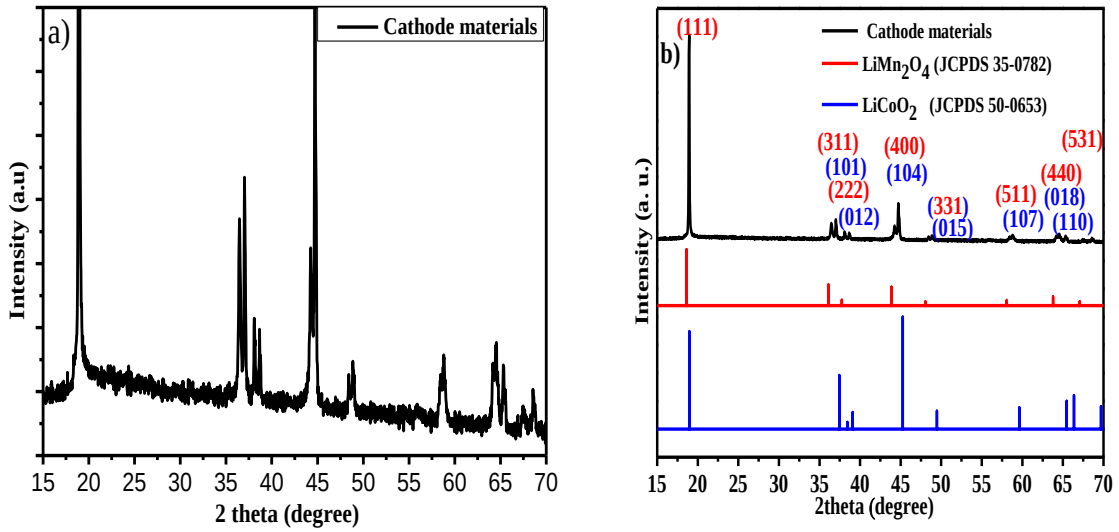


Fig. 3. *(a) XRD pattern of the materials on cathode of the battery; (b) XRD pattern of the materials in compare with JCPDS 35-0782 and JCPDS 50-0653.*

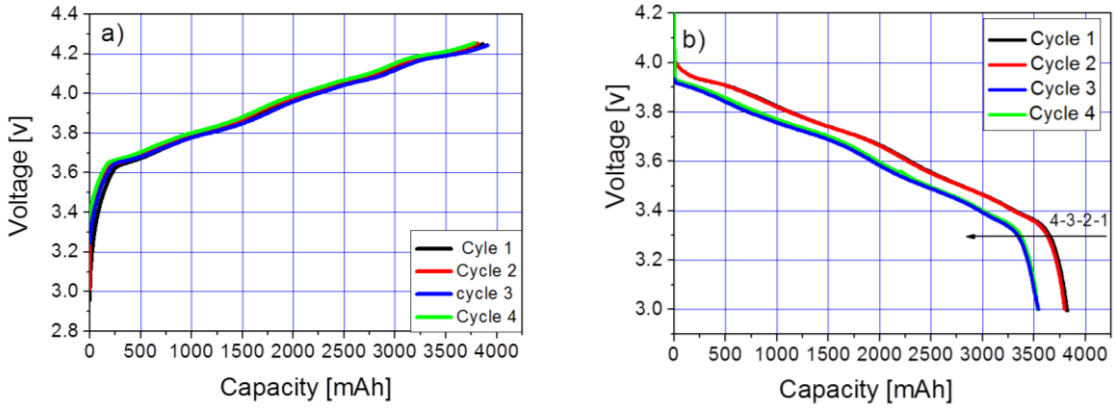


Fig. 4. *(a) The first to the fourth charge cycle profiles and (b) discharge cycle profiles.*

Figure 3a shows the XRD pattern of the materials on the battery cathode. To recognize what are the structures of the materials, the XRD pattern of the materials is put in the same graph with two JCPDS 35-0782 ($LiMn_2O_4$) and JCPDS 50-0653 ($LiCoO_2$) cards in Fig. 3b.

The results show that the cathode compounds have two structures, one is the spinel structure which is identical with the structure of the $LiMn_2O_4$ oxide and another one is P2 layer structure which is the same with the structure of the $LiCoO_2$ oxide.

Figure 4a and Fig. 4b show the battery voltage versus capacity during the charging and discharging processes at current density of 0.2 C between 3.0 - 4.25 V. The charge curves are similar and almost overlap which means the Li^+ ions can insert into the electrode materials reversibly (Fig. 4a). The discharge curves have the same shape which means the Li^+ can extract from the electrode reversibly (Fig. 4b). The average charge voltage plateau and discharge plateau are appropriate 4.0 V and 3.7 V, respectively. The difference between charge voltage plateau and discharge voltage plateau is about 0.25 V suggested that the de-lithiation and the lithiation processes have a slightly polarization.

The charge capacities, discharge capacities and the corresponding coulomb efficiencies, which are the ratios of discharge capacities and charge capacities, of the first to the fourth cycles are listed in Table 1. It can be seen that the discharge capacities are gradually decreased and the Coulombic efficiencies are over 90%. The decrease of the discharge capacity *is assigned to* the formation of the solid-electrolyte interface layer during the initial lithiation, which forms a barrier to the flow of lithium ions through the electrode [13]. This reduces the access to the active material.

Table 1. **The charge and discharge capacities and coulomb efficiencies of the first to the fourth cycles**

Cycle number	Charge capacity (mAh)	Discharge capacity (mAh)	Coulomb efficiency (%)
1	3863	3820	98.89
2	3809	3796	99.65
3	3911	3542	90.57
4	3778	3523	93.25

Table 2. **Chemical components (wt%) of the anode at different zones**

Zone	1	2	3
C	77.04	75.44	76.49
O	17.64	17.93	17.68
F	3.87	4.75	4.49
Si		0.17	
P	1.09	1.47	1.2
S	0.15	0.07	0.04
Ca			0.03
Fe	0.21	0.18	0.07

Total	100	100	100

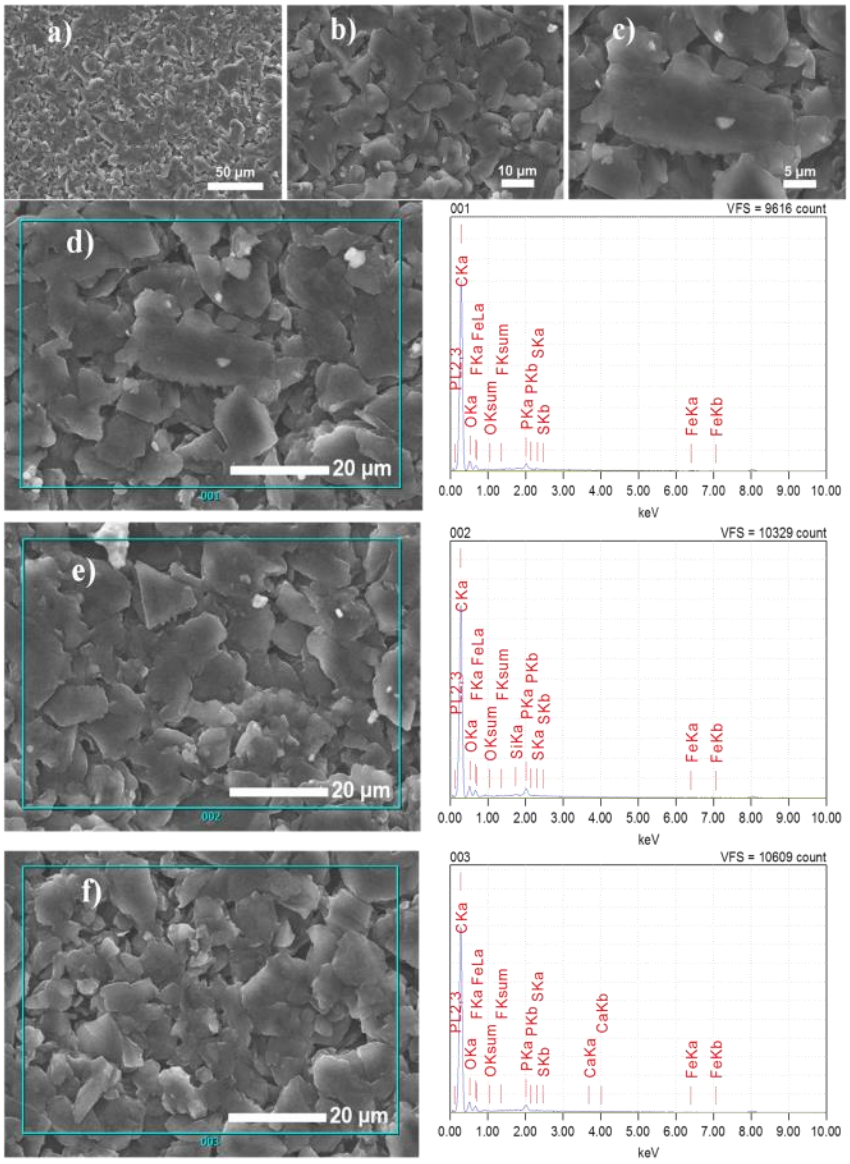


Fig. 5. *(a, b, c) SEM images of anode of the battery; (d, e, f) EDS analysis result at 3 zones for anode of the battery*

Figure 5a, 5b, 5c show the SEM images of the material that the anode of the battery is made of at different magnifications. Figure 5d, 5e, 5f show EDS analysis result at different zones of the anode. Table 1 reports the chemical analysis of the anode of the battery. As can be seen from Table 2, no other elements were detected in amount much higher than 1% but C, O, F. They are the main components in the anode in all three zones (C, O, F about 77%, 17%, 4%, respectively). Fe, P, S are presented in the anode with a very small in proportion. Especially, no Si and Ca were detected in zone 1, 3 and zone 1, 2, respectively.

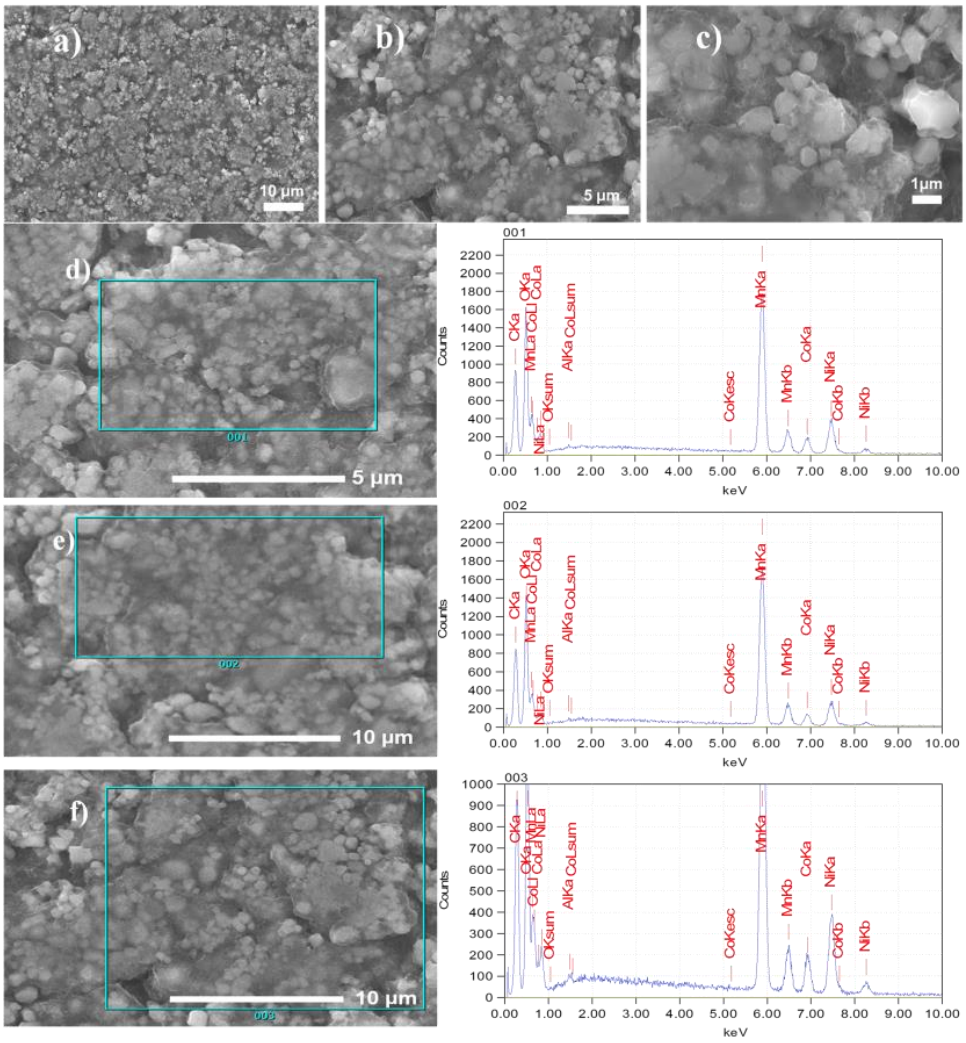


Fig. 6. *(a, b, c) SEM images of the cathode; (d, e, f) EDS analysis result at 3 zones of the cathode.*

Figure 6a, 6b, 6c show the SEM images of the material on the anode of the battery at different magnifications. Figure 6d, 6e, 6f show EDS analysis results at different zones of the cathode. Chemical analyses of the materials on the cathode are summarized in Table 3.

Table 3. **Chemical components (wt%) of the cathode at different zones**

Zone	1	2	3
C	22.96	22.71	24.67
O	24.27	24.51	23.35
Al	0.14	0.22	0.23
Mn	34.72	37.89	32.21
Co	5.12	4.28	5.6
Ni	12.78	10.41	13.94

Total	100	100	100

As can be seen from Table 3, Al element was detected in amount much smaller than 1%. The Mn, Co, Ni - transition metals present at about 34%, 5% and 12%, respectively, suggesting the presence of several lithium compounds that contain transition metal oxides (Mn, Co, Ni).

3. CONCLUSION

A commercial lithium-ion battery was disassembled to investigate. The characterization has been carried out by analysis means using EDS, XRD and SEM after having dismantled the battery into the single components. The results show that the anode is dominantly made of graphite while the cathode is made of LiM_2O_4 and $LiMO_2$ oxides (M = Mn, Ni, Co, Al or its combinations). The charge and discharge profiles at current density of 0.2 C (800 mA), between 3.0 - 4.25 V show that the battery has the first discharge capacity of 3820 mAh and the capacity is gradually decreased during cycling. The average working voltage is about 3.7 V which is well consisted with the rated voltage released by the provided company.

Acknowledgements

This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 103.02-2018.22.

Notes: The authors declare no competing financial interest.

REFERENCES

1. D. Ouyang, M. Chen, J. Liu, R. Wei, J. Weng, J. Wang (2018), *Investigation of a commercial lithium-ion battery under overcharge/over-discharge failure conditions*, RSC Adv.8 33414-33424.

2. N. Imanishi, T. Horiuchi, A. Hirano, Y. Takeda (2001), *Lithium intercalation mechanism of iron cyanocomplex*, Studies in Surface Science and Catalysis 132 935-938.

3. H. Qiao, Q. Wei (2012), *Functional Nanofibers in Lithium-ion batteries*, in: Q. Wei (Eds.), Functional Nanofibers and their Applications, Woodhead Publishing Series in Textiles, Oxford, UK, pp. 197-208.

1. 7. K. Jalkanen, J. Karppinen, L. Skogstrom, T. Laurila, M. Nisula, K. Vuorilehto **(**2015), *Cycle aging of commercial NMC/graphite pouch cells at different temperatures*, Appl. Energy 154**,** p.160-172.

2. 8. B. Gerelt-Od, H. Kim, U. J. Lee, J. Kim, N. Kim, Y. J. Han, H. Son, S. Yoon (2018), *Potential Dependence of Gas Evolution in 18650 Cylindrical Lithium-Ion Batteries Using In-Situ Raman Spectroscopy*, J. Electrochem. Soc. 165 A168-A174.

10. W. Waag, S. Käbitz, D. U. Sauer (2013), *Experimental investigation of the lithium-ion battery impedance characteristic at various conditions and aging states and its influence on the application*, Appl. Energy 102 885-897.

11. L. Bodenes, R. Naturel, H. Martinez, R. Dedryvere, M. Menetrier, L. Croguennec, J. Peres, C. Tessier, F. Fischer (2013), *Lithium secondary batteries working at very high temperature: Capacity fade and understanding of aging mechanisms*, J. of Power Sources 236 265-275.

12. D. Aurbach, B. Markovsky, A. Rodkin, M. Cojocaru, E. Levi, H. Kim (2002), *An analysis of*

rechargeable lithium-ion batteries after prolonged cycling, Electrochim. Acta 47 1899-1911.

13. N. Philip, P. Stefano, W. Martin (2013), *Interface Investigations of a Commercial Lithium Ion Battery Graphite Anode Material by Sputter Depth Profile X-ray Photoelectron Spectroscopy*, *Langmui*r 29 5806-5816.

14. Wang, S. Kadam, H. Li, S. Shi, Y. Qi (2018), *Review on modeling of the anode solid electrolyte interphase (SEI) for lithium-ion batteries*, Npj Comput. Mater. 4 1-26.

KHẢO SÁT CÁC ĐẶC TRƯNG ĐIỆN HÓA CỦA PIN LITI-ION THƯƠNG MẠI DẠNG TRỤ

Tóm tắt: *Pin liti-ion thương mại dạng trụ, kiểu dáng 26650, có dung lượng danh định 4000 mAh đã được tháo dỡ phục vụ nghiên cứu cấu trúc và thành phần cấu tạo của vật liệu điện cực. Các phép phân tích nhiễu xạ tia X (X-ray), hiển vi điện tử quét (SEM), phổ tán xạ năng lượng tia X (EDX) cho thấy vật liệu dương cực là hỗn hợp của các ôxít* $LiMn_2O_4$ *và* $LiMO_2$ *(M = Mn, Co, Ni), vật liệu âm cực là graphit. Vật liệu dương cực được cấu tạo từ các hạt ôxít tương đối đồng đều với đường kính trung bình trong khoảng từ 1-3 µm, vật liệu âm cực là các hạt graphit với đường kính trung bình khoảng 10 µm. Dung lượng xả của pin ở chu kỳ đầu tiên là 3820 mAh (tương ứng khoảng 95,5% dung lượng danh định). Dung lượng pin giảm dần trong các chu kì phóng, nạp tiếp theo. Hiệu điện thế hoạt động trung bình của quá trình phóng là 3.7 V, giá trị này tương đồng với hiệu điện thế hoạt động danh định của pin.*

Từ khóa*: Pin liti-ion, pin thương mại, vật liệu âm cực, vật liệu dương cực, hợp chất liti.*

FLAPW METHOD AND PRACTICAL CALCULATION FOR Gd CRYSTAL: ELECTRONIC STRUCTURE AND MAGNETIC PHASE STABILITY

Tran Van Quang

University of Transport and Communications

Abstact: *Description and understanding of electronic structures and magnetic properties of gadolinium Gd have been challenging. Especially, its magnetic phase stability of gadolinium has been in debate for a long time. In this report, the precise all-electron full-potential linearized augmented plane wave (FLAPW) method is introduced to study properties of Gd. Due to strongly localized f-states, the calculation may lead to weird results depending on defined parameters. The calculations including both 4f-core and 4f-band models are performed. The analysis of the electronic structure and magnetic phase stability are shown and discussed. All the results are good agreement with available experiments and previous theoretical reports.*

Keywords: *Gd phase stability, band structure, f-core model, f-band model, FLAPW method.*

Received: 15 March 2020

Accepted for publication: 20 April 2020

Email: tranquang@utc.edu.vn

1. INTRODUCTION

In modern material science, the economy-efficient approach to explore is using the density functional theory [1] proposed by Honhenberg, Kohn and Sham. The core of the theory is Kohn – Sham equation (in atomic unit) [2],

$$\left\{-\frac{\Delta}{2} + v(\vec{r}) + \int \frac{\rho(\vec{r}\prime)}{|\vec{r}-\vec{r}\prime|} d\vec{r}' + v_{xc}[\rho(\vec{r})]\right\}\psi_i = \varepsilon_i\psi_i, \quad (1)$$

$$\rho(\vec{r}) = \sum_{i=1}^{N} n_i \langle\psi_i|\psi_i\rangle, \quad (2)$$

is electron density, n the occupation number, $v_{xc}[\rho]=\delta E_{xc}[\rho]/\delta\rho$ the exchange-corelation potential, and v the external potential. One can solve this equation self-consistenly [3]. The

seft-consistently converged solution obtained gives us information of the ground states, e.g. eigenvalues ε_i, total energy, forces, and etc. [3] Nevertheless, this task is very demanding and the method to solve is still being developed in different ways, e.g. to deal with exchange correlation potential [4–6] and to develop numerical methods. Practically, when working on a magnetic system, many local minima may occur basically which infer multi-solutions. Some of the solutions are therefore unphysical meaning. Especially, in the case of strongly localized system such as Gd bulk (the well-known rare earth materials), the calculations may contain gosh states which originate from the strongly localized f states. As in general, the electron-nuclear interaction is given by the bare Cloulomb interaction whereas exchange correlation is very tough to describe. The strongly localized f states affect drastically in both of them. There are two classes of electrons: valence electrons (participate actively in chemical bonding), and core electrons (tightly bound to the nuclei, do not participate in bonding and to be treated as frozen orbitals). There is a third class of electrons called semi-core electrons. The f electrons usually are in this class. Its wave functions polarizes. There are two way to treat the problam: pseudopotential methods and all-electron methods. The precise all-electron full-potential linearized augmented plane wave (FLAPW) method is one of the most precise all electron method [3–5]. In this report we will present some matrix elements within FLAPW method. The exchange correlation potential will be treated by using local density approximation (LDA) [7]. The numerical results will be shown and discussed.

2. CONTENT

2.1. Hamiltonian matrix in FLAPW method

To solve Kohn-Sham equation, orbitals ψ are written as a linear combination of a complete basis set, i.e.

$$\psi_i(\vec{r}) = \sum_{\alpha}^{M} c_{i\alpha}\varphi_\alpha(\vec{r}). \quad \text{(M is dimension of the basis orbitals)} \tag{3}$$

For the specific basis set ϕ, in FLAPW, it is chosen by deviding space into interstitial and muffin-tin regions (here we are interested in 3D bulk calculations, if 2D oe 1D needed vacuum should be included) [3]

$$\varphi_{E_l}^{(\vec{k})} = \begin{cases} \dfrac{1}{\sqrt{\Omega}}\sum_{\vec{G}} c_{\vec{G}} e^{i(\vec{k}+\vec{G})\vec{r}} \ \text{ in the interstitial region}: \ |\vec{r}-\vec{R}| > r_{mt} \\ \sum_{l,m}\left[a_{lm}^{\mu,(\vec{k})}(E_l)u_l^\mu(r) + b_{lm}^{\mu,(\vec{k})}(E_l)\dot{u}_l^\mu(r)\right]Y_{lm}(\vec{r}^{\mu}) \\ \text{in the } a\text{tomic region muffin-tin } \mu: \ |\vec{r}-\vec{R}| < r_{mt} \end{cases} \tag{4}$$

In the muffin-tin region, these two radial wave functions are (i) the solutions of the radial Schrödinger equation, u_l, solved at a fixed energy, E_l; (atom unit)

$$\left(-\frac{d^2}{dr^2}+\frac{l(l+1)}{r^2}u_l(r)+V(r)-E_1\right)ru_l(r)=0, \tag{5}$$

and (ii) their derivatives $\dot{u}_l$. Y_{lm} are spherical harmonics and the coefficients a_{lm} and b_{lm} are determined by the requirement that the plane waves and their radial derivatives are continuous at the muffin-tin boundary. For the potential, there is no shape approximation assumed [3,8]

$$V(\vec{r})=\begin{cases}\sum_{\vec{G}} V_I^{\vec{G}} e^{i\vec{G}\vec{r}} & \text{interstitial}\\ \sum_{lm} V_{MT}^{lm}(r)Y_{lm}(\vec{r}_\mu) & \text{muffin-tin}\end{cases}, \tag{6}$$

Accordingly, hamiltonian and overlap matrices consist of two contributions from the two regions where space is divided, i.e. $H=H_I+H_{MT}$ and $S=S_I+S_{MT}$ in which *I* stands for "Interstitial" and *MT* "muffin-tin"

Contribution of muffin-tins. Let denote the quantum states as follow: $lm \to L;\ \ u_l Y_{lm} \to \varphi_L$. The contribution of muffin-tin to the Hamiltonian and overlap matrices are is given by inserting Eqs. (3,4) into Eqs. (2) and (1) to obtain

$$H_{MT}^{\vec{G}\prime\vec{G}}(\vec{k})=\sum_\mu \int_{MT^\mu} d\vec{r} \sum_{L\prime}\left(a_{L\prime}^{\mu\vec{G}}(\vec{k})\varphi_{L\prime}^{\alpha}(\vec{r})+ b_{L\prime}^{\mu\vec{G}}(\vec{k})\dot{\varphi}_{L\prime}^{\alpha}(\vec{r})\right)^* H_{MT^\alpha}\sum_L\left(a_L^{\mu\vec{G}}(\vec{k})\varphi_L^{\alpha}(\vec{r})+b_L^{\mu\vec{G}}(\vec{k})\dot{\varphi}_L^{\alpha}(\vec{r})\right), \tag{7}$$

$$S_{MT}^{\vec{G}\prime\vec{G}}(\vec{k})=\sum_\mu \int_{MT^\mu} d\vec{r} \sum_{L\prime}\left(a_{L\prime}^{\mu\vec{G}}(\vec{k})\varphi_{L\prime}^{\alpha}(\vec{r})+ b_{L\prime}^{\mu\vec{G}}(\vec{k})\dot{\varphi}_{L\prime}^{\alpha}(\vec{r})\right)^* \sum_L\left(a_L^{\mu\vec{G}}(\vec{k})\varphi_L^{\alpha}(\vec{r})+b_L^{\mu\vec{G}}(\vec{k})\dot{\varphi}_L^{\alpha}(\vec{r})\right), \tag{8}$$

These contain the following type of matrix elements

$$t_{L'L}^{\alpha\varphi\varphi}(\vec{k}) = \int_{MT^{\mu}} d\vec{r}\left(\varphi_{L'}^{\alpha}(\vec{r})\right)^{*} H_{MT^{\alpha}} \varphi_{L}^{\alpha}(\vec{r}) \tag{9}$$

H can be splited into two parts, the spherical H_{sp} and the nonspherical contributions V_{ns}, i.e.

$$H_{MT^{\alpha}} = H_{sp}^{\alpha} + V_{ns}^{\alpha} \tag{10}$$

Note that $\phi_{L}^{\alpha}, \dot{\phi}_{L}^{\alpha}$ can be chosen to diagonalize H_{sp}

$$H_{sp}^{\alpha}\varphi_{L}^{\alpha} = E_{l}\ \varphi_{L}^{\alpha}, \tag{11}$$

$$H_{sp}^{\alpha}\dot{\varphi}_{L}^{\alpha} = E_{l}\ \dot{\varphi}_{L}^{\alpha} + \varphi_{L}^{\alpha}. \tag{12}$$

Taking inner product with $\langle\varphi_{L}^{\alpha}|, \langle\dot{\varphi}_{L}^{\alpha}|$ respectively gives

$$\langle\varphi_{L'}^{\alpha}\left|H_{sp}^{\alpha}\varphi_{L}^{\alpha}\right\rangle_{MT^{\alpha}} = E_{l}\ \delta_{ll'}\delta_{mm'}, \qquad \langle\varphi_{L'}^{\alpha}|\varphi_{L}^{\alpha}\rangle_{MT^{\alpha}} = \delta_{ll'}\delta_{mm'}, \tag{13}$$

$$\langle\dot{\varphi}_{L'}^{\alpha}\left|H_{sp}^{\alpha}\varphi_{L}^{\alpha}\right\rangle_{MT^{\alpha}} = 0, \qquad \langle\dot{\varphi}_{L'}^{\alpha}|\varphi_{L}^{\alpha}\rangle_{MT^{\alpha}} = 0, \tag{14}$$

$$\langle\varphi_{L'}^{\alpha}\left|H_{sp}^{\alpha}\dot{\varphi}_{L}^{\alpha}\right\rangle_{MT^{\alpha}} = \delta_{ll'}\delta_{mm'}, \quad \langle\varphi_{L'}^{\alpha}\left|H_{sp}^{\alpha}\dot{\varphi}_{L}^{\alpha}\right\rangle_{MT^{\alpha}} = \langle\varphi_{L'}^{\alpha}\left|E_{l}\ \dot{\varphi}_{L}^{\alpha} + \varphi_{L}^{\alpha}\right\rangle_{MT^{\alpha}} = \delta_{ll'}\delta_{mm'}, \tag{15}$$

$$\langle\dot{\varphi}_{L'}^{\alpha}\left|H_{sp}^{\alpha}\dot{\varphi}_{L}^{\alpha}\right\rangle_{MT^{\alpha}} = E_{l}\ \delta_{ll'}\delta_{mm'}\langle\dot{u}_{l}^{\alpha}|\dot{u}_{l}^{\alpha}\rangle_{MT^{\alpha}}, \qquad \langle\dot{\varphi}_{L'}^{\alpha}|\dot{\varphi}_{L}^{\alpha}\rangle_{MT^{\alpha}} = \delta_{ll'}\delta_{mm'}\langle\dot{u}_{l}^{\alpha}|\dot{u}_{l}^{\alpha}\rangle_{MT^{\alpha}}. \tag{16}$$

It is noted that the potential is also expanded by using spherical harmonics, i.e.

$$V^{\alpha}(\vec{r}) = \sum_{L''} V_{L''}^{\alpha}(\vec{r}) Y_{L''}(\vec{r}). \tag{17}$$

Thus, hamiltonian matrix is obtained

$$H_{MT}^{\vec{G}'\vec{G}}(\vec{k}) = \\ = \sum_{\mu}\sum_{L'}\left[\begin{array}{c}\left(a_{L'}^{\mu\vec{G}'}(\vec{k})\right)^{*} t_{L'L}^{\alpha\varphi\varphi} a_{L}^{\mu\vec{G}}(\vec{k}) + \left(b_{L'}^{\mu\vec{G}'}(\vec{k})\right)^{*} t_{L'L}^{\alpha\dot{\varphi}\dot{\varphi}} b_{L}^{\mu\vec{G}}(\vec{k}) + \\ + \left(a_{L'}^{\mu\vec{G}'}(\vec{k})\right)^{*} t_{L'L}^{\alpha\varphi\dot{\varphi}} b_{L}^{\mu\vec{G}}(\vec{k}) + \\ + \left(b_{L'}^{\mu\vec{G}'}(\vec{k})\right)^{*} t_{L'L}^{\alpha\dot{\varphi}\varphi} a_{L}^{\mu\vec{G}}(\vec{k})\end{array}\right] \tag{18}$$

Where

$$t_{L'L}^{\alpha\varphi\varphi} = \sum_{l"} I_{l'll"}^{\alpha uu} G_{l'll"}^{m'mm"} + \delta_{ll'}\delta_{mm'}E_l,$$
$$G_{l'll"}^{m'mm"} = \int Y_{lm}^{*} Y_{l'm'} Y_{l"m"} d\Omega\,, \tag{19}$$
$$I_{l'll"}^{\alpha uu} = \int u_{l'}^{\alpha}(r) u_{l}^{\alpha}(r) V_{l"}^{\alpha}(r) r^2 dr$$

.

Similarly, the overlap matrix is

$$S_{MT}^{\vec{G}'\vec{G}}(\vec{k}) = \sum_{\mu} \sum_{L'} \left[\left(a_{L'}^{\mu\vec{G}'}(\vec{k}) \right)^{*} a_{L}^{\mu\vec{G}}(\vec{k}) + \left(b_{L'}^{\mu\vec{G}'}(\vec{k}) \right)^{*} b_{L}^{\mu\vec{G}}(\vec{k}) \right] \langle \dot{u}_l^{\alpha} | \dot{u}_l^{\alpha} \rangle_{MT^{\alpha}} \tag{20}$$

The interstitial contribution. Using basis function (4) for the interstitial region, the hamiltonian matrix is derived by noting that the kinetic energy is diagonal in momentum space and the potential is local, diagonal in real space and of convolution form in momentum space,

$$H_I^{\vec{G}\vec{G}'}(\vec{k}) = -\frac{\hbar^2}{2m}|\vec{G} + \vec{k}|^2 \delta_{\vec{G}\vec{G}'} + V(\vec{G} - \vec{G}'); \;\; S_I^{\vec{G}\vec{G}'} = \delta_{\vec{G}\vec{G}'}, \tag{21}$$

The muffin-tin *a*- and *b*-coefficients are determined by expanding planewave into spherical harmonics using Rayleigh expansion, i.e.

$$e^{i\vec{K}\vec{r}} = 4\pi \sum i^l j_l(rK) Y_L^{*}(\vec{K}) Y_L\ (\vec{r}), \tag{22}$$

where $r = |\vec{r}|;\ \vec{K} \equiv \vec{G} + \vec{k};\ K = |\vec{K}|$. The requirement of continuity of the wave functions at the muffin-tin boundary leads the coefficients *a* and *b* [9]

$$a_L^{\mu\vec{G}}(\vec{k}) = e^{i\vec{K}\vec{\tau}^{\mu}} \frac{4\pi i^l}{W} Y_L^{*}(\vec{R}^{\mu}\vec{K}) \left[\dot{u}_l(R_{MT^{\alpha}}) K j'_l\ (R_{MT^{\alpha}}K) - \dot{u}'_l(R_{MT^{\alpha}}) K j_l\ (R_{MT^{\alpha}}K) \right], \tag{23}$$

$$b_L^{\mu\vec{G}}(\vec{k}) = e^{i\vec{K}\vec{\tau}^{\mu}} \frac{4\pi i^l}{W} Y_L^{*}(\vec{R}^{\mu}\vec{K}) \left[u'_l(R_{MT^{\alpha}}) K j_l\ (R_{MT^{\alpha}}K) - u_l(R_{MT^{\alpha}}) K j'_l\ (R_{MT^{\alpha}}K) \right], \tag{24}$$

With

$$W = \dot{u}_l(R_{MT^\alpha})u'_l(R_{MT^\alpha}) - u_l(R_{MT^\alpha})\dot{u}'_l(R_{MT^\alpha}). \quad (25)$$

The density therefore can be obtained by tanking inner product from Eq. (2)

2.2. Numerical results

In numberical calculation for Gd pristine crystal, we used hexagonal structure with lattice constants *a* =6.89 au and *c*=10.92 au. The star-function cut-off, *Gmax*, is 11.5. The plane-wave cut off *Kmax* is 3.8. The spin polarization has been included. For the k-point mesh, we use 17×17×9 Monkhorst-Pack grids. The initial spin polarization is provided by starting magnetic moments of 7.0μ_B and 7.0μ_B. At first, Gd-4f states are treated as core. In this model so-called 4f-core model, we vary the lattice constants *a* and *c* and calculate the corresponding total energies. The results are presented in FIG. 1. This calculation shows the equilibrium lattice constants, i.e. *a* = 6.79 au and *c* = 10.80 au.

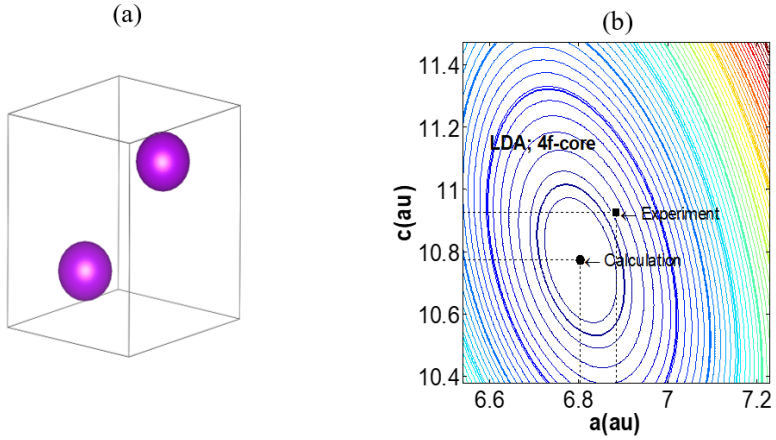


Fig. 1. *(a) Crystal structure of Gd and (b) its energy mesh. The minimum value infers the equilibirum lattice constants.*

***Table 1*. Equilibrium lattice constants and total magnetic moments within LDA calculation in 4f-core model together with experimental result.**

	a(au)			$\mu_{tot}(\mu B)$
LDA	6.79	c(au)	c/a	7.81
Experiment	6.88	10.92	1.59	7.63

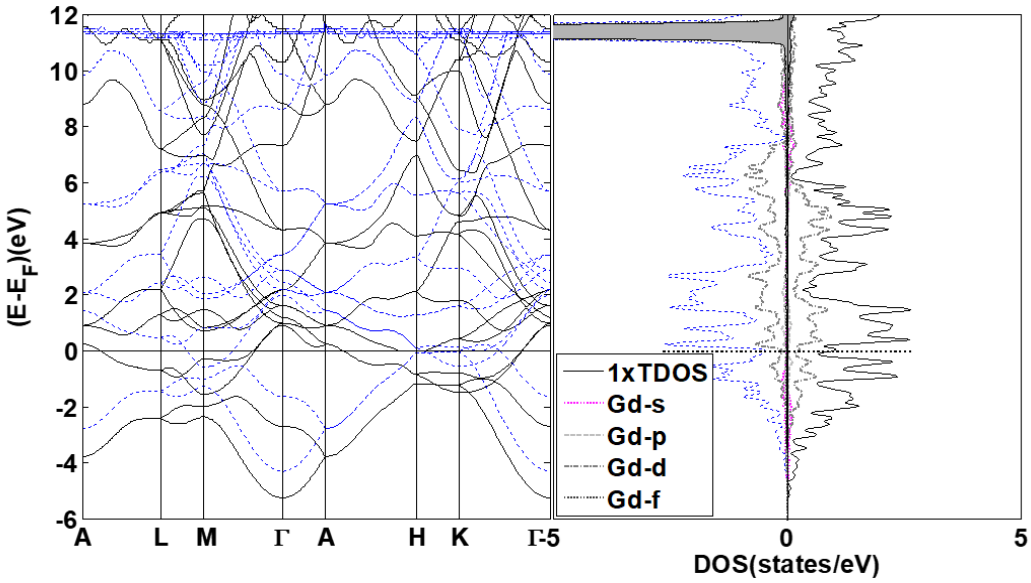


Fig. 2. *Band structure calculation and Density of states within LDA calculation and 4f-core model*

As can be seen, in the 4f-core model, calculation using LDA gives slightly underestimated equiriblium lattice constants as it does [6,10]. The results still are very well consistent with experiments and theoretical reports earlier [11]. The magnetic moment has been reported to be 7.41μ_B whereas our result shows 7.81μ_B and the experiment result is 7.63μ_B. Our calculated result is only 2.4% larger than the experimental value.

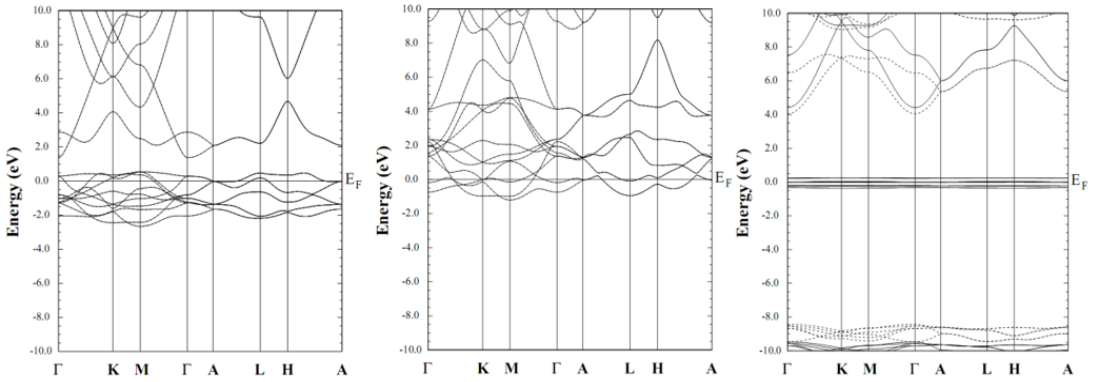


Fig. 3. *Band structure of Gd in 4f-band model with different Kmax values. Ghost states result in weird band structures.*

To continue, we examine the band structure of Gd. The results are presented in FIG. 2. As can be seen, 4f bands are disappeared from the valence band structure. It is well agreed with results reported of Ph Kurz et al. using all-electron FLAPW-FLEUR package [11].

In 4f-band model in which 4f electrons is treated as valence electrons, by using the experimental lattice constants, we found that some gosh states occur. These lead to weird results as shown in FIG. 3. These unrelevant results stem from chosing unappropriate parameters such as *Kmax* values and the gosh states appears during self-consistently solving.

Therefore, the parameters invoked must be opted to be very careful. After a number of tests, here we present the calculations with Gmax= 11.5, Kmax=3.8. We obtained relevant results, as presented in FIG. 4 for LDA calculation

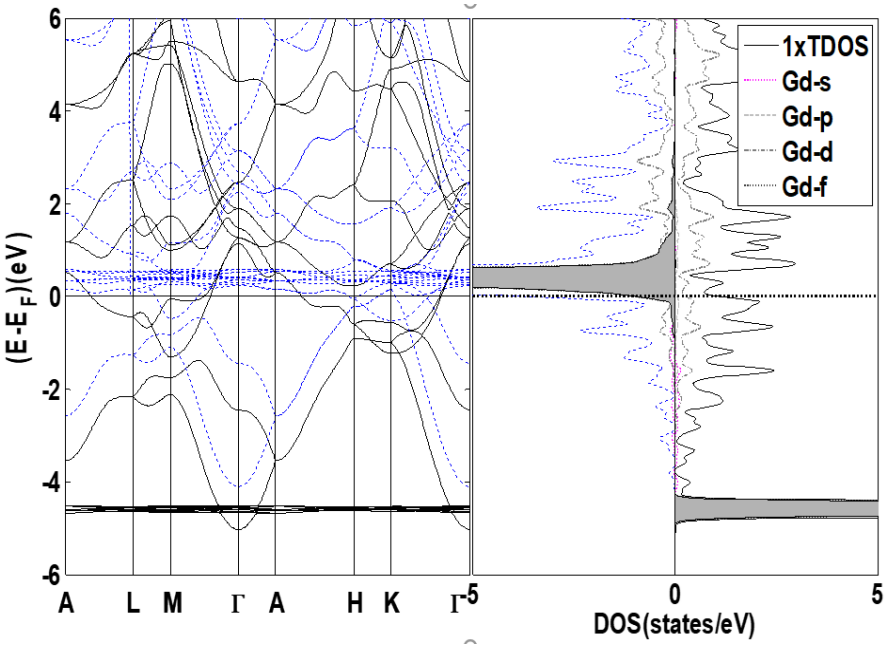


Fig. 4. *Band structure calculation and DOS within LDA calculation and 4f-band model*

As can be seen, within LDA calculation, Gd-4f states localize strongly at around -4.5 eV from the Fermi energy for majority spin and right beside the Fermi energy for the minority spin. The latter alters the band near the Fermi energy thereby the chemical bonding and the phase stability of Gd crystal. To take into account the effect of on-site interactions from f bands, Hubbard U correction is adapted, i.e. LDA+U calculation with the correlation energies of U_d = 5.0eV; J_d = 1.0eV and U_f = 7.7eV; J_d = 0.7eV [11–17]. The calculated electronic band structure is presented in FIG. 5. As shown, the on-site interaction with U and J corrections pushes majority and minority spins away. The majority spin locates at ~ -10.3 eV (deep) below Fermi energy. This explains why 4f-core model works for some cases, e.g. band structure as presented above, and 4f electrons play as semi-core electrons. The minority spin is at ~1.8eV above Fermi energy. The calculated results are excellent agreement with previous publications [11–14,18,19]. Note that all the calculations have been done by assuming that FM phase is stable. Next step, we will demonstrate that FM ordering is indeed stable.

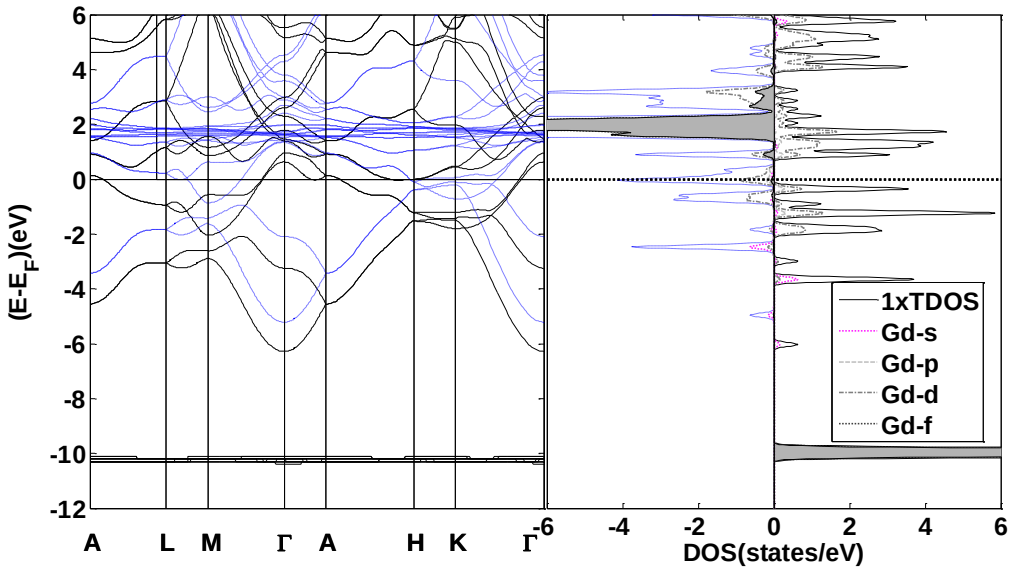


Fig. 5. *Band structure calculation and DOS within LDA+U calculation and 4f-band model*

Table 2. **Total energy (TE, in hatree) and energy differences, dE (in meV), between two phases FM and AFM**

	Approximation	TE, AFM phase	TE, FM phase	dE(meV) $=E_{FM}-E_{AFM}$
Calculation 1	LDA	-22545.5176615281	-22545.5173656716	8.1
	LDA+U	-22545.3641565677	-22545.3677646254	-98.2
Calculation 2	LDA	-22545.5176887987	-22545.5173456476	9.3
	LDA+U	-22545.3639499519	-22545.3678762767	-106.8

In order to do this, we carefully consider two sets of calculations. In *calculation 1***,** lattice constants are taken from Shick et al. [13] and we let x-axis be along [110] direction. Numer of states are 90 of which the highest state is about 54 eV above E_F. In *calculation 2***,** lattice constants are taken from Kurz et al. [11] and x-axis is along [010] direction. The number of states are 40 of which the highest state is about 19 eV above E_F. Basically, these two results of calculations should not be much different. For each calculation, we align magnetic moments to be parallel each other for FM and antiparallel for AFM and fix them during the self-consistent process to search for the minimum energy within both LDA and LDA+U calculations. The total energies are obtained by solving Eq. (1). We tabulate the results in Table II. Indeed, there are not much different between the two results of calculations. Accordingly, the calculated results show that in LDA calculation, the AFM is more stable with 8~9 meV lower than those of FM. Hower, in LDA+U calculation, the FM phase is more stable with 98~107meV lower than those of AFM. Our results are well agreement with results of Harmon el al. [20] in which LMTO+ASA calculation had been performed. And they found that within LDA calculation the energy difference is 8.2meV/atoms with AFM stable. In LDA+U calculation, the difference is -56.4meV/atom with FM stable. Shick et

al. [13] also found that FM is stable with energy different of about 63 meV (even this number is not clearly indicated for specific configurations in the paper) using LDA+U calculation within all-electron method. Kurz et al. [11] by using FLAPW-FLUER packages also demonstrated that the AFM phase is more stable over FM with -69meV in LDA calculation whereas the FM becomes more stable with energy difference of 34meV in the LDA+U calculation. By using the self-consistent semi-relativistic TB-LMTO-ASA method, Jenkins et al. [21] also argued that AFM is stable within LDA calculation with energy difference of 9.2 meV per atom. In another work, they used FP-LMTO method to prove both LDA and GGA giving AFM stable whereas in LMTO-ASA method, LDA gives AFM stable and GGA gives FM stable, with the energy difference of about 6mRy [22]. Petersen et al. [23] also used pseudo-potential method implemented by VASP package to testify that the orbital moment is very small and in GGA-PBE scheme, the energy different is ΔE=-7meV/atom with AFM stable whereas in GGA+U, calculated energy diference is 69 meV/atoms with FM stable. Our calculated results are excellent agreement with all these publications. And also 4f bands should be treated as valence bands with the Hubbard correction included, i.e. +U implementation [13].

3. CONCLUSION

FLAPW method is a very precise computational method to solve the modern material problems. It can well describe any system without shape approximation within atomic muffin-tin area, especially for dealing with the system with core structure, e.g. polarized wave functions. The use of input parameters should be very careful to obtain relevant results in the f compounds. The calculations applied for Gd show that Gd-4f can be treated either core, semi-core or valence states in some particular cases. The LDA scheme gives underestimated equilibrium lattice constants. Beyond this, it predicts excited f states to localize strongly near Fermi enery thereby the valence band close to Fermi level. Moreover, LDA calculation leads to AFM stable over FM phase whereas in LDA+U calculation, FM phase is more stable. This is reason giving rise to LDA+U implemented throughout the study and it should be invoked in studies of f-electron compounds. All the results from LDA and LDA+U calculations are well consistent with previous publications, especially for the proof of magnetic phase stability.

Acknowledgments

The author thanks Prof. Miyoung Kim at Sookmyung Women's University for the help and supervision.

REFERENCES

1. P. Hohenberg and W. Kohn (1964), Phys. Rev. 136, B864.
2. W. Kohn and L. J. Sham (1965), Phys. Rev. 140, A1134
3. E. Wimmer, H. Krakauer, M. Weinert, and A. J. Freeman (1981), Phys. Rev. B 24, 864
4. S. Youn and A. J. Freeman (2001), Phys. Rev. B 63, 085112.
5. M. Kim, A. J. Freeman, and C. B. Geller (2005), Phys. Rev. B 72, 035205.
6. T. Van Quang and K. Miyoung (2016), J. Korean Phys. Soc. 68, 393

7. U. Von Barth and L. Hedin (1972), J. Phys. C Solid State Phys. 5, 1629.
8. M. Weinert, G. Schneider, R. Podloucky, and J. Redinger (2009), J. Phys. Condens. Matter 21, 084201.
9. E. Wimmer, H. Krakauer, M. Weinert, and A. J. Freeman (1981), Phys. Rev. B 24, 864.
10. S. Youn and A. J. Freeman (2001), Phys. Rev. B 63, 085112.
11. M. Kim, A. J. Freeman, and C. B. Geller (2005), Phys. Rev. B 72, 035205.
12. T. Van Quang and K. Miyoung (2016), J. Korean Phys. Soc. 68, 393.
13. U. Von Barth and L. Hedin (1972), J. Phys. C Solid State Phys. 5, 1629.
14. M. Weinert, G. Schneider, R. Podloucky, and J. Redinger (2009), J. Phys. Condens. Matter 21, 084201.
15. S. Blügel and G. Bihlmayer (2006), Comput. Nanosci. 31, 85.
16. T. Van Quang and M. Kim (2017), J. Appl. Phys. 122, 245104.
17. P. Kurz, G. Bihlmayer, and S. Blügel (2002), J. Phys. Condens. Matter 14, 6353.
18. S. Abdelouahed, N. Baadji, and M. Alouani (2007), Phys. Rev. B 75, 094428.
19. A. B. Shick, A. I. Liechtenstein, andW. E. Pickett (1999), Phys. Rev. B 60, 10763.
20. C. Santos, W. Nolting, and V. Eyert (2004), Phys. Rev. B 69, 214412
21. T. Van Quang and M. Kim (2013), J. Appl. Phys. 113, 17A934.
22. T. Van Quang and M. Kim (2014), IEEE Trans. Magn. 50, 1000904.
23. T. Van Quang and M. Kim (2016), J. Korean Magn. Soc. 26, 39.
24. J. O. Dimmock and A. J. Freeman(1964), Phys. Rev. Lett. 13, 750.
25. L. Oroszlány, A. Deák, E. Simon, S. Khmelevskyi, and L. Szunyogh (2015), Phys. Rev. Lett. 115, 096402.
26. B. N. Harmon, V. P. Antropov, A. I. Liechtenstein, I. V. Solovyev, and V. I. Anisimov (1995), J. Phys. Chem. Solids 56, 1521.
27. A. C. Jenkins and W. M. Temmerman (1999), J. Magn. Magn. Mater. 198, 567.
28. A. C. Jenkins, W. M. Temmerman, R. Ahuja, O. Eriksson, B. Johansson, and J. Wills (2000), J. Phys. Condens. Matter 12, 10441.
29. M. Petersen, J. Hafner, and M. Marsman (2006), J. Phys. Condens. Matter 18, 7021.

PHƯƠNG PHÁP FLAPW VÀ ỨNG DỤNG NGHIÊN CỨU TINH THỂ Gd: CẤU TRÚC ĐIỆN TỬ VÀ TÍNH BỀN VỮNG CỦA CÁC PHA TỪ FM VÀ AFM

Tóm tắt: *Mô tả và hiểu biết về cấu trúc điện tử và thuộc tính từ tính của gadolinium là một vấn đề đầy thử thách. Đặc biệt, cấu trúc từ bền vững nhất của Gd đã gây nhiều tranh luận trong một thời gian dài. Trong báo cáo này, phương pháp thế toàn phần sóng phẳng gia tăng được giới thiệu để nghiên cứu các tính chất của Gd. Do các trạng thái f định xứ rất mạnh, phép tính có thể cho kết quả rất lạ lùng phụ thuộc vào tham số định nghĩa khác nhau. Việc tính toán bao gồm các mô hình 4f-core và 4f-band được thực hiện. Phân tích cấu trúc điện tử và độ ổn định pha từ được trình bày và thảo luận. Tất cả các kết quả là phù hợp tuyệt vời với các kết quả thực nghiệm và báo cáo lý thuyết trước đó.*

Từ khoá: *Cấu trúc Gd bền vững, cấu trúc, mô hình f-lõi, mô hình f-band, mô hình f hóa trị, phương pháp FLAPW.*

HIGH ENERGY SCATTERING AMPLITUDE IN THE LINEARIZED GRAVITATIONAL THEORY

Do Thu Ha

Hanoi University of Science

***Abstract**: The asymptotic behavior of the elastic scattering amplitude by the exchange of graviton between two scalar particles at high energies and fixed momentum transfers is reconsidered in the Logunov-Tavkhelidze equation in the linearized gravitational theory. The corrections to the eikonal approximation in the quasi-potential approach of relative order 1/p is developed with the principal contributions at high energy. The eikonal expression of scattering amplitude and the formal first correction are derived. The Yukawa potential is applied to discuss the results.*

***Keywords:** Field theory, scattering amplitude, eikonal approximation.*

Received: 11 March 2020

Accepted for publication: 20 April 2020

Email: thuhahunre@gmail.com

1. INTRODUCTION

The eikonal approximation (which is also called straight-line path approximation) is an effective method of calculating the scattering amplitude at high energies and is studied by many authors in quantum field theory [1-4], and in quantum gravity theory recently [5-13]. However, in different approaches, only the main term of amplitude was considered, while the first correction does not have an explicit solution. Researches [9,10] in which the path-integral method with a modified perturbation theory and Logunov-Tavkhelidze quasi-potential are used to give the analytic expression of the first correction. Thus, the advantage of the quasi-potential approach is affirmed and need to be studied more deeply.

The aim of this paper is to make a more detailed investigation of the quasi-potential approach by solving the quasi-potential equation [9-10] to find the eikonal scattering amplitude and the first correction at high energies and minor momentum transfers.

The paper is organized as follows. In section 2.1, an eikonal approximation for the scattering amplitude and the first correction are derived by using a quasi-potential approach in the coordinate representation. This result is applied to the Yukawa potential in section 2.2.

The last section, we draw our conclusion.

2. CONTENT

2.1. Correction terms of scattering amplitude

First, we will derive homogenous equation for one-time wave function of an interactive two scalar particle system. To do this, we start from 4-time Green function $G_{ab}(x, y; x', y')$ which must be satisfied the Bethe-Salpeter equation [14], and can be written down in a symbolic form [14,15].

$$G = \alpha G^0 + \alpha^{-1} G^0 K G, \qquad \alpha = (2\pi)^4, \tag{1}$$

where $G^0$is the Green function of free particles, and the kernel K can be found by the perturbative method.

Solving (1) by using the reduction technique with a relation between the 2-time Green function $\tilde{G}_{ab}$and 4-time Green functionG_{ab}, following the procedure of ref. [16], we have an explicit equation for the 2-time Green function in momentum representation.

$$F(p'^2; E^2)\tilde{G}(p', p, E) - \int dq^3 V(p', q, E)\tilde{G}(q, p, E) = \delta(p - p') \tag{2}$$

where $F_{a,b}(p^2; E^2) = (p^2 + m_{a,b}^2 - E^2)\sqrt{p^2 + m_{a,b}^2}$, and $V(p', q, E)$ is a potential matrix.

From (2) and the relation $\tilde{G}_{ab}(t, \vec{x}, \vec{y}) = \sum_n \varphi_n(t; \vec{x}, \vec{y})\, \varphi_n^+(t; \vec{x}, \vec{y})$ between the 1-time wave function and the 2-time Green function, the homogenous equation of 1-time wave function will be

$$(\vec{p}^2 - E^2 + m^2)\psi(\vec{p}) = \int \mathrm{d\,q} \rightarrow \frac{\mathrm{V[(\,p \rightarrow - \,q \rightarrow\,)^2; E]}\psi(\vec{q})}{\sqrt{\mathrm{m}^2 + \vec{q}^2}} \tag{3}$$

Considering Eq. (3) in the coordinate representation with a purely imaginary local quasi-potential $V(\vec{r}; E) = ipEv(\vec{r})$in which $v(\vec{r})$is a smooth positive function and $p = |\vec{p}|$. At high energies and small scattering angles, wave function $\psi_p(\vec{r})$ can be written in the form $\psi_p(\vec{r}) = e^{ipz}F_p(\vec{r})$, $F_p(\vec{r})\big|_{z\rightarrow-\infty} = 1$. By the way of expanding terms in inverse powers of momentum, and keeping only terms of the order 1/p takes the form, the solution of Eq. (3) will be [18,19]

$$F_p(\vec{r}) = \exp\left[-\frac{z\theta(z)\gamma(\rho)}{2ip} - \int_{-\infty}^{z} v(\rho, z')\, dz' - \frac{1}{2ip}\int_{-\infty}^{z} \tilde{\chi}^{(1)}(\rho, z')\, dz'\right], \tag{4}$$

where

$$\tilde{\chi}^{(1)}(\rho,z)=\chi^{(1)}(\rho,z)-\theta(z)\gamma(\rho),\ \int_{-\infty}^{z}\chi^{(1)}(\rho,z')\,dz'=\int_{-\infty}^{z}\tilde{\chi}^{(1)}(\rho,z')\,dz'+z\theta(z)\gamma(\rho)$$

$$\chi^{(0)}=v(r),\ \chi^{(1)}=-3\left[\partial_z v(r)-v(r)^2-\eta_\perp(\vec{r},v)\right],\ \eta_\perp(\vec{r},\chi)=-\int_{-\infty}^{z}\nabla_\perp^2\chi(\rho,z')\,dz'+\left(\int_{-\infty}^{z}\nabla_\perp\chi(\rho,z')\,dz'\right)^2.$$

The scattering amplitude is related to the wave function as follows

$$T(\Delta^2;E)=\frac{1}{(2\pi)^3}\int dr\,\psi^{*(0)}_k V(E;r)\,\psi_p(r), \tag{5}$$

where, $\Delta^2=(p-k)^2=\Delta_\perp^2+\Delta_z^2=-t$ and$\Delta_z=\frac{\Delta_\perp^2}{2p}+O\left(\frac{1}{p^2}\right)$.

Substituting (4) into (5) and integrating by part, we obtain

$$T(\Delta^2;E)=T^{(0)}(\Delta^2;E)+\frac{1}{2ip}T^{(1)}(\Delta^2;E)+\ldots, \tag{6}$$

where the eikonal approximation for the amplitude is

$$T^{(0)}(\Delta^2;E)=-2ipE\frac{1}{(2\pi)^3}\int d^2\rho e^{i\rho\Delta_\perp}(e^{-\int_{-\infty}^{\infty}v(\rho,z')\,dz'}-1), \tag{7}$$

and the first correction in this approximation

$$\begin{aligned}&T^{(1)}(\Delta^2;E)\\&=2ipE\frac{1}{(2\pi)^3}\Bigg\{\int d^2\rho\, e^{i\rho\Delta_\perp}e^{-\int_{-\infty}^{\infty}v(\rho,z')\,dz'}3\int_{-\infty}^{\infty}v^2(\rho,z)\,dz\int_{-\infty}^{\infty}dz\,z\,v(\rho,v)\\&-\int d^2\rho dz\ e^{i\rho\Delta_\perp}\Delta_\perp^2 e^{-\int_{-\infty}^{z}v(\rho,z')\,dz'}\\&+\int d^2\rho\, e^{i\rho\Delta_\perp}\int_{-\infty}^{\infty}dz\,\eta_\perp(\rho,v)\left(e^{-\int_{-\infty}^{z}v(\rho,z')\,dz'}-\ e^{\int_{-\infty}^{\infty}v(\rho,z')\,dz'}\right)\Bigg\}.\end{aligned} \tag{7}$$

Using method of integration by part [9] and quasi-potential approach in the momentum representation [10] these results (7), (8) can also be found.

Now, let us consider the case where momentum transfers t = 0 and the quasi-potential has the Gaussian form $V(E;\Delta^2)=i$sge^{at}, $t=-\Delta^2$which the corresponding form in the coordinate representation is

$$V(E;r)=isg\sqrt{\pi/a}\,e^{-r^2/4a}. \tag{9}$$

Since$t=\Delta_\perp^2+\Delta_z^2=0$, it follows$\Delta_\perp=0$. Substituting (9) into (8), and noticing that on the mass shell$p^2=E^2-m^2\ \Rightarrow p\propto E\propto\sqrt{s}$, the first correction term will be

$$T^{(1)}(\Delta^2=0;E) \propto 3isg\frac{g}{\pi\sqrt{8\pi a}}\int d^2\rho e^{-\rho^2/2a}e^{2i\chi_0} + isg\frac{1}{8\pi^2 a}\int d^2\rho e^{-\rho^2/4a}(1-)$$
$$\times \int_{-\infty}^{\infty} dz \int_{-\infty}^{z} V(z')dz'\left(\exp\left[2i\chi\int_{-\infty}^{z}V(z')\,dz'\right]-\exp\left[2i\chi\int_{-\infty}^{\infty}V(z')\,dz'\right]\right), \quad (10)$$

Where $2i\chi = -4\pi g e^{-\rho^2/4a}, V(z) = \frac{1}{\sqrt{4\pi a}}e^{-z^2/4a}$.

The similar result Eq. (10) is also found by the Born approximation in momentum representation [7].

2.2. Asymptotic behavior of the scattering amplitude at high energies

In the previous section, the general form of the scattering amplitude of two scalar particles is found in the potential $V(\vec{r};E)$.

Now, let us consider a particular example in which the graviton exchange[1] [9] the quasi-potential increases with energy $V(r,s) = (\kappa^2 s e^{-\mu r}/2\pi r)$. Substituting this Yukawa potential into (7), (8) and noticing that at high energies, $p \propto E \propto \sqrt{s}$, we find the leading term of the scattering amplitude

$$T^{(0)}(\Delta^2;E) \propto \frac{\kappa^2 s}{(2\pi)^4}\left(\frac{1}{\mu^2-t}-\frac{\kappa^4}{2(2\pi)^2}F_1(t)+\frac{\kappa^4}{3(2\pi)^5}F_2(t)\right) \quad (11)$$

and the first correction term

$$T^{(1)}(\Delta^2;E) = \frac{3i\kappa^6}{(2\pi)^6}\left(F_1(t)-\frac{2\kappa^3}{(2\pi)^3}F_2(t)\right) \quad (12)$$

Where

$$F_1(t) = \frac{1}{t\sqrt{1-\frac{4\mu^2}{t}}}\ln\left|\frac{1-\sqrt{1-4\mu^2/t}}{1+\sqrt{1-4\mu^2/t}}\right| \quad (13)$$

---

[1] The model of interaction of a scalar "nucleons" with a gravitational field in the linear approximation to $h_{\mu\nu}(x)$ $L(x) = L_{0,\varphi}(x) + L_{0.grav}(x) + L_{\text{int}}(x)$ where

$L_{0,\varphi}(x) = \frac{1}{2}\left[\partial^\mu\varphi(x)\partial_\mu\varphi(x) - m^2\varphi^2(x)\right]$; $L_{\text{int}}(x) = -\frac{\kappa}{2}h^{\mu\nu}(x)T_{\mu\nu}(x)$;

$T_{\mu\nu}(x) = \partial_\mu\varphi(x)\partial_\nu\varphi(x) - \frac{1}{2}\eta_{\mu\nu}\left[\partial^\mu\varphi(x)\partial_\mu\varphi(x) - m^2\varphi^2(x)\right]$; and $T_{\mu\nu}(x)$ is the energy momentum tensor of the scalar field. The coupling constant κ is related to the Newton constant of gravitation G by$\kappa^2 = 32\pi G = 32\pi l_{PL}^2$. $l_{PL} = 1,6.10^{-33}cm$ is the Planck length.

$$F_2(t) = \int_0^1 dy \frac{1}{(ty+\mu^2)(y-1)} \ln\left|\frac{\mu^2}{y(ty+\mu^2-t)}\right| \quad (14)$$

These results (11), (12) have similar forms in [9]. Moreover, from these equations we see that the first correction term of the eikonal expression of the scattering amplitude at high energies and fixed momentum transfers increases rapidly in the linearized gravitational theory. Comparison of these above potentials has made it possible to draw the following conclusions: in the model with the scalar exchange, the total cross section σ_t decreases as $(1/s)$, and only the Born term predominates in the entire eikonal equation; the vector model leads to a total cross section σ_t approaching a constant value as $s \to \infty, (t/s) \to 0$. In both cases, the eikonal phases are purely real and consequently the influence of inelastic scattering is disregarded in this approximation, $\sigma^{in} = 0$. In the case of graviton exchange the Froissart limit is violated. A similar result is also obtained in Ref. [20] with the eikonal series for reggeized graviton exchange.

3. CONCLUSION

The asymptotic behavior of the scattering amplitude at high energies and fixed momentum transfers has been studied within a quasi-potential approach in the coordinate representation in the linearized gravitational theory. The obtained results of eikonal expression of the scattering amplitude and the corresponding first correction term coincide with the results found by other authors [9-10]. The Yukawa potential has been used to concretize the results.

Acknowledgements

The author is grateful to Prof. Nguyen Suan Han for his suggestions of the problem and many useful comments. This work is supported in part by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under grant number 103.01-2018.42 and the project 911 of Hanoi University of Science – VNUHN.

REFERENCES

1. M. Barbashov, S. P. Kuleshov, V. A. Matveev, V. N. Pervushin, A. N. Sissakian and A. N. Tavkhelidze (1970), *Phys.Lett. 33B 484.*

2. E.Eichen and R. Jackiw (1971), Phys. Rev. D4 439.

3. G.'tHooft (1987), Phys. Lett. 198B 61.

4. L. N. Lipatov(1992, 1991), Phys. Lett. B116 411; Nucl. Phys.B365 614.

5. M. Fabbrichesi, R. Pettorino, G. Veneziano, G. A. Vilkovisky (1994), Nucl. Phys. B419 147.
6. R. Kirschner (1995), Phys. Rev. D52, N4 2333.
7. Nguyen Suan Han and E. Ponna (1997), Nuovo Cimento A, 110 459.
8. Nguyen Suan Han (2000), Eur. Phys. J. C16 547.
9. Nguyen Suan Han, Nguyen Nhu Xuan (2002), Eur. Phys. J.C24 643;
10. Nguyen Suan Han, Do Thu Ha, Nguyen Nhu Xuan (2019), Eur. Phys. J. C, 79:835.
11. D. Amati, M. Ciafaloni and G. Veneziano (1990), Nucl. Phys.B347 550.
12. E. Verlinde and H. Verlinde (1992), Nucl. Phys. B371 246.
13. D. Kabat and M. Ortiz (1992), Nucl. Phys. B388 570.
14. E. Salpeter and H. Bethe (1951), Phys. Rev. 84 1232.
15. M. Gell-man and F. Low (1951), Phys. Rev. 84 350.
16. Nguyen Van Hieu and R. N. Faustov, *Dubna, preprint P 1235 1963.*
17. V. A. Shcherbina (1973), Theor. and Math. Phys.14 253.
18. P.S. Saxon and L. I. Schiff (1957), Nuovo Cim., 6, 614.
19. V.R. Garsevanishvili, S.V. Goloskokov, V.A. Matveev, L.A. Slepchenko, and A.N. Tavkhelidze (1971), Theor. and Math. Phys 6, 36.
20. I. Muzinich, M. Soldate (1998), Phys. Rev. D 37, 353.

BIÊN ĐỘ TÁN XẠ NĂNG LƯỢNG CAO TRONG LÝ THUYẾT HẤP DẪN TUYẾN TÍNH

Tóm tắt: *Dáng điệu tiệm cận của biên độ tán xạ đàn hồi bằng cách trao đổi graviton giữa hai hạt vô hướng ở vùng năng lượng cao và xung lượng truyền cố định được nghiên cứu lại qua phương trình chuẩn thế Logunov-Tavkhelidze trong lý thuyết hấp dẫn tuyến tính. Các bổ chính đối với gần đúng eikonal, bằng cách tiếp cận khai triển gần đúng theo lũy thừa của bậc* $1/p$ *được phát triển, cùng với các đóng góp chính ở vùng năng lượng cao. Biểu thức eikonal của biên độ tán xạ và bổ chính đầu tiên chính thức đã thu được. Thế Yukawa được áp dụng để thảo luận kết quả.*

Từ khóa: *Lý thuyết trường, biên độ tán xạ, gần đúng eikonal.*

SYNTHESIS, STRUCTURE OF SOME α, β - UNSATURATED KETONES FROM ACETONE AND ALDEHYDES

Pham Van Hoan, Nguyen Xuan Trinh

Hanoi Metropolitan University

Abstract: *Diverse biological activities of α, β- unsaturated ketones such as antibacterial, anti-fungal, weed and insecticide killer, anti-cancer of liver, lung.. have been mentioned in many studies. Otherwise, α,β-unsaturated ketones are derivatives that play a very important role in the synthesis of heterocyclic compounds with very plentiful activity. There have been a number of studies synthesizing [1,2,3,4,5,6,7] and converting them into diverse heterocyclic compounds [8, 9, 10,..] showing the important role of α, β- unsaturated ketones thank to their diverse applications. This paper introduces the synthesis some α, β-unsaturated ketones from acetone and some aldehydes.*

Keywords: *α, β- unsaturated ketones, IR and NMR spectrum, structure.*

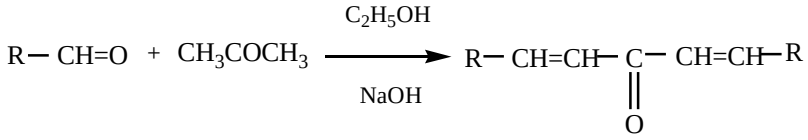
Received: 19 March 2020

Accepted for publication: 20 April 2020

Email: pvhoan@hnmu.edu.vn

1. INTRODUCTION

Synthesis α,β-unsaturated ketones are synthesized according to general chemical equations:



Here: R is hydrocarbon radical.

The reaction is carried out in an ethanol environment with sodium hydroxide catalyst, the ratio of substances involved (aldehyde/acetone) in the reaction is 2: 1 in moles. After recrystallization, the product is measured for melting point, checked for purity by thin silica gel chromatography in solvent system of n-hexane and ethyl *acetate according to appropriate volume ratio and spectrometry to determine structure.*

2. CONTENT

2.1 Confirm structure of synthesized compounds

The synthesized substances after measuring the melting point are stable and checked by thin layer chromatography silica gel showing round and neat marks will be determined by modern spectroscopic methods (IR, ^{1}H NMR).

2.1.1. Measuring the melting temperature

Melting temperature of synthesized compounds was measured on Gallenkamp (UK) at the Organic Chemistry Laboratory, Department of Chemistry, Hanoi National University of Education.

2.1.2. Measuring IR spectrum

Infrared spectrum of substances is recorded on PTS-6000 (Bio-Rad, USA) at the Infrared Spectroscopy Room, Institute of Chemistry – Vietnam Academy of Science and Technology, tablet form with solid KBr.

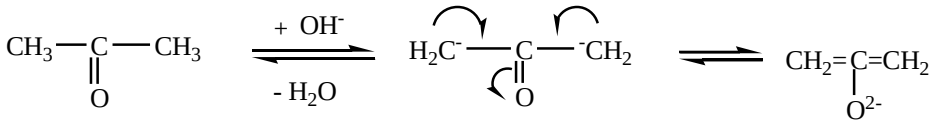
2.1.3. Measuring nuclear magnetic resonance spectrum

^{1}H NMR spectrum of substances is measured in solvent DMSO or $CDCl_3$ (TMS is the standard) by Brucker Avance 500 MHz machine at The Nuclear Magnetic Resonance Spectroscopy Room of Chemistry Institute – Vietnam Academy of Science and Technology.

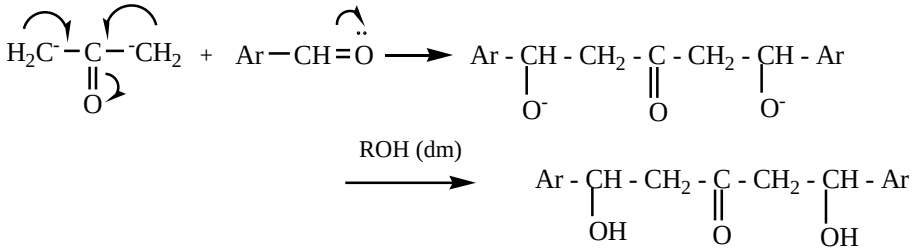
2.2. Disscuss results

2.2.1. Disscuss results of synthesis α,β- unsaturated ketones

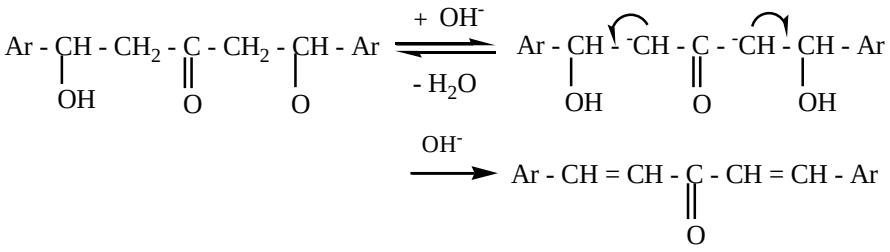
There are many ways to synthesize α, β– unsaturated ketones, but the most common is the aldol – croton condensation reaction. In the synthesis process, the used catalyst used was 20% sodium hydroxide solution. In terms of reaction conditions, we used a simpler, cheaper agent than the previous authors [1; 2; 3] because the authors also closed the cumarin loop before the synthesis of ketones. The α,β– unsaturated ketones are synthesized according [5; 6; 7] with fairly high performance. The reaction went through 2 stages of add-separation as follows [5]: Addition phase: The role of the base in the aldol additive reaction is to activate the methylene component so that it is easily added to the carbonyl group of the aldehyde molecule**.**



The resulting carbanion attacks the carbonyl group of the aldehydes molecule to form aldol compounds:

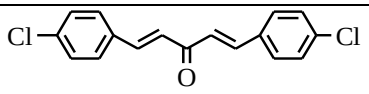
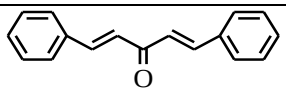
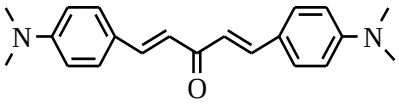
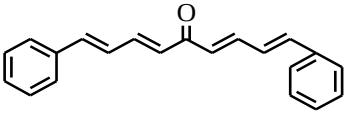


Separation phase: The product of the reaction usually has a trans configuration. The decomposition reaction takes place by E1cb reaction mechanism:



The intermediate product is a carbanion; The less negatively charged carbanion the more stable it is, the higher the efficiency. In the process of synthesizing α, β– unsaturated ketones, we only use one ketone (acetone) so the efficiency and speed of the reaction depend on the aldehyde nature. The higher the density of positive charge on the C atom in the –CHO group of the aldehyde molecule, the easier the reaction will occur. Such substituents having –I, –C effects will increase the positive charge density on the carbon atoms in the –CHO group and vice versa. Reaction performance values are perfectly consistent than expected.

Table 1. Some physical characteristics of α, β– unsaturated ketones

	α, β– unsaturated ketones	Some physical characteristics
A		Needle-shaped crystals, light yellow, m.p 196^0-$196,5^0$C. Efficiency 73%.
B		Flake crystals, bright yellow m.p 111 - 112^0C. Efficiency 70%.
C		needle-shaped crystals, dark red; m.p 193 – $193,5^0$C. Efficiency 65%.
D		needle-shaped crystals, dark yellow; m.p 148 - $148,5^0$C. Efficiency 72%.

Comparing the conditions of synthesis A, B, C, D, we found that: for A, B, D the reaction was easier (at normal temperature) than C (mixed boiling). The reason is that 2 $N(CH_3)$ groups causes a strong + C effect that reduces the positive charge density on carbon atoms

in the –CHO group of 4-N,N-dimethylaminobenzaldehyde, which reduces the reactivity. This is entirely consistent with the theoretical basis. Synthesized α, β– unsaturated ketones are all crystals, well soluble in solvents DMF, ether, acetone, DMSO, slightly soluble in alcohol and insoluble in water.

2.2.2. Research results of molecular structure

2.2.2.1. Infrared spectrum (IR) of α, β– unsaturated ketones

On the infrared spectrum of α, β– unsaturated ketones appear patterned at about 1630–1652 cm^{-1}, having a strong intensity, typical for the valence of C = O. The valence band of C = C group (the alkene associated with the aromatic nucleus and the C = O group) and C = C (the aromatic nucleus) in the lower region is about 1572–1600cm^{-1}, has a stronger intensity than the C=O. Especially at 980–999cm^{-1} there is a characteristic oscillation for non-flat deformation of vinyl group with trans-configuration when joining into combination with C = O group.

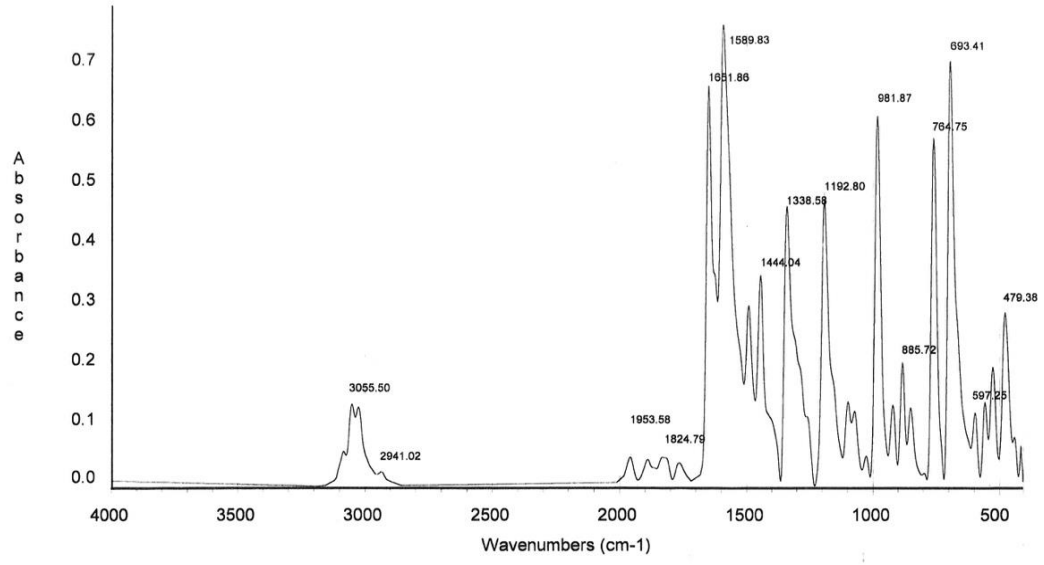


Fig. 1*. Infrared spectrum of compound B*

On the IR spectrum of α, β– unsaturated ketones we find: There was no signals of characteristic valence oscillation of CH bonding of -CHO group ($\nu_{CH\ aldehyde}$ = 2700–2900 cm^{-1}, usually show 2 signals, includes one at 2700 cm^{-1}). Valence oscillation of C=O group has lower frequency than C=O of initial ketone ($\nu_{C=O\ saturated}$ = 1700-1720 cm^{-1}) and aldehydes ($\nu_{C=O}$ = 1680–1715 cm^{-1}). Valence oscillation of C=C group has lower frequency but intensity is higher than isolated C=C group ($\nu_{C=C\ isolated}$ = 1620 –1680 cm^{-1}). The reason is that the combination of C=O group with C=C group reduces the bonding multiple, so they oscillate at smaller frequencies. Thus, the results obtained when analyzing the IR spectrum

above show that the synthesized α, β– unsaturated ketones are consistent/match with the initial estimates.

Table 2. **Results of IR interpret of α, β– unsaturated ketones (R-CH=CH-)$_2$C=O**

Kí hiệu	R	$\nu_{CH\ Ar,\ alkene}$	$\nu_{CH\ sat.}$	$\nu_{C=O}$	$\nu_{C=C,\ C=N}$	γ-CH= (*trans*)
A	4–ClC_6H_4–	3012,5	-	1626,8	1586,3	982,3
B	C_6H_5–	3055,5	-	1651,8	1589,8	981,8
C	4–$(CH_3)_2NC_6H_4$–	2896,3	2803,8	1636,7	1518,2; 1596,3	986,4
D	C_6H_5–CH=CH–	3036,4	-	1639,9	1572,7; 1491,4	998,2

2.2.2.2. ^{1}H NMR spectrum of α, β– unsaturated ketones

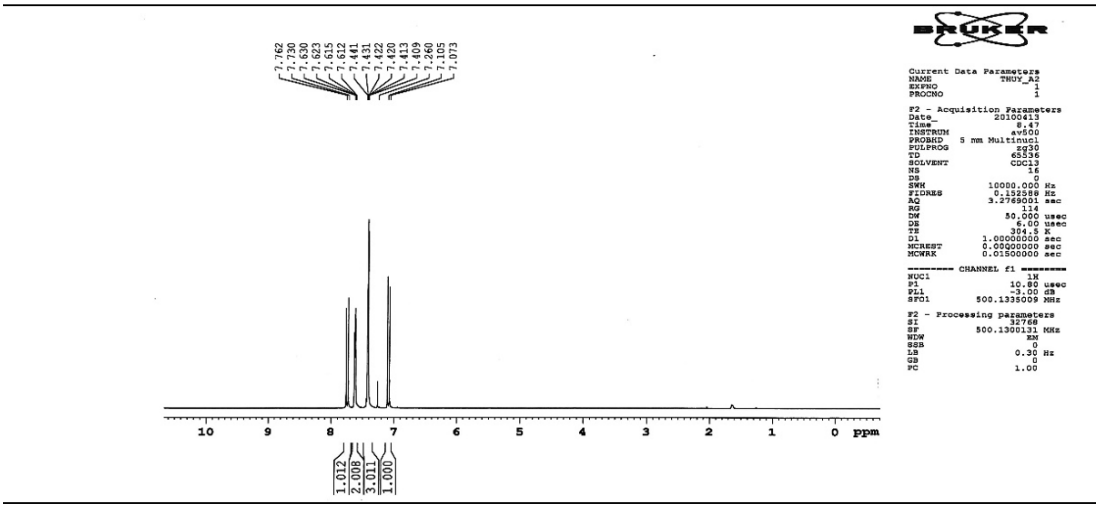


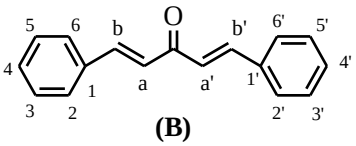
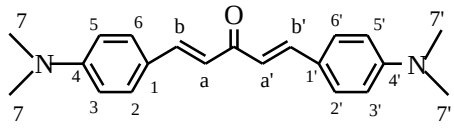
Fig. 2. *^{1}H NMR spectrum of compound B in $CDCl_3$*

On the 1H NMR spectrum of α, β– unsaturated ketones, protons gathered into signal clusters with chemical shift in the range of 6.6 – 7.8 ppm. This is the typical chemical shift for unsaturated protons of = CH– group and there is a pair of doublets in the form of the roof effect in the range of 6.9 – 7.8 ppm with spin – spin interaction constant J = 15 – 16 Hz of group protons –CH = CH– proving that they exist in *trans*-configuration (consistent with IR spectrum). In addition, on the spectrum of compound C, there is also a characteristic signal of CH_3–N–Ar group.

Because α, β–unsaturated ketones are symmetric molecules, there are many equivalent protons, moreover, the protons of the benzene ring have quite close displacement in the range of 6.6 – 7.62 ppm. Based on the relative strength, chemical displacement and spin – spin interaction constant, the ^{1}H NMR spectrum analysis results of B and C are shown in the following table 3.

Table 3. Results of ^{1}H NMR interpret of B, C (δ ppm)

(J (Hz); s: singlet; d: doublet; dd: doublet–doublet; t: triplet; q: quartet; m: multiplet)

Positions	 **(B)**	 **(C)**
$H^{a, a'}$	7,09; d; 16; 1H	6,88; d; 15,5; 1H
$H^{b, b'}$	7,74; d; 16; 1H	7,68; d; 16; 1H
$H^{2, 2'}$; $H^{6, 6'}$	7,62; d-d; 6,5; 1,5; 2H	7,5; d; 9; 2H
$H^{3, 3'}$; $H^{5, 5'}$	7,41; d-d; 5; 1; 2H	6,69; d; 8,5; 2H
$H^{4, 4'}$	7,43; d; 5; 1H	-
H^{7}, $H^{7'}$	-	3,03; s; 3H

Thus, the results of the above analysis show that: α, β– unsaturated ketones synthesized enough protons as expected. The relative intensity, chemical displacement as well as the spin – spin separation constant of the protons are consistent with the originally expected structure. The results allow proper conclusions about the structure of α, β– unsaturated ketones synthesized.

3. CONCLUSION

Four α, β– unsaturated ketones have synthesized from acetone and aldehydes that have not found in references.

The combination of interpret results of IR and ^{1}H NMR spectrums confirmed structure of synthesized compounds.

REFERENCES

1. Nguyen Minh Thao, Nguyen Van Thuan, Tran Hoai Nam, Nguyen Ngoc Thanh (2009) Synthesis some α,β– unsaturated ketones from 3-acetyl-6-methoxy-4-methylcumarin. Journal of Chemistry, V.47(4), pp.461-465 (in Vietnamese).

2. Duong Ngoc Toan, Nguyen Minh Thao (2019), *Synthesis some α,β– unsaturated ketones from 6-acetyl-5-hydroxy-4-methylcumarin*. J. of Science and Technology, Thai Nguyen University, 200(07), pp. 215-221 (*in Vietnamese*).

3. Nguyen Minh Thao, Nguyen Van Vinh, Tran Quoc Toan, Nguyen Duc Chinh, Dong Thi Duyen (2009), *Synthsis some α,β– unsaturated ketones from methylcumarin*. Journal of Chemistry, V.47(1), pp. 22-27 (*in Vietnamese*).

4. Nakamura Y., and et al, (2004), *A tropical ginger sesquiterpene, activates phase II drug metabolizing enzymes*. FEBS Lett., 572(1-3), pp. 245-250.

5. Egi M., Yamaguchi Y., Fujiwara N., Akai S., (2008), *Mo-Au combo catalysis for rapid 1,3-rearrangement of propargyl alcohols into α,β-unsaturated carbonyl compounds*. Org. Lett., 10, pp. 1867-1870.

6. Madkour H. M. F., (1993), *Reactions with 2-methyl-8-phenyl-benzopiran-4-one and Its Derivatives.* Rev. Roum. Chim., 38, pp. 11171125.
7. Valenta P. P., Drucker N. A., Bode J. W., Walsh PP. J. (2009), *Simple One-pot Conversion of Aldehydes and Ketones to Enals*. Org. Lett., 11, pp. 2117-2119.
8. Mamta Rani, Yushup Mohamad (2014), *Synthesis, study in vivo antibacterial activities of some 5-(thiophene-2-yl)phenylpyrazoline derivaties.* J. Of Saudi Chemical Society, 2014, 18, 411-417.
9. Mohamad Yushup, Payal Jain (2014), *Synthetic and biological studies of pyrazolines and related heterocyclic compounds.* Arabian Journal of Chemistry, 2014, 7, 533-596.
10. Naveen Tanwer, Rajneesh Kaur, Devika Rana, Raman Singh, Kuldeep Singh (2015), *Synthesis and characterization of pyrazoline derivatives*. Journal of Integrated Science and Technology, 2015, 3(2), 39-41.

TỔNG HỢP, CẤU TRÚC MỘT SỐ KETONE A, B- KHÔNG NO TỪ ACETONE VÀ ALDEHYDE

Tóm tắt: *Các ketone α,β- không no có hoạt tính rất đa dạng, như kháng khuẩn, chống nấm, diệt cỏ dại và trừ sâu, chống ung thư gan, phổi,... đã được đề cập trong nhiều công trình nghiên cứu [1,2,3,4,5,6,7]. Ngoài ra, các ketone α,β- không no còn là loại dẫn xuất có vai trò rất quan trọng trong việc tổng hợp các hợp chất chứa dị vòng có hoạt tính rất phong phú. Đã có nhiều nghiên cứu tổng hợp [1,2,3,4,5,6,7,...] và chuyển hóa chúng thành các hợp chất chứa dị vòng [8, 9, 10,..] cho thấy vai trò quan trọng của các ketone α,β- không no. Bài viết này giới thiệu kết quả nghiên cứu tổng hợp, cấu trúc một số ketone α,β- không no từ acetone và aldehyde.*

Từ khóa: *Xeton α, β- không no, phổ IR và phổ NMR, cấu trúc phân tử.*

CONSIDER SOME PRIMARY PROBLEMS FROM PERSPECTIVE OF ALGEBRAIC STRUCTURES

Dong Huu Mau

Hanoi Metropolitan University

Abstract: *This paper considers the nature of some primary problems from the perspective of the algebraic structure of groups, rings, fields, etc. Thereby explaining the rationale for the solution to those problems.*

Keywords: *Algebraic structure, primary problem, field extension, number set.*

Received: 10 March 2020

Accepted for publication: 20 April 2020

Email: dhmau@hnmu.edu.vn

1. INTRODUCTION

In high school, we often study primary problems with basic solutions on arithmetic properties, equivalent transformations, or acknowledged methods. For example, when solving a product equation on the real number set $A(x).B(x) = 0$, we give the equivalent of solving two equations $A\ (x)\ =\ 0$ or $B\ (x)\ =\ 0$; or solving a first-degree equation with an unknown $ax\ =\ b$ where $a, b \in \mathbb{R}, a \neq 0$, we divide both sides by a and get the solution $x\ = \frac{b}{a}$. Besides, the number sets $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$ and $\mathbb{C}$ also are studied, but the construction of them is not paying attention. Here we will consider the nature of the solutions of the above problems by the properties of algebraic structures such as group, ring, integer, and field properties of the set considered respectively.

2. CONTENT

First we will review some algebraic structures.

2.1. Group, ring and field

A *group* is a set G together with a binary operation, denote by (*), such that:

1) Associativity: For any $x, y, z \in\ G$, we have $(x\ *\ y) * z\ =\ x * (y * z)$.

2) Identity: There exists an $e \in G$ such that $e\ *\ x\ =\ x\ *\ e\ =\ x$ for any $x \in G$. We say that e is an identity element of G.

3) Inverse: For any $x \in\ G$, there exists a $y \in\ G$ such that $x\ *\ y\ =\ e\ =\ y\ *\ x$. We

say that y is an inverse of x.

A *ring* A is a set with two binary operations (addition and multiplication) such that:

i) A is an abelian group with respect to addition (so that A has a zero element, denoted by 0, and every $x \in A$ has an inverse, $-x$).

ii) Multiplication is associative: $(xy)z = x(yz)$ for all $x, y, z \in A$.

iii) Multiplication distributive over addition:

$$x(y+z) = xy + xz, (y+z)x = yx + zx \text{ for all } x, y, z \in A.$$

A *zero-divisor* in a ring A is an element x which "divides 0", i.e., for which there exists $y \neq 0$ in A such that $xy = 0$. In other hand, if $xy = 0$ then $x = 0$ or $y = 0$.

A ring A, which is commutative ($xy = yx$), has an identity element (denoted by 1) and has no zero-divisor (and in which $1 \neq 0$), is called an *integral domain*.

A *field* is an integral domain A in which every non-zero element is a unit.

A given binary relation ~ on a set X is said to be an *equivalence relation* if and only if it is reflexive, symmetric and transitive. That is, for all a, b and c in X:

- $a \sim a$. (Reflexivity)
- $a \sim b$ if and only if $b \sim a$. (Symmetry)
- if $a \sim b$ and $b \sim c$ then $a \sim c$. (Transitivity)

The equivalence class of $a \in X$ under ~, denoted $\bar{a}$, is defined as $\bar{a} = \{b \in X | b \sim a\}$. The set

$$X/\sim = \{\bar{a} | a \in X\}$$

is called the quotient set of X by ~.

A field extension E/F is called a *simple extension* if there exists an element θ in E with $E = F(\theta)$. The element θ is called a primitive element, or generating element, for the extension; we also say that E is generated over F by θ. In other hand, E is a simple extension of F generated by θ then it is the smallest field which contains both F and θ.

2.2 Product equation

The rationale for explaining the solution of a product equation is the notion of zero-divisor. For the sets, $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \text{and } \mathbb{C}$ have no zero-divisor so that $ab = 0$ implies $a = 0$ or $b = 0$.

Consider the product equation $A(x).B(x) = 0$ on $\mathbb{Z}$ (or $\mathbb{Q}, \mathbb{R}, \mathbb{C}$). In fact, this is the problem of finding $x \in \mathbb{Z}$ (or $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ respectively) such that $A(x).B(x) = 0$. For each $x \in \mathbb{Z}$ (or $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ respectively) then $A(x), B(x)$ are elements of $\mathbb{Z}$ (or $\mathbb{Q}, \mathbb{R}, \mathbb{C}$ respectively). For the sets, $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \text{and } \mathbb{C}$ have no zero-divisor so that $A(x).B(x) = 0$ implies $A(x) = 0$ or $B(x) = 0$. Reserve, if $A(x) = 0$ or $B(x) = 0$ then $A(x).B(x) = 0$ is obviously. Therefore, $A(x).B(x) = 0$ is equivalent to $A(x) = 0$ or $B(x) = 0$.

The same explanation for the solution of a product equation in the case of two variables, three variables, etc.

Example 1. Consider the equation $(x+3)(2x+4)=0$ (1) on $\mathbb{R}$. For every $x \in \mathbb{R}$, we have $x+3 \in \mathbb{R}, 2x+3 \in \mathbb{R}$. Since $\mathbb{R}$ has no zero-divisor so that (1) be equivalent to $x+3=0$ or $2x+4=0$, lead to $x=-3$ and $x=-2$ satisfies equation (1).

We have another way to explain equation (1) as follows: The expressions $x+3$ and $2x+4$ are two elements of the polynomial ring $\mathbb{R}[x]$. Since $\mathbb{R}$ is a field, $\mathbb{R}[x]$ is an integral domain, so that $\mathbb{R}[x]$ has no zero-divisor. Thus (1) is equivalent to $x+3=0$ or $2x+4=0$.

2.3. First degree equation with an unknown and difference between two numbers

The nature of these problems is the existence of the symmetric element x' of an element x in group X. If the operation on X is a multiplication, we call x' as the inverse of x. If the operation on X is an addition, we call x' the opposite element of x.

If b has the symmetric element then we can define the division (dividing a by b is multiplying a by the inverse element of b) and subtract (a minus b equals a plus with the opposite element of b).

Consider $\mathbb{Z}, \mathbb{Q}, \mathbb{R},$ and $\mathbb{C}$ together with addition, they are groups, so every element has an opposite element. Therefore, we always solve the problem of finding the difference of two numbers a and b on these sets. The result is the sum of a and the opposite element of b, that is, $a-b = a + (-b)$.

Consider the first-degree equation with an unknown $ax=b$ with the coefficients $a, b \in \mathbb{R}$, $a \neq 0$ (similar to $\mathbb{Q}$ or $\mathbb{C}$). Since $\mathbb{Q}, \mathbb{R},$ and $\mathbb{C}$ are fields so a has an inverse element a^{-1}. Now we multiply both sides by the inverse element of a, we get $x = a^{-1}b = \frac{b}{a}$.

2.4. Number sets

In this section, we consider the construction of number sets from the perspective of algebraic structures.

2.4.1. Integer set

In elementary school teaching, integers are often intuitively defined as the (positive) natural numbers, zero, and the negations of the natural numbers. However, this style of definition leads to many different cases (each arithmetic operation needs to be defined on each combination of types of integer) and makes it tedious to prove that these operations obey the laws of arithmetic. Therefore, in modern set-theoretic mathematics a more abstract construction, which allows one to define the arithmetical operations without any case distinction, is often used instead. The integers can thus be formally constructed as the equivalence classes of ordered pairs of natural numbers (a,b). Now we will construct the set of integers $\mathbb{Z}$:

Consider $\mathbb{N} \times \mathbb{N} = \{(a,b) | a,b \in \mathbb{N}\}$ and we define an a binary relation $\sim$ on these pairs with the following rule: : $(a,b) \sim (c,d)$ if and only if $a+d=b+c$. It is easy to check, reflexive property: $(a,b) \sim (a,b)$, symmetric property: if $(a,b) \sim (c,d)$ then $(c,d) \sim (a,b)$

and transitive property: if $(a,b)\sim(c,d)$ and $(c,d)\sim(e,f)$ then $(a,b)\sim(e,f)$. Therefore, $\sim$ is an equivalence relation on $\mathbb{N}\times\mathbb{N}$. Set $\mathbb{Z} = \mathbb{N}\times\mathbb{N}/\sim$ is a set of equivalence classes. With $x\in\mathbb{Z}$ and $(a,b)\in x$, we denote $x = \overline{(a,b)}$. Now, we equippe two binary operations on $\mathbb{Z}$, one called addition and the other called multiplication:

$$\overline{(a,b)} + \overline{(c,d)} = \overline{(a+c,b+d)}$$
$$\overline{(a,b)}.\overline{(c,d)} = \overline{(ac+bd,ad+bc)}.$$

It is easy to prove $\mathbb{Z}$ with two binary operations is a commutative ring whose the additive identity is $\overline{(0,0)} = \overline{(n,n)}$ and the multiplicative identity is $\overline{(1,0)} = \overline{(n',n)}$, forall n $\in\mathbb{N}$, the inverse of $\overline{(a,b)}$ is $\overline{(b,a)}$ which is denoted $-\overline{(a,b)}$. Now, we prove $\mathbb{Z}$ has no nonzero zero divisors: if $\overline{(a,b)}.\overline{(c,d)} = \overline{(ac+bd,ad+bc)} = \overline{(n,n)}$ then $ac+bd = ad+bc$ or $(a-b)(c-d) = 0$. Without loss of generality we may assume that $-b\in\mathbb{N}$ and $c-d\in\mathbb{N}$. Then we have $a-b=0$ or $c-d=0$, lead to $a=b$ or $c=d$. In case of $a=b$ then $\overline{(a,b)} = \overline{(0,0)}$; and when $c=d$ then $\overline{(c,d)} = \overline{(0,0)}$. Thus $\mathbb{Z}$ is an integer domain. Consider a map

$$f:\mathbb{N}\to\mathbb{Z}, a\mapsto\overline{(a,0)}.$$

Since $\overline{(a,0)} = \overline{(b,0)}$ if and only if $a=b$. Hence f is an injection. Forall $a,b\in\mathbb{N}$, we have

$$f(a+b) = \overline{(a+b,0)} = \overline{(a,0)} + \overline{(b,0)} = f(a)+f(b),$$
$$f(ab) = \overline{(ab,0)} = \overline{(a,0)}.\overline{(b,0)} = f(a)f(b).$$

Consequently, f is both a homomorphism of the monoids with addition and both a homomorphism of the monoids with multiplication. Since f is an injection, it follows that f is both a monomorphism of the monoids with addition and both a monomorphism of the monoids with multiplication.

Since $\overline{(a,b)} = \overline{(a,0)} + \overline{(b,0)} = \overline{(a,0)} - \overline{(b,0)} = f(a)-f(b)$. Hence, forall elements in $\mathbb{Z}$ have the form $f(a)-f(b)$.

The pair $(\mathbb{Z},f)$ defined as above is unique, differing from one isomorphism, meaning that if there is a pair (P,g) where P is a ring and $g:\mathbb{N}\to P$ is both a monomorphism of monoids with addition and both a monomorphism of monoids with multiplication, the elements of P have the form $g(a)-g(b)$, there exists an isomorphism $\varphi:\mathbb{Z}\to P$ such that $\varphi.f = g$. This is evidenced by a map:

$$\varphi:\mathbb{Z}\to P, f(a)-f(b)\mapsto g(a)-g(b).$$

It is tobe noticed that, f is a monomorphism. Hence, we can identical element $n\in\mathbb{N}$ with element $\overline{(n,0)}\in\mathbb{Z}$. This leads to $\mathbb{N}\subset\mathbb{Z}$. Therefore, every $x\in\mathbb{Z}, x=\overline{(a,b)}$, we have: if $a\geq b$ then $x = \overline{(a-b,0)} = a-b$. And if $a<b$ then $x = -\overline{(b,a)} = -\overline{(b-a,0)} = -(b-a)$. It follows that every $x\in\mathbb{Z}$, either $x\in\mathbb{N}$ or $-x\in\mathbb{N}$.

The ring $\mathbb{Z}$ above is called *ring of integers.*

2.4.2. Rational set

In mathematics at high school, a rational number is a number that can be expressed as the quotient or fraction $\frac{a}{b}$ of two integers, a numerator a, and a non-zero denominator b. Since b may be equal to 1, every integer is a rational number. The set of all rational numbers often referred to as "the rationals", is usually denoted by $\mathbb{Q}$.

In algebraic structures, the set of all rational numbers is constructed as a field. Rational numbers can be formally defined as equivalence classes of pairs of integers (a, b) such that $b \neq 0$, for the equivalence relation defined by $(a, b) \sim (c, d)$ if, and only if $ad = bc$. With this formal definition, the fraction $\frac{a}{b}$ becomes the standard notation for the equivalence class of (a, b). Now we will construct a field of rational numbers from an integer ring:

Consider $\mathbb{Z} \times \mathbb{Z}^* = \{(a, b) \mid a, b \in \mathbb{Z}, b \neq 0\}$, where $\mathbb{Z}^*$ is the set of nonzero integer numbers. We define an a binary relation $\sim$ on these pairs with the following rule: $(a, b) \sim (c, d)$ if and only if $a.d = b.c$. It is easy to check that $\sim$ is an equivalence relation on $\mathbb{Z} \times \mathbb{Z}^*$. Set $\mathbb{Q} = \mathbb{Z} \times \mathbb{Z}^* / \sim$ is the set of equivalence classes.

With $x \in \mathbb{Q}$ and $(a, b) \in x$, we denote $x = \overline{(a, b)}$. Now, we equippe two binary operations on $\mathbb{Q}$, one called addition and the other called multiplication:

$$\overline{(a, b)} + \overline{(c, d)} = \overline{(ad + bc, bd)}, \overline{(a, b)}.\overline{(c, d)} = \overline{(ac, bd)}.$$

It is easy to prove $\mathbb{Q}$ with two binary operations above is a field whose the additive identity is $\overline{(0,1)}$ and the multiplicative identity is $\overline{(1,1)}$. Consider the map

$$f: \mathbb{Z} \to \mathbb{Q}, a \mapsto \overline{(a, 1)}.$$

It is clear that f is a ring monomorphism. Hence, we can identical element $a \in \mathbb{Z}$ with element $\overline{(a, 1)} \in \mathbb{Q}$. This leads to $\mathbb{Z} \subset \mathbb{Q}$. And every $x \in \mathbb{Q}$ can be written

$$x = \overline{(a, b)} = \overline{(a, 1)}.\overline{(1, b)} = f(a)f(b)^{-1} = ab^{-1}.$$

The pair $(\mathbb{Q}, f)$ defined as above is unique, differing from one isomorphism, meaning that if there is a pair (P, g) where P is a field and $g: \mathbb{Z} \to P$ is a ring monomorphism, the elements of P have the form $x = g(a)g(b)^{-1}$, $a, b \in \mathbb{Z}, b \neq 0$, there exists an isomorphism

$$\varphi: \mathbb{Q} \to P, f(a)f(b)^{-1} \mapsto g(a)g(b)^{-1}$$

such that $\varphi.f = g$.

The fact that $\mathbb{Q}$ is the field of quotients of the integral domain of integers. The field $\mathbb{Q}$, which is constructed above, is called *field of rational numbers*. Every rational number may be expressed in the form $\frac{a}{b}$, which is called fraction, where a and b are integer numbers and $b \neq 0$. Then two fractions are equal, write

$$\frac{a}{b} = \frac{c}{d} \text{ if and only if } ad = bc;$$

two fractions are added as follows:

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd};$$

and the rule for multiplication is:

$$\frac{a}{b} . \frac{c}{d} = \frac{ac}{bd}.$$

2.4.3. Real set

In basis mathematics, a real number is a value of a continuous quantity that can represent a distance along a line. The real numbers include all the rational numbers, such as the integer −3 and the fraction 5/4, and all the irrational numbers, such as $\sqrt{3}, \sqrt[3]{4}, e, \pi$. There are also many ways to construct the real number system, for example, starting from natural numbers, then defining rational numbers algebraically, and finally defining real numbers as equivalence classes of their Cauchy sequences or as Dedekind cuts, which are certain subsets of rational numbers. Another possibility is to start from some rigorous axiomatization of Euclidean geometry (Hilbert, Tarski, etc.) and then define the real number system geometrically. All these constructions of the real numbers have been shown to be equivalent, that is the resulting number systems are isomorphic. In this section, we describe a constructive way set of real numbers from Cauchy sequences.

Let X be the set of all Cauchy sequences of rational numbers. That is, sequences $x_1, x_2, \dots, x_n, \dots$ of rational numbers such that for every rational $\mathcal{E} > 0$, there exists an integer N such that for all natural numbers $m, n > N$ we have $|x_m - x_n| < \mathcal{E}$.

Cauchy sequences $\{x_n\}$ and $\{y_n\}$ can be added and multiplied as follows:

$$\{x_n\} + \{y_n\} = \{x_n + y_n\}, \{x_n\}.\{y_n\} = \{x_n . y_n\}.$$

It is easy to prove X with two binary operations above is a commutative ring whose the additive identity is $\{0\}_{n\in\mathbb{N}}$ and the multiplicative identity is $\{1\}_{n\in\mathbb{N}}$. And subset

$$I = \left\{\{x_n\} \in X | \lim_{n\to+\infty} x_n = 0\right\}$$

is an ideal of X. Therefore, the quotient X/I is a field. The inverse of nonzero elements $\alpha = \overline{\{x_n\}} = \{x_n\} + I \in X/I$ is $\overline{\{y_n\}} \in X/I$ where

$$y_n = \begin{cases} 0 \; if \; n \le n_1 \\ \frac{1}{x_n} \; if \; n > n_1 \end{cases}.$$

The field X/I, which is denoted by $\mathbb{R}$, is called *field of real numbers*.

2.4.4. Complex set

In high school, a complex number is defined by giving its form ($z = a + bi;\ a, b \in \mathbb{R}$) with the appearance of an imaginary number i that satisfies $i^2 = -1$.

In algebraic structure, the complex set $\mathbb{C}$ is constructed from the set $\mathbb{R}^2 = \{(a, b) | a, b \in \mathbb{R}\}$ with the addition and multiplication operations defined: For all $(a, b), (c, d) \in \mathbb{R}^2$ we have

$$(a, b) + (c, d) = (a + c, b + d);\ (a, b).(c, d) = (ac - bd; ad + bc).$$

The set $\mathbb{R}^2$ together with the two operations is a field, which is called a *complex field* $\mathbb{C}$.

It is clear that the map $f: \mathbb{R} \to \mathbb{C}, a \mapsto (a, 0)$.

is a field monomorphism. Hence, we can identical element $a \in \mathbb{R}$ with element $(a, 0) \in \mathbb{C}$. This leads to $\mathbb{R} \subset \mathbb{C}$.

Let $i = (0, 1) \in \mathbb{C}$. We have $i^2 = (0, 1).(0, 1) = (-1, 0) = -1$ and every $x = (a, b) \in \mathbb{C}$ can be written $x = (a, b) = (a, 0) + (0, b) = (a, 0) + (b, 0).(0, 1) = a + b.i$.

From the perspective of field extension theory, $\mathbb{C}$ is a simple extension of $\mathbb{R}$ with algebraic element i on $\mathbb{R}$, that is, $\mathbb{C} = \mathbb{R}(i)$.

3. CONCLUSION

In this article, I present an explanation for solutions to product equation with a variable, first-degree equation with an unknown, problems of finding differences between two numbers or construction of number sets from the perspective of algebraic structures.

REFERENCES

1. M. F. Atiyah, I. G. Macdonald (1969), *Introduction to Commutative Algebra*, Massachusetts : Addison-Wesley Publishing.
2. Patrict Morandi (1996), *Field and Galois Theory*, Springer.
3. Bui Huy Hien (2005), *General Algebra Exercises*, Education Publishing.
4. Hoang Xuan Sinh (2015), *General Algebra*, Education Publishing.
5. Hoang Quoc Viet, Dam Van Nhi (2017), *Number Theory and Polynomial Bases*, University of Education Publishing.

XEM XÉT MỘT SỐ BÀI TOÁN SƠ CẤP DƯỚI GÓC NHÌN CỦA CÁC CẤU TRÚC ĐẠI SỐ

Tóm tắt: *Bài viết này xem xét bản chất của một số bài toán sơ cấp dưới góc nhìn của cấu trúc đại số nhóm, vành, trường, từ đó giải thích cơ sở lí luận cho lời giải của bài toán đó.*

Từ khóa*: Cấu trúc đại số, bài toán sơ cấp, mở rộng trường.*

SOME PROBABILITY MODELS FOR ARTIFICIAL INTELLIGENCE

Nguyen Van Tuan

Hanoi Metropolitan University

Abstract: *In this article, we will show some probability models for artificial intelligence. They are Bayesian and linear regression models. Each model is illustrated by its own example.*

Keywords: *Probability model; AI; Bayesian rule; Bayesian networks; linear regression model.*

Received: 18 March 2020

Accepted for publication: 20 April 2020

Email: tuannv@daihocthudo.edu.vn

1. INTRODUCTION

In recent years mathematicians have been interested in studying the application of probability models to Artificial Intelligence (AI). This is because of the probability of having great applications in AI and life. The main researching areas are generalizations of using probabilities in AI or using probabilities for specific issues of AI. For example, M. Kukacka [2] gave an overview of Bayesian Methods in Artificial Intelligence; Sunghae [6] has studied a probability learning model for constructing artificial minds; Credit risk analysis using machine and deep learning models was introduced by Peter and others [4],…

In this paper, we have applied the probability for AI. This work is built as follow: in Section 2, Bayesian models for AI and their application are presented. The Linear regression models in prediction are established in Section 3. In the last Section, Section 4, conclusion is presented.

2. CONTENT

2.1. Bayesian models for AI

2.1.1. Bayesian rule and applications

1. Bayesian rule

Assume that events $\{H_i\}_{i=1}^{n}$ form a partition of the sample space Ω, i.e.

$$\Omega = \bigcup_{i=1}^{n} H_i,\ H_i \cap H_j = \emptyset, i \neq j.$$

Then an event A in Ω, we have

$$P(H_i|A) = \frac{P(H_i)P(A|H_i)}{\sum_{i=1}^{n} P(H_i)P(A|H_i)}, i = 1, 2, \dots, n.$$

Applying Bayesian rule, we can solve many problems of AI that have applications in real-life.

2. Some examples

Example 1. In a factory machines (I), (II) and (III) are all producing springs of the same length. Of their production, machines (I), (II) and (III) produce 2%, 1% and 3% defective springs, respectively. Of the total production of springs in the factory, machine (I) producces 35%, machine (II) producces 25%, machine (III) producces 40%.

If one of the factory's springs is randomly selected, the possibility of obtaining a spring is made by which machine.

Solution

Call A, B, C as random events to choose the spring made by machines (I), (II) and (III) produce. Then {A, B, C} form a complete system.

Let H be an incident that selects a factory product.

We have $P(A) = 0.35;\ P(A) = 0.35;\ P(A) = 0.35;\ P(H|A) = 0.02;\ P(H|A) = 0.02;\ P(H|A) = 0.02.$

If one spring is selected at random from the total springs, the propability that it is defected equals:

$$P(H) = P(A)P(H|A) + P(B)P(H|B) + P(C)P(H|C) = 0.0218.$$

Applying Baysian rule, we get

$$P(A|H) = \frac{P(A)P(H|A)}{P(H)} = \frac{70}{215};$$

$$P(B|H) = \frac{P(B)P(H|B)}{P(H)} = \frac{25}{215};$$

$$P(C|H) = \frac{P(C)P(H|C)}{P(H)} = \frac{120}{215}.$$

So most likely the spring obtained by the machine (III).

Example 2. The probability of a certain medical test being positive is 90%, if a patient has flu. 1% of the population has the disease, and the test records a false positive 5% of the time. If a man in that region received a positive test, how much is his probability of having flu?

Solution.

Let H be the event of selecting a person who has the flu in that area.

Let A be the positive test event for a person with the flu in that area.

By hypothesis, we have

$$P(H) = 0.01; P(A|H) = 0.90;\ P(A|\bar{H}) = 0.05.$$

Because events $\{H_i\}_{i=1}^{n}$ form a partition of the space Ω, so we get:

$$P(A) = P(H)P(A|H) + P(\bar{H})P(A|\bar{H}) = \ 0.0585.$$

Applying Bayesian rule, we have

$$P(H|A) = \frac{P(H)P(A|H)}{P(A)} = \frac{0.01 \times 0.9}{0.0585} = 0.15$$

So if a man in that region received a positive test, then his probability of having flu is 0.15.

2.2. Bayesian netwwork

Bayes networks have many applications in AI. Here, we present some basis for Bayes networks application in AI.

1. Definition. A Bayesian network (BN) is defined by the following elements [2]:

• A set of nodes, where each node represents a single variable;

• A set of directed connections on these nodes, forming a directed acyclic graph, where a link specifies a dependence relationship between variables;

• A conditional probability table (CPT) for each node in the graph, specifying a probability distribution of the corresponding variable conditioned by its parents in the graph (i.e. $P(X_i|Parents\ (X_i))$).

We can retrieve the probability of any event in a system described by a Bayesian network using the following formula:

$$P(x_1, \dots, x_n) = \prod_{i=1}^{n} (x_i|\text{Parents}(X_i)),$$

where parents(X_i) denotes the specific values of the X_i 's parent variables. This implies that the Bayesian network fully describes the full joint distribution of the system.

2. Some rules of Bayesian networks

Let A, B, C are random variables. We say that A and B independent with given C, if we know C, evidence of B does not change the likelihood of A.

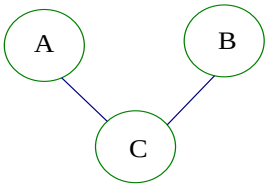
If A and B independent given C then $P(A|B,C) = P(A|C)$. When we have

i) $P(B|A,C) = P(B|C)$;

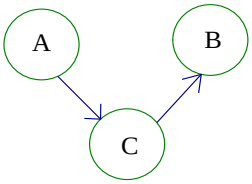
ii) $P(A,B,C) = P(A|C)P(B|C)P(C)$;

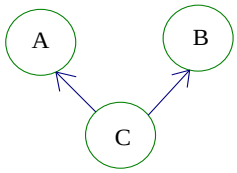
iii) Graphs of A, B and C as following:

+ Undirected graphs



+ Directed acyclic graphs



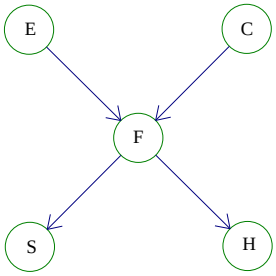


3. Example of causal models

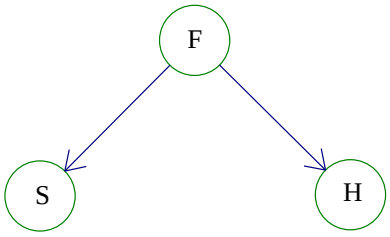
We can use graphical to represent causal models. Let E, C, F, S, H are random variables, where E: electric; C: cigaratte; F: fire; S: smoke; H: heat.

We know that, fire can be caused by an electrical problem or by a cigarette. Smoke and Heat are results of firing.

Smoke and Heat is the consequence of burning. Therefore, we have the following causal model:



Now we consider a problem: Let we have a graphical model:



Each factor is described by a conditional probability table

F	$P(S)$	$P(\overline{S})$
F	0.01	0.99
T	0.9	0.1

F	$P(H)$	$P(\overline{H})$
F	0.001	0.999
T	0.99	0.01

$P(F)$	$P(\overline{F})$
0.1	0.9

Calculate the probability of fire when smoke has occurred.

Solution.

From the hypothesis, we have

$$\begin{cases} P(F) = 0.1 \\ P(\overline{F}) = 0.9, \end{cases} \quad \begin{cases} P(S|\overline{F}) = 0.01 \\ P(\overline{S}|\overline{F}) = 0.99 \\ P(S|F) = 0.9 \\ P(\overline{S}|F) = 0.1, \end{cases} \quad \begin{cases} P(H|\overline{F}) = 0.001 \\ P(\overline{H}|\overline{F}) = 0.999 \\ P(H|F) = 0.99 \\ P(\overline{H}|F) = 0.01. \end{cases}$$

The probability of fire and smoke is

$$P(S,H) = \sum_{H} P(F,S,H) = \sum_{H} P(H|F)P(S|F)P(F)$$

$$= P(H|F)P(S|F)P(F) + P(\overline{H}|F)P(S|F)P(F)$$

$$= 0.9 \times 0.99 \times 0.1 + 0.90 \times 0.01 \times 0.1$$

$$= 0.09.$$

The probability of seeing smoke is

$$P(S) = \sum_{F}\sum_{H} P(F,S,H) = \sum_{F}\sum_{H} P\,(H|F)P(S|F)P(F)$$

$$= P(H|F)P(S|F)P(F) + P(\overline{H}|F)P(S|F)P(F) +$$

$$+ P(H|\overline{F})P(S|\overline{F})P(\overline{F}) + P(\overline{H}|\overline{F})P(S|\overline{F})P(\overline{F}) = 0.99 \times 0.9 \times 0.1 +$$

$$0.01 \times 0.9 \times 0.1 + 0.001 \times 0.01 \times 0.9 + 0.999 \times 0.01 \times 0.9 = 0.099.$$

So the probability of fire if we see smoke it is defined

$$P(F|S) = \frac{P(F,S)}{P(S)} = \frac{0.09}{0.099} = 0.92$$

2.3. Some linear regression models for AI

Linear regression can be applied in AI [6]. In this section we will build linear regression models and get some illustrative examples.

2.3.1. Linear regression models for AI

We observe paired data points $\{(x_i, y_i)\}_{i=1}^n$, where assume that as a function of x_i, each y_i is generated by using some true underlying line $Y = \beta_0 + \beta_1 X$ that is evalute at x_i. Formally,

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i, i = 1, \dots, n. \quad (3.1)$$

We will model ε_i as being Gaussian: $\varepsilon \sim N(0, \sigma^2)$. We find β_0, β_1 by solving the following optimization problem:

$$min_{\beta_0,\beta_1} \sum_{i=1}^{n} [y_i - (\beta_0 + \beta_1 x_i)]^2. \quad (3.2)$$

Given a set of points $\{(x_i, y_i)\}_{i=1}^n$, the solution is [6]:

$$\widehat{\beta_1} = \frac{\overline{XY} - \overline{X}\,\overline{Y}}{\overline{X}^2 - (\overline{X})^2} = r_{XY} \frac{\sqrt{MS_Y}}{\sqrt{MS_X}},$$

$$\widehat{\beta_0} = \overline{Y} - \widehat{\beta_1}\overline{X},$$

where

$$MS_X = \overline{X}^2 - \left(\overline{X}\right)^2, MS_Y = \overline{Y}^2 - \left(\overline{Y}\right)^2, \overline{XY} = \frac{\sum_{i=1}^n x_i y_i}{n}, \overline{X}^2 = \frac{\sum_{i=1}^n {x_i}^2}{n}, \overline{Y}^2 = \frac{\sum_{i=1}^n {y_i}^2}{n}.$$

Example 1. On 100 field plots of the same size, people apply different amounts of fertilizer. Then, study the relationship between fertilizer and yield. The result is given in the table below (X is the amount of fertilizer, Y is the yield).

Y \ X	1	2	3	4	5
14	10	8			
15		12	7		
16			28	6	
17				8	9
18					12

a) If the amount of manure is 3.44, what is the estimate yield?

b) Estimate yield if the amount of fertilizer is 5.65.

Solution. We have

$$\sum_{i=1}^{100} x_i = 1 \times 10 + (8+12) \times 2 + (7+28) \times 3 + (6+8) \times 4 + (9+12) \times 5 = 316.$$

$$\overline{X} = \frac{\sum_{i=1}^{100} x_i}{n} = \frac{316}{100} = 3.16.$$

$$\sum_{i=1}^{100} y_i = (1+8) \times 14 + (12+7) \times 15 + (6+28) \times 16 + (8+9) \times 17 + 12 \times 8 = 1586.$$

$$\overline{y} = \frac{\sum_{i=1}^{100} y_i}{n} = \frac{1586}{100} = 15.86.$$

$$\sum_{i=1}^{100} {x_i}^2 = 1154; \sum_{i=1}^{100} {y_i}^2 = 25308; \sum_{i=1}^{100} x_i y_i = 5156;$$

$$100 \sum_{i=1}^{100} x_i y_i - \sum_{i=1}^{100} x_i \left(\sum_{i=1}^{100} y_i \right) = 14424;$$

$$\sqrt{MS_X} = \sqrt{100 \sum_{i=1}^{100} {x_i}^2 - \left(\sum_{i=1}^{100} x_i \right)^2} = 124.67;$$

$$\sqrt{MS_Y} = \sqrt{100 \sum_{i=1}^{100} {y_i}^2 - \left(\sum_{i=1}^{100} y_i \right)^2} = 124.11;$$

$$S\,Ir_{XY} = \frac{14424}{124.67 \times 124.11} = 0.93.$$

So we get

$$\hat{Y} = \widehat{\beta_0} + \widehat{\beta_1} X \Rightarrow \hat{Y} = 12.93 + 0.93X.$$

a) If $X = 3.44$ then $\hat{Y} = 16.13$;

b) If $X = 5.65$ then $\hat{Y} = 18.19$.

Example 2. The same pill was given to 5 patients of different ages. Study the time to completely disintegrate the drug in each person's body. Specific results are as follows:

X: age (year)	Y: Decomposition time (minute)
30	15
25	28
65	30

50	22
40	24

a) What conclusions can be drawn about the relationship between the time of drug decomposition and ages.

b) What is the time predict for drug decomposition by 42 years old?

Solution.

We have

$$\sum_{i=1}^{5} x_i = 210;\ \sum_{i=1}^{5} x_i = 119;\ \sum_{i=1}^{5} x_i\, y_i = 5160;\ \sum_{i=1}^{5} {x_i}^2 = 9850.$$

$$\widehat{\beta_1} = \frac{n\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} {x_i}^2 - (\sum_{i=1}^{n} x_i)^2} = \frac{5 \times 5160 \ - \ 210 \times 119}{5.9850 \ - \ 210^2} = 0.16;$$

$$\widehat{\beta_0} = \frac{\sum_{i=1}^{n} y_i - \widehat{\beta_1}\sum_{i=1}^{n} x_i}{n} = \frac{5 \times 5160 \ - \ 210 \times 119}{5.9850 \ - \ 210^2} = 17.08.$$

So we get

$$\hat{Y} = 17.08 + 0.16X. \tag{3.3}$$

By (3.3) we say that

a) Every 10 years of age, the decomposition time will increase by 1.6 minutes.

b) The time for decomposition of 42-year-old patient is 23.8 minutes.

2.3.2. Multiple linear regressions

Let $Y, X_2, \dots, X_p$ are p radom variables, where $p \in \boldsymbol{N}^*,\ p > 2$. We find linear model of Y for $X_2, \dots, X_p$ as following:

$$f_Y(X_2, \dots, X_p) = \beta_1 + \beta_2 X_2 + \cdots + \beta_p X_p,$$

which observation matrix

$$\begin{pmatrix} Y_1 & X_{12} & X_{13} \dots & X_{1p} \\ Y_2 & X_{22} & X_{23} \dots & X_{2p} \\ \dots & \dots & \dots \dots & \dots \\ Y_n & X_{n2} & X_{n3} \cdots & X_{np} \end{pmatrix}.$$

We get regression model

$$Y_i = \beta_1 + \beta_2 X_{i2} + \cdots + \beta_p X_{ip} + \varepsilon_i, \tag{3.4}$$

where $\varepsilon \sim N(0, \sigma^2)$.

Let

$$Y = [Y_1, \dots, Y_n]^T;\ X = \begin{pmatrix} 1 & X_{12} & X_{13} \dots & X_{1p} \\ 1 & X_{22} & X_{23} \dots & X_{2p} \\ \dots & \dots & \dots & \dots \\ 1 & X_{n2} & X_{n3} \cdots & X_{np} \end{pmatrix};\ X_j = [X_{1j}, \dots, X_{nj}]^T, j = 2, \dots, p; \quad (3.5)$$

$$\beta = [\beta_1, \dots, \beta_p]^T;\ \varepsilon = [\varepsilon_1, \dots, \varepsilon_n]^T.$$

$$Y = [Y_1, \dots, Y_n]^T;\ X = \begin{pmatrix} 1 & X_{12} & X_{13} \dots & X_{1p} \\ 1 & X_{22} & X_{23} \dots & X_{2p} \\ \dots & \dots & \dots & \dots \\ 1 & X_{n2} & X_{n3} \cdots & X_{np} \end{pmatrix};\ X_j = [X_{1j}, \dots, X_{nj}]^T, j = 2, \dots, p; \quad (3.6)$$

Using notations (3.5) and (3.6), we can write model (3.4) as following:

$$Y = X\beta + \varepsilon.$$

To determine β we have to solve the optimization problem

$$min_\beta \sum_{i=1}^{n} (y_i - X_i\beta)^2 \quad (3.7)$$

We can use some basic linear algebra to solve this problem and find the result:

$$\hat{\beta} = (X^T X)^{-1} X^T Y. \quad (3.8)$$

Example 3. A Corporation has 21 stores in provinces and cities. Let Y be the revenue of the Corporations (10^3USD), X_1 is the population (thousand), X_2 is the average income per person (10^3 USD) in the provinces and cities where the business is located. Assume the following table of business data is available:

Provinces and cities	X_1 (Thousand)	X_2 (10^3USD)	Y (10^3USD)
1.	68.5	16.7	174.4
2.	45.2	16.8	164.4
3.	91.3	18.2	224.2
4.	47.8	16.3	154.6
5.	46.9	17.3	181.6
6.	66.1	10.2	207.5
7.	49.5	15.9	152.8
8.	52.0	17.2	163.2
9.	48.9	16.6	145.4
10.	38.4	16.0	137.2
11.	87.9	18.3	241.9
12.	72.8	17.1	191.1
13.	88.4	17.4	232.0

14.	42.9	15.8	145.3
15.	52.5	17.8	161.1
16.	85.7	18.4	209.7
17.	43.1	16.5	146.4
18.	51.7	16.3	144.0
19.	89.6	18.1	232.6
20.	82.7	19.1	224.1
21.	52.3	16.0	166.5

Predict the revenue of this Corporation.

Solution.

Use model (3.4) which $p = 3$ to solve this problem.

We have

$$X = \begin{pmatrix} 1 & 68.5 & 16.7 \\ 1 & 42.2 & 16.8 \\ \dots & \dots & \dots \\ 1 & 52.3 & 16.0 \end{pmatrix}; Y = \begin{pmatrix} 174.4 \\ 164.4 \\ \dots \\ 166.5 \end{pmatrix}.$$

After calculating we obtained

$$X^T X = \begin{pmatrix} 21.0 & 1302.4 & 360.0 \\ 1302.4 & 87707.9 & 22609.2 \\ 360 & 22609.2 & 6190.3 \end{pmatrix}, X^T Y = \begin{pmatrix} 3820 \\ 249643 \\ 66073 \end{pmatrix},$$

$$(X^T X)^{-1} = \begin{pmatrix} 29.5740 & 0.0718 & -1.9820 \\ 0.0718 & 0.00037 & -0.0055 \\ -1.9820 & -0.0055 & 0.1356 \end{pmatrix}$$

And

$$\hat{\beta} = (X^T X)^{-1} X^T Y = \begin{pmatrix} -68.609 \\ 1.455 \\ 9.488 \end{pmatrix}.$$

So we have the result

$$\widehat{Y} = -68.609 + 1.455X_1 + 9.488X_2. \quad (3.9)$$

From (3.9) we can predict the results:

a) If the average income is constant and the population increases by one thousand people, the sale turnover of the Corporation increases by 1455 USD;

b) If the population is constant and the income per capita increases by one thousand dollars, the sale turnover increases by 9488 USD.

3. CONCLUSION

In this article, some probabilistic models for AI are presented. Using Bayesian models and linear regression models we could make predictions in medicine as well as production and business. At the same time given models claim effective applications of AI technology.

REFERENCES

1. Christopher M. Bishop, *Bayesian Methods for Neural Networks*, Technical Report: NCRG/95/009.
2. M. Kukacka (2010), *Bayesian Methods in Artificial Intelligence*, WDS'10 Proceedings of Contributed Papers, Part I, 25–30, 2010.
3. P. Larranaga, S. Moral (2011), *Probabilistic graphical models in artificial intelligence*, Applied Soft Computing 11 (2011) 1511–1528.
4. Peter Martey Addo, Dominique Guegan and Bertrand Hassani (2018), *Credit Risk Analysis Using Machine and Deep Learning Models*, Risks 2018, 6, 38.
5. Robert J. Sternberg (2005), *The Theory of Successful Intelligence*, Revista Interamericana de Psicología/Interamerican Journal of Psychology, 2005, Vol. 39, Num. 2 pp. 189-202.
6. Sunghae Jun (2017), *A Probability Learning Model for Constructing Artificial Minds*, International Journal of Applied Engineering, Vol. 12, Num. 12, pp. 3542-3545.
7. Nguyen Cao Van, Tran Thai Ninh (2005), *Text book Probability theory and mathematical statistics*, Statistical Publisher.

MỘT SỐ MÔ HÌNH XÁC SUẤT CHO TRÍ TUỆ NHÂN TẠO

Tóm tắt: *Trong bài báo này chúng ta sẽ trình bày một vài mô hình xác suất được áp dụng trong AI. Đó là các mô hình Bayes và hồi quy tuyến tính. Các ví dụ cũng được đưa ra để minh họa cho từng mô hình.*

Từ khóa: *Mô hình xác suất, trí tuệ nhân tạo, công thức Bayes, mạng Bayes, mô hình hồi quy tuyến tính.*

GAUGE SECTOR OF THE MINIMAL 3-3-1-1 MODEL

Dang Quoc Phong

Hanoi Architectural University

Abstract: *In this work, we consider the gauge boson sector of the minimal 3-3-1-1 model. We identify the standard model gauge bosons and the new gauge particles. We prove that the number of the massive gauge bosons match those of the Goldstone bosons, leaving only photon and gluon massless. The dark matter vector is discussed.*

PACS numbers: 12.60.-i

Keywords: *Dark matter, 3-3-1-1 model.*

Received: 22 March 2020

Accepted for publication: 20 April 2020

Email: tphongvltckt@yahoo.com

1. INTRODUCTION

The standard model cannot explain the neutrino masses and the dark matter abundance of the universe [1]. The new approaches based on the 3-3-1-1 symmetry not only address such questions but also yield interested novel consequences [2, 3].

Yet, the gauge sector of the minimal 3-3-1-1 model was not investigated. In this work, we diagonalize the relevant gauge sector and prove that it contains appropriate gauge spectrum of the standard model. Additionally, the new gauge bosons are obtained, attracting attention at the current colliders.

In Sec. II, we introduce the minimal 3-3-1-1 model. In Sec. III, we diagonalize the gauge sector, identifying relevant mass spectrum. We conclude this work in Sec. IV.

2. CONTENT

2.1. Minimal 3-3-1-1 model

The gauge symmetry is given by,

$$SU(3)_C \otimes SU(3)_L \otimes U(1)_X \otimes U(1)_N, \quad (1)$$

which is an extension of the 3-3-1 model [6–8] by supposing that $B - L$ closes algebraically with $SU(3)_L$ [2, 3]. The electric charge is obtained as

$$Q = T_3 - \sqrt{3}T_8 + X, \tag{2}$$

corresponding to the minimal fermion content of the minimal 3-3-1 model [6]. For the minimal fermion content, the baryon minus lepton number is embedded as

$$\boldsymbol{B - L = -\left(4/\sqrt{3}\right)T_8 + N} \tag{3}$$

After symmetry breaking, a matter parity arises to be

$$\mathrm{P} = (-1)^{3(\mathrm{B}-\mathrm{L})+2\mathbf{s}} \tag{4}$$

Under the 3-3-1-1 symmetry, the mentioned fermion representations are

$$\psi_{aL} \equiv \begin{pmatrix} \nu_{aL} \\ e_{aL} \\ e^c_{aR} \end{pmatrix} \sim (1,3,0,-1/3), \quad \nu_{aR} \sim (1,1,0,-1), \tag{5}$$

$$Q_{\alpha L} \equiv \begin{pmatrix} d_{\alpha L} \\ -u_{\alpha L} \\ J_{\alpha L} \end{pmatrix} \sim (3,3^*,-1/3,-1/3), \quad Q_{3L} \equiv \begin{pmatrix} u_{3L} \\ d_{3L} \\ J_{3L} \end{pmatrix} \sim (3,3,2/3,1), \tag{6}$$

$$u_{aR} \sim (3,1,2/3,1/3), \quad d_{aR} \sim (3,1,-1/3,1/3), \tag{7}$$

$$J_{\alpha R} \sim (3,1,-4/3,-5/3), J_{3R} \sim (3,1,5/3,7/3), \tag{8}$$

which is sufficient for the anomaly cancellation [4, 5].

To break the 3-3-1-1 symmetry, we introduce the following scalars,

$$\eta = \begin{pmatrix} \eta_1^0 \\ \eta_2^- \\ \eta_3^+ \end{pmatrix} \sim (1,3,0,2/3), \quad \rho = \begin{pmatrix} \rho_1^+ \\ \rho_2^0 \\ \rho_3^{++} \end{pmatrix} \sim (1,3,1,2/3), \tag{9}$$

$$\chi = \begin{pmatrix} \chi_1^- \\ \chi_2^{--} \\ \chi_3^0 \end{pmatrix} \sim (1,3,-1,-4/3), \quad \phi \sim (1,1,0,2), \tag{10}$$

with vacuum expectative values,

$$\langle \eta \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} u \\ 0 \\ 0 \end{pmatrix}, \quad \langle \rho \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ v \\ 0 \end{pmatrix}, \quad \langle \chi \rangle = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 0 \\ w \end{pmatrix}, \quad \langle \phi \rangle = \frac{1}{\sqrt{2}} \Lambda \tag{11}$$

The total Lagrangian is given by

$$\begin{aligned}\mathcal{L} &= \sum_F \bar{F} i\gamma^\mu D_\mu F + \sum_S (D^\mu S)^\dagger (D_\mu S) - V + \mathcal{L}_Y \\ &\quad -\frac{1}{4}(G_{i\mu\nu}G_i^{\mu\nu} + A_{i\mu\nu}A_i^{\mu\nu} + B_{\mu\nu}B^{\mu\nu} + C_{\mu\nu}C^{\mu\nu})\end{aligned} \tag{12}$$

where F and S indicate to the fermion multiplets and scalar multiplets, respectively. The covariant derivative takes the form,

$$D_\mu = \partial_\mu + ig_s t_i G_{i\mu} + igT_i A_{i\mu} + ig_X X B_\mu + ig_N N C_\mu, \tag{13}$$

where (t_i, T_i, X, N), (g_s, g, g_X, g_N), and (G_i, A_i, B, C) denote the generators, gauge couplingconstants, and gauge bosons of the 3-3-1-1 groups, respectively. The field strength tensors are

$$Gi\mu v = \partial\mu Giv - \partial vGi\mu - gsfijkGj\mu Gkv, \tag{14}$$

$$Ai\mu v = \partial\mu Aiv - \partial vAi\mu - gfijkAj\mu Akv, \tag{15}$$

$$B\mu v = \partial\mu Bv - \partial vB\mu, \qquad C\mu v = \partial\mu Cv - \partial vC\mu, \tag{16}$$

where f_{ijk} is the fine structure constant of SU(3) group.

Since the Yukawa Lagrangian and scalar potential are not investigated in this work, they are being skipped.

2.2. Gauge spectrum

In this section, let us consider the gauge boson spectrum. The gauge bosons obtain masses when the scalar fields develop VEVs. Therefore, the mass Lagrangian of the gauge bosons is given by

$$\boldsymbol{\mathcal{L}_{mass}^{gause}} = \sum_{\boldsymbol{S=\eta,\rho,\chi,\phi}} (\boldsymbol{D_\mu\langle S\rangle}) + (\boldsymbol{D^\mu\langle S\rangle} \tag{17}$$

Where

$$\begin{aligned}D_\mu\langle\eta\rangle &= (\partial_\mu + ig_s G_{i\mu}t_i + igA_{i\mu}T_i + ig_X B_\mu X + ig_N C_\mu N)\langle\eta\rangle \\ &= \frac{igu}{2\sqrt{2}}\begin{pmatrix} A_{3\mu} + \frac{1}{\sqrt{3}}A_{8\mu} + \frac{4}{3}t_N C_\mu \\ \sqrt{2}W_\mu^- \\ \sqrt{2}X_\mu^+ \end{pmatrix},\end{aligned} \tag{18}$$

Where $\boldsymbol{t_N = \frac{g_N}{g}}$ (19)

And

$$W_\mu^\pm = \frac{A_{1\mu} \mp iA_{2\mu}}{\sqrt{2}}, \quad X_\mu^\pm = \frac{A_{4\mu} \pm iA_{5\mu}}{\sqrt{2}}, \quad Y_\mu^{\pm\pm} = \frac{A_{6\mu} \pm iA_{7\mu}}{\sqrt{2}}$$

$$D_\mu\langle\rho\rangle = \frac{igv}{2\sqrt{2}} \begin{pmatrix} \sqrt{2}W_\mu^+ \\ -A_{3\mu} + \frac{1}{\sqrt{3}}A_{8\mu} + 2t_X B_\mu + \frac{4}{3}t_N C_\mu \\ \sqrt{2}Y_\mu^{++} \end{pmatrix}, \tag{20}$$

in which $\boldsymbol{t_X = \frac{g_X}{g}}$

$$D_\mu\langle\chi\rangle = \frac{igw}{2\sqrt{2}} \begin{pmatrix} \sqrt{2}X_\mu^- \\ \sqrt{2}Y_\mu^{--} \\ -\frac{2}{\sqrt{3}}A_{8\mu} - 2t_X B_\mu - \frac{8}{3}t_N C_\mu \end{pmatrix} \tag{21}$$

$$D_\mu\langle\phi\rangle = \sqrt{2}igt_N C_\mu \Lambda. \tag{22}$$

Combining (17), (18), (20), (21) and (22), we have

$$\begin{aligned} \mathcal{L}_{\text{mass}}^{\text{gause}} &= \frac{g^2u^2}{8}[2W_\mu^+W^{\mu-} + 2X_\mu^+X^{\mu-} + (A_{3\mu} + \frac{1}{\sqrt{3}}A_{8\mu} + \frac{4}{3}t_N C_\mu)^2] \\ &+ \frac{g^2v^2}{8}[2W_\mu^+W^{\mu-} + 2Y_\mu^{++}Y^{\mu--} + (-A_{3\mu} + \frac{1}{\sqrt{3}}A_{8\mu} + 2t_X B_\mu + \frac{4}{3}t_N C_\mu)^2] \\ &+ \frac{g^2w^2}{8}[2X_\mu^+X^{\mu-} + 2Y_\mu^{++}Y^{\mu--} + (-\frac{2}{\sqrt{3}}A_{8\mu} - 2t_X B_\mu - \frac{8}{3}t_N C_\mu)^2] \\ &+ 2g^2t_N^2\Lambda^2 C_\mu C^\mu \\ &= \mathcal{L}_{mass}^{changed} + \mathcal{L}_{mass}^{neutral} \end{aligned} \tag{23}$$

The mass Lagrangian of the charged scalars is written as

$$\begin{aligned} \mathcal{L}_{\text{mass}}^{\text{gause}} &= \frac{g^2u^2}{8}(2W_\mu^+W^{\mu-} + 2X_\mu^+X^{\mu-}) \\ &+ \frac{g^2v^2}{8}(2W_\mu^+W^{\mu-} + 2Y_\mu^{++}Y^{\mu--}) \\ &+ \frac{g^2w^2}{8}(2X_\mu^+X^{\mu-} + 2Y_\mu^{++}Y^{\mu--}) \\ &= \frac{g^2}{4}(u^2+v^2)W_\mu^+W^{\mu-} + \frac{g^2}{4}(u^2+w^2)X_\mu^+X^{\mu-} + \frac{g^2}{4}(v^2+w^2)Y_\mu^{++}Y^{\mu--} \end{aligned} \tag{24}$$

The gauge bosons $W^\pm$, $X^\pm$ and $Y^{\mp\mp}$ by themselves are physical fields with the respective masses,

$$m_W^2 = \frac{g^2}{4}(u^2+v^2) \quad m_X^2 = \frac{g^2}{4}(u^2+w^2), \quad m_Y^2 = \frac{g^2}{4}(v^2+w^2) \tag{25}$$

Note that the mass of W boson implies $u^2+v^2=v_w^2=$ (246 GeV)^2. Due to the conditions, $w\gg u\sim v$, we have $m_W\ll m_X\simeq m_Y\simeq$ TeV. The X and Y are new gauge bosons with the large masses given in w scale.

The mass Lagrangian of the neutral scalar fields is written as

$$\begin{aligned}\mathcal{L}_{\rm mass}^{\rm neutral} &= (u^2+v^2)\frac{g^2}{8}A_{3\mu}^2+(u^2+v^2+4w^2)\frac{g^2}{24}A_{8\mu}^2\\ &+(v^2+w^2)\frac{g^2}{2}t_X^2B_\mu^2+(u^2+v^2+4w^2+9\Lambda^2)\frac{2g^2}{9}t_N^2C_\mu^2\\ &+(u^2-v^2)\frac{g^2}{4\sqrt{3}}A_{3\mu}A_{8\mu}-v^2\frac{g^2}{2}t_XA_{3\mu}B_\mu+(u^2-v^2)\frac{g^2}{3}t_NA_{3\mu}C_\mu\\ &+(v^2+2w^2)\frac{g^2}{2\sqrt{3}}t_XA_{8\mu}B_\mu+(u^2+v^2+4w^2)\frac{g^2}{3\sqrt{3}}t_NA_{8\mu}C_\mu\\ &+(v^2+2w^2)\frac{2g^2}{3}t_Xt_NB_\mu C_\mu\\ &=\frac{1}{2}\begin{pmatrix}A_{3\mu} & A_{8\mu} & B_\mu & C_\mu\end{pmatrix}M^2\begin{pmatrix}A_3^\mu\\ A_8^\mu\\ B^\mu\\ C^\mu\end{pmatrix}.\end{aligned}\tag{26}$$

The squared mass matrix of neutral gauge bosons is found to be,

$$M^2=\begin{pmatrix}(u^2+v^2)\frac{g^2}{4} & (u^2-v^2)\frac{g^2}{4\sqrt{3}} & -v^2\frac{g^2}{2}t_X & (u^2-v^2)\frac{g^2}{3}t_N\\ (u^2-v^2)\frac{g^2}{4\sqrt{3}} & (u^2+v^2+4w^2)\frac{g^2}{12} & (v^2+2w^2)\frac{g^2}{2\sqrt{3}}t_X & (u^2+v^2+4w^2)\frac{g^2}{3\sqrt{3}}t_N\\ -v^2\frac{g^2}{2}t_X & (v^2+2w^2)\frac{g^2}{2\sqrt{3}}t_X & (v^2+w^2)g^2t_X^2 & (v^2+2w^2)\frac{2g^2}{3}t_Xt_N\\ (u^2-v^2)\frac{g^2}{3}t_N & (u^2+v^2+4w^2)\frac{g^2}{3\sqrt{3}}t_N & (v^2+2w^2)\frac{2g^2}{3}t_Xt_N & (u^2+v^2+4w^2+9\Lambda^2)\frac{4g^2}{9}t_N^2\end{pmatrix}.\tag{27}$$

The neutral gauge bosons ($A_{3\mu}$,$A_{8\mu}$,B_μ,C_μ) mix via the mass matrix M^2. It is easily checked that M^2 has a zero eigenvalue (the photon mass) and respective eigenstate independent of VEVs as follows,

$$m_A^2=0,\quad A_\mu=\frac{t_X}{\sqrt{1+4t_X^2}}A_{3\mu}-\frac{\sqrt{3}t_X}{\sqrt{1+4t_X^2}}A_{8\mu}+\frac{1}{\sqrt{1+4t_X^2}}B_\mu\tag{28}$$

q where, $t_X=g_X/g=s_W/\sqrt{1-4s_W^2}$, with $s_W=e/g$ is the sine of the Weinberg angle. The Eq (28) can be rewritten as

$$m_A^2=0,\quad A_\mu=\frac{t_X}{\sqrt{1+4t_X^2}}A_{3\mu}-\frac{\sqrt{3}t_X}{\sqrt{1+4t_X^2}}A_{8\mu}+\frac{1}{\sqrt{1+4t_X^2}}B_\mu\tag{28}$$

$$A_\mu=s_WA_{3\mu}+c_W(-\sqrt{3}t_WA_{8\mu}+\sqrt{1-3t_W^2}B_\mu\tag{29}$$

while the terms in the parentheses are the weak hypercharge, $\boldsymbol{Y = -\sqrt{3}t_W A_{8\mu} + \sqrt{1-3t_W^2}B_\mu}$ *(or* $\boldsymbol{Y = -\sqrt{3}T_8 + X}$*)*

The remaining states of the boson in the standard model Z and the new bosons Z' and C are

$$Z_\mu = c_W A_{3\mu} - s_W(-\sqrt{3}t_W A_{8\mu} + \sqrt{1-3t_W^2}B_\mu \tag{30}$$

$$Z'_\mu = \sqrt{1-3t_W^2}A_{8\mu} + \sqrt{3}t_W B_\mu), \tag{31}$$

$$Z_N = C_\mu. \tag{32}$$

The mass matrix M^2 can be diagonalized via several steps. In the first step, we change the basis to $(A_{3\mu},A_{8\mu},B_\mu,C_\mu) \rightarrow (A,Z,Z^0,C)$,

$$\begin{pmatrix} A_{3\mu} \\ A_{8\mu} \\ B_\mu \\ C_\mu \end{pmatrix} = U_1 \begin{pmatrix} A \\ Z \\ Z' \\ C \end{pmatrix}, \quad U_1 = \begin{pmatrix} s_W & c_W & 0 & 0 \\ -\sqrt{3}s_W & \sqrt{3}s_W t_W & \sqrt{1-3t_W^2} & 0 \\ c_W\sqrt{1-3t_W^2} & -s_W\sqrt{1-3t_W^2} & \sqrt{3}t_W & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{33}$$

In this new basis, the mass matrix M^2 becomes

$$M'^2 = U_1^T M^2 U_1 = \begin{pmatrix} 0 & 0 \\ 0 & M_s'^2 \end{pmatrix} \tag{34}$$

Where

$$M_s'^2 = U_1^T M^2 U_1 = \begin{pmatrix} m_Z^2 & m_{ZZ'}^2 & m_{ZC}^2 \\ m_{ZZ'}^2 & m_{Z'}^2 & m_{Z'C}^2 \\ m_{ZC}^2 & m_{Z'C}^2 & m_C^2 \end{pmatrix} = g^2 \times$$

$$\begin{pmatrix} \frac{u^2+v^2}{4c_W^2} & \frac{(1-4s_W^2)u^2-(1+2s_W^2)v^2}{4\sqrt{3}c_W^2\sqrt{1-4s_W^2}} & \frac{(u^2-v^2)t_N}{3c_W} \\ \frac{(1-4s_W^2)u^2-(1+2s_W^2)v^2}{4\sqrt{3}c_W^2\sqrt{1-4s_W^2}} & \frac{(1-4s_W^2)^2u^2+(1+2s_W^2)^2v^2+4c_W^4w^2}{12(1-4s_W^2)c_W^2} & \frac{t_N[(1-4s_W^2)u^2+(1+2s_W^2)v^2+4c_W^2w^2]}{3\sqrt{3}c_W\sqrt{1-4s_W^2}} \\ \frac{(u^2-v^2)t_N}{3c_W} & \frac{t_N[(1-4s_W^2)u^2+(1+2s_W^2)v^2+4c_W^2w^2]}{3\sqrt{3}c_W\sqrt{1-4s_W^2}} & \frac{4}{9}(u^2+v^2+4w^2+9\Lambda^2)t_N^2 \end{pmatrix} \tag{35}$$

In the limit, $u, v \ll \Lambda, w$, we have $m_Z^2, m_{ZZ'}^2, m_{ZC}^2 \ll m_{Z'}^2, m_{Z'C}^2, m_C^2$. Hence, in the second step, the mass matrix M^{02} can be diagonalized by using seesaw formula to separate the light state (Z) from the heavy states (Z^0,C). We denote the new basis as (A, Z_1, Z_1', C_1) so that A,Z_1 are physical fields and decoupled while the rest mix,

$$\begin{pmatrix} A \\ Z \\ Z' \\ C \end{pmatrix} = U_2 \begin{pmatrix} A \\ Z_1 \\ Z_1' \\ C_1 \end{pmatrix}, \quad M''^2 = U_2^T M'^2 U_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & m_{Z_1}^2 & 0 \\ 0 & 0 & M_s''^2 \end{pmatrix}, \tag{36}$$

Where

$$M_s''^2 = \begin{pmatrix} m_{Z'}^2 & m_{Z'C}^2 \\ m_{Z'C}^2 & m_C^2 \end{pmatrix}, \quad m_{Z_1}^2 \simeq m_Z^2 - \mathcal{E} \begin{pmatrix} m_{ZZ'}^2 \\ m_{ZC}^2 \end{pmatrix} \tag{37}$$

The m_{Z1} is the mass of Z_1 light state, while $M_s''^2$ is the mass sub-matrix of Z_1', C_1 heavy states. By the virtue of seesaw approximation, we have

$$U_2 \simeq \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \mathcal{E} \\ 0 & -\mathcal{E}^T & 1 \end{pmatrix}, \quad \mathcal{E} \equiv \begin{pmatrix} m_{ZZ'}^2 & m_{ZC}^2 \end{pmatrix} \begin{pmatrix} m_{Z'}^2 & m_{Z'C}^2 \\ m_{Z'C}^2 & m_C^2 \end{pmatrix}^{-1} \tag{38}$$

The E is a two-component vector given by

$$\mathcal{E}_1 \simeq -\frac{\sqrt{1-4s_W^2}\{4s_W^2 w^2(u^2+v^2) + 3\Lambda^2[-u^2+v^2+2(2u^2+v^2)s_W^2]\}}{4\sqrt{3}(1-s_W^2)^2 w^2 \Lambda^2} \ll 1 \tag{39}$$

$$\mathcal{E}_2 \simeq \frac{s_W^2(u^2+v^2)}{4(1-s_W^2)^{3/2} t_N \Lambda^2} \ll 1, \tag{40}$$

which are suppressed at the leading order $u, v \ll \Lambda, w$. When neglecting the mixing of Z,Z^0 and C, m_{Z1} ' m_Z, we can identify $Z_1 \equiv Z, Z_1' \equiv Z'$ and $C_1 \equiv C$. For the final step, it is easily to diagonalize M''^2 (or $M_s''^2$) to obtain two remaining physical states, denoted by Z_2 and Z_N, such that

$$\begin{pmatrix} A \\ Z_1 \\ Z_1' \\ C_1 \end{pmatrix} = U_3 \begin{pmatrix} A \\ Z_1 \\ Z_2 \\ Z_N \end{pmatrix}, \quad U_3 \simeq \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & c_\xi & -s_\xi \\ 0 & 0 & s_\xi & c_\xi \end{pmatrix}, \tag{41}$$

$$M'''^2 = U_3^T M''^2 U_3 = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & m_{Z_1}^2 & 0 & 0 \\ 0 & 0 & m_{Z_2}^2 & 0 \\ 0 & 0 & 0 & m_{Z_N}^2 \end{pmatrix}. \tag{42}$$

The mixing angle and new masses are given by

$$t_{2\xi} = \frac{-2m_{Z'C}^2}{m_C^2 - m_{Z'}^2} \simeq -\frac{8}{\sqrt{3}} \frac{\sqrt{1-4s_W^2} c_W t_N w^2}{12(1-4s_W^2)t_N^2 \Lambda^2 + \left[\frac{16}{3}(1-4s_W^2)t_N^2 - c_W^2\right] w^2} \tag{43}$$

$$m^2_{Z_2} \simeq \frac{g^2}{2}\left(4(\Lambda^2+\frac{4}{9}w^2)t^2_N+\frac{w^2c^2_W}{3(1-4s^2_W)} -\sqrt{\left[4(\Lambda^2+\frac{4}{9}w^2)t^2_N-\frac{w^2c^2_W}{3(1-4s^2_W)}\right]^2+\frac{64}{27}\frac{c^2_Wt^2_Nw^4}{1-4s^2_W}}\right), \quad (44)$$

$$m^2_{Z_N} \simeq \frac{g^2}{2}\left(4(\Lambda^2+\frac{4}{9}w^2)t^2_N+\frac{w^2c^2_W}{3(1-4s^2_W)} +\sqrt{\left[4(\Lambda^2+\frac{4}{9}w^2)t^2_N-\frac{w^2c^2_W}{3(1-4s^2_W)}\right]^2+\frac{64}{27}\frac{c^2_Wt^2_Nw^4}{1-4s^2_W}}\right). \quad (45)$$

The Z_2 and Z_N are heavy particles with the masses in w scale.

The physical fields are related to the gauge states as

$$\begin{pmatrix} A_{3\mu} \\ A_{8\mu} \\ B_\mu \\ C_\mu \end{pmatrix} = U \begin{pmatrix} A \\ Z_1 \\ Z_2 \\ Z_N \end{pmatrix}, \quad U = U_1U_2U_3, \quad (46)$$

Where

$$U = \begin{pmatrix} s_W & c_W & c_W(\mathcal{E}_1c_\xi+\mathcal{E}_2s_\xi) & c_W(\mathcal{E}_2c_\xi-\mathcal{E}_1s_\xi) \\ -\sqrt{3}s_W & \sqrt{3}s_Wt_W-\mathcal{E}_1\sqrt{1-3t^2_W} & \sqrt{3}(\mathcal{E}_1c_\xi+\mathcal{E}_2s_\xi)s_Wt_W+c_\xi\sqrt{1-3t^2_W} & \sqrt{3}(\mathcal{E}_2c_\xi-\mathcal{E}_1s_\xi)s_Wt_W-s_\xi\sqrt{1-3t^2_W} \\ c_W\sqrt{1-3t^2_W} & -\sqrt{3}\mathcal{E}_1t_W-s_W\sqrt{1-3t^2_W} & \sqrt{3}c_\xi t_W-(\mathcal{E}_1c_\xi+\mathcal{E}_2s_\xi)s_W\sqrt{1-3t^2_W} & -\sqrt{3}s_\xi t_W-(\mathcal{E}_2c_\xi-\mathcal{E}_1s_\xi)s_W\sqrt{1-3t^2_W} \\ 0 & -\mathcal{E}_2 & s_\xi & c_\xi \end{pmatrix}.$$

In the limit, $\{u^2,v^2\}/\{w^2,\Lambda^2\} \ll 1$ and means that the standard model Z boson by itself is a physical field Z ' Z_1 and do not mix with the new gauge bosons $Z_{2,N}$.

The ρ-parameter (or $\Delta\rho \equiv \rho-1$ used below) that is due to the contribution of the new physics comes from the tree-level mixing of Z with Z^0 and C, which can be evaluated as [?]

$$(\Delta\rho)_{tree} = \frac{m^2_W}{c^2_Wm^2_{Z_1}}-1=\frac{m^2_Z}{m^2_Z-\mathcal{E}(m^2_{ZZ'},\,m^2_{ZC})^T}-1\simeq\frac{\mathcal{E}(m^2_{ZZ'},\,m^2_{ZC})^T}{m^2_Z} \quad (47)$$

Here, notice that $m_W = c_W m_Z$ and $m^2_Z \sim m^2_{ZZ}0 \sim m^2_{ZC}$. We get the deviation as

$$(\Delta\rho)_{tree} \simeq \frac{[(1-4s^2_W)u^2-(1+2s^2_W)v^2]^2}{4c^4_Wv^2_ww^2}+\frac{s^4_Wv^2_w}{c^4_W\Lambda^2} \quad (48)$$

We study the two (u,w) parameters in two cases $\Lambda \gg w$ and $\Lambda \sim$ w, both cases are being limited by the experimental Δρ parameters $0.00016 \le \Delta\rho \le 0.00064$. The case $\Lambda \gg w$ yields results similar to the 3-3-1 model [6]. In the case $\Lambda \sim$ w, we choose Λ = 2w, to the new physics scale w > 1.4 TeV.

3. CONCLUSION

We have found that the W,Z_1,A fields are the standard model gauge bosons. The fields Z_2,Z_N and X,Y are new gauge bosons. There are 9 massive gauge bosons matching 9

Goldstone bosons in the scalar sector. Constraint from the rho parameter yields that $\Lambda \sim w \sim$ TeV. Therefore, the new gauge bosons such as Z_2,Z_N and X,Y can be investigated by the current colliders.

Acknowledgments

I would like to thank Hanoi Architectural University for supporting.

REFERENCES

1. K. A. Olive et al. (Particle Data Group) (2014), Chin. Phys. C 38, 090001.
2. P. V. Dong, T. D. Tham, and H. T. Hung (2013), Phys. Rev. D 87, 115003.
3. P. V. Dong, D. T. Huong, F. S. Queiroz, and N. T. Thuy (2014), Phys. Rev. D 90, 075021.
4. D. T. Huong, P. V. Dong, C. S. Kim, and N. T. Thuy (2015), Phys. Rev. D 91, 055023.
5. P. V. Dong, Phys. Rev. D 92, 055026 (2015); P. V. Dong and D. T. Si, Kinetic mixing effect in the 3-3-1-1 model, arXiv:1510.06815 [hep-ph].
6. D. J. Gross and R. Jackiw (1972), Phys. Rev. D 6, 477.
7. H. Georgi and S. L. Glashow (1972), Phys. Rev. D 6, 429.
8. J. Banks and H. Georgi (1976), Phys. Rev. D 14, 1159.
9. S. Okubo (1977), Phys. Rev. D 16, 3528.
10. A. Doff and F. Pisano (2000), Mod. Phys. Lett. A 15, 1471.
11. P. V. Dong and H. N. Long (2006), Int. J. Mod. Phys. A 21, 6677.
12. F. Pisano and V. Pleitez (1992), Phys. Rev. D 46, 410.
13. P. H. Frampton (1992), Phys. Rev. Lett. 69, 2889.
14. R. Foot, O. F. Hernandez, F. Pisano and V. Pleitez (1993), Phys. Rev. D 47, 4158.
15. M. Singer, J. W. F. Valle and J. Schechter (1980), Phys. Rev. D 22, 738.
16. J. C. Montero, F. Pisano and V. Pleitez (1993), Phys. Rev. D 47, 2918.
17. R. Foot, H. N. Long and Tuan A. Chan (1994), Phys. Rev. D 50, 34.
18. P. V. Dong, L. T. Hue, H. N. Long, and D. V. Soa (2010), Phys. Rev. D 81, 053004.
19. P. V. Dong, H. N. Long, D. V. Soa, and V. V. Vien (2011), Eur. Phys. J. C 71, 1544.
20. P. V. Dong, H. N. Long, C. H. Nam, and V. V. Vien (2012), Phys. Rev. D 85, 053001.

CƠ CHẾ GAUGE CỦA MÔ HÌNH 3-3-1-1 TỐI THIỂU

Tóm tắt: *Trong bài báo này, chúng tôi nhận diện mô hình hạt gauge chuẩn và các hạt gauge mới. Ta có thể phân biệt được hạt gauge chuẩn và hạt gauge mới. Chúng tôi chứng minh rằng số lượng của hạt gauge trùng với số lượng của hạt Goldstone, chỉ để lại hạt photon và gluon không khối lượng do đó vector vật chất tối được thảo luận.*

PACS numbers: 12.60.-i

Từ khoá: *Vật chất tối, mô hình 3-3-1-1.*

FUZZY INITIAL VALUE PROBLEM FOR FRACTIONAL DIFFERENTIAL EQUATION ON LINEAR CORRELATED FUZZY FUNCTION SPACE

Hoang Thi Phuong Thao
University Of Languages And International Studies

Abstract: *The paper focuses on linear correlated fuzzy function space* $\mathbb{R}_{\mathcal{F}_{(A)}}$ *with its special properties . Firstly, we show some special properties of the norm operator on two spaces* $\mathbb{R}_{\mathcal{F}_{(A_1)}}, \mathbb{R}_{\mathcal{F}_{(A2)}}$ *with* $A_2 = kA_1, A_1, A_2$ *given fuzzy numbers. Besides, we study the fractional differential equation in the space* $\mathbb{R}_{\mathcal{F}_{(A)}}$ *. By combining the special properties of space with fractional order calculations and the fixed point theorem, we have built assumptions to ensure that the problem has unique solution.*

Key words: *Fractional equation, Fréchet derivative, Caputo-Fréchet derivative*

Received: 26 March 2020
Accepted for publication: 21 April 2020
Email: hpthao@flssl.edu.vn.

1. INTRODUCTION

In mathematics, fuzzy sets (uncertain sets) are somewhat like sets whose elements have degrees of membership. Fuzzy sets were introduced independently by L.A. Zadeh [38] in 1965 as an extension of the classical sets. But applications of fuzzy sets in inventory control problems were around 15-20 years back. Among these works, one can refer the works of Mandal and Maiti [24], Wee et al. [37]. In [38], Zadeh defined operations on set of fuzzy numbers with the following subtraction $u \ominus_H v = w \Leftrightarrow u = v + w$. It is well-known that this usual Hukuhara difference between two fuzzy numbers only exists under very restrictive conditions [12]. To fix it, L. Stefanini introduced the fuzzy gH-difference in [34,35]. The gH-difference of two fuzzy numbers exists under much less restrictive conditions, however it does not always exist [35]. This makes the space of fuzzy numbers not a Banach space. To overcome this difficulty, Estevão Esmi [4] presented a practical way to introduce the space

of linear correlated fuzzy numbers $\mathbb{R}_{\mathcal{F}_{(A)}} = \{\psi_A(q,r) / (q,r) \in \mathbb{R}^2\}$ where ψ_A is the operator that associates each vector $(q,r) \in \mathbb{R}^2$ with a fuzzy number $q[A]^\alpha + r$ where $[A]^\alpha$ stands for the α-level of a fuzzy number A for all $\alpha \in [0,1]$. The set $\mathbb{R}_{\mathcal{F}_{(A)}}$ equipped with addition and and scalar product defined in naturally 3.7 in [4] is a Banach space.

Hukuhara differentiability of fuzzy-valued functions is generalization of Hukuhara differentiability of set-valued functions. This differentiability is based on Hukuhara difference. Hukuhara introduced this difference (subtraction) of two sets in [10]. He introduced the notions of integral and derivative for set-valued mappings and considered the relationship between them. This derivative is widely studied and analysed by researchers for set-valued as well as fuzzy-valued functions. A wide range of applications of Hukuhara derivatives are studied in fuzzy differential equations and fuzzy optimization problems. Unfortunately, the derivative is very restrictive. Its existence is based on certain conditions. Estevão Esmi [4] gave the concept of the derivative Fréchet. Accordingly, the derivative of function $f(t) = \psi_A(q(t), r(t))$ is calculated through the derivative of $q(t), r(t)$.

The concepts of fractional derivatives for a fuzzy valued function are either based on the notion of Hukuhara derivative (H-derivative) or on the notion of strongly generalized derivative (G-derivative). The concept of Hukuhara derivative is old and well known [10], Puri and Ralescu [31] and the concept of G-derivative was recently introduced by Bede and Gal [2]. The Fréchet derivative was introduced the first time by D. Behmardi and E. Dehghan, Nayeri in [8]. Since then, many research works about Fréchet derivative have been published in [36].

Fractional calculus and fractional differential equations arise naturally in a variety of fields such as rheology, viscoelasticity, electrochemistry, diffusion process etc [6,7,8,9]They are usually applied to replace the derivative time in a given evolution equation by a derivative of fractional order. One can find applications of fractional differential equations in signal processing and in the complex dynamic in biological tissues [23,24,25]. For a general overview, we refer the reader to the monographs of Samko et al. [33], Podlubny [32], Kilbas et.al [13,14] and the papers [15,16, 26, 33, 34]. Some new research results for the fractional equation can be mentioned as [18,19,20,21,23,25,26]

This paper has 3 main results: Firstly, based on Fréchet derivative given by E. Esmi [4] , we introduce the definition of Fréchet-Caputo fractional derivative. The paper focuses on exploiting the application of the Fréchet derivative to the system of fuzzy frational equations. Applying the Lipschitz fixed point theorem, we make assumptions about the system of equations to have solutions. Finally, we present an example to illustrate the result.

2. CONTENT

2.1. Preliminaries

2.1.1. The space of linear correlated fuzzy numbers

Denote by $\mathbb{R}_{\mathcal{F}}$ the space of all fuzzy numbers on the real line. According to [1], the characteristic properties of a fuzzy number u are presented via its α-cuts or level sets, which are defined by

$$[\boldsymbol{u}]^{\boldsymbol{\alpha}} = \begin{cases} \{\boldsymbol{x} \in \mathbb{R}: \boldsymbol{u}(\boldsymbol{x}) \geq \boldsymbol{\alpha}\} \; \boldsymbol{if}\; \boldsymbol{\alpha} \in (\mathbf{0}, \mathbf{1}] \\ \overline{\{\boldsymbol{x} \in \mathbb{R}: \boldsymbol{u}(\boldsymbol{x}) > \mathbf{0}\}} \; \boldsymbol{if}\; \boldsymbol{\alpha} = \mathbf{0}. \end{cases}$$

In addition, it is well-known that the level sets of u can be rewritten in the parametric form $[\boldsymbol{u}]^{\boldsymbol{\alpha}} = [\boldsymbol{u}_{\boldsymbol{\alpha}}^{-}, \boldsymbol{u}_{\boldsymbol{\alpha}}^{+}]$and the diameter of $[\boldsymbol{u}]^{\boldsymbol{\alpha}}$is given by $\boldsymbol{len}[\boldsymbol{u}]^{\boldsymbol{\alpha}} = \boldsymbol{u}_{\boldsymbol{\alpha}}^{+} - \boldsymbol{u}_{\boldsymbol{\alpha}}^{-}$for each $\boldsymbol{\alpha} \in [\mathbf{0}, \mathbf{1}]$

The space $(\mathbb{R}_{\mathcal{F}}, d_{\infty})$endowed with the supremum metric

$$d_{\infty}(u, v) = \sup_{0 \leq \alpha \leq 1} d_H([u]^{\alpha}, [v]^{\alpha})$$ for all $u, v \in \mathbb{R}_{\mathcal{F}}$, is a complete metric space (see [1]).

Definition 2.1.1 [4]

For each$\boldsymbol{A} \in \mathbb{R}_{\boldsymbol{\mathcal{F}}}$, define a mapping $\boldsymbol{\psi_A}: \mathbb{R} \times \mathbb{R} \to \mathbb{R}_{\boldsymbol{\mathcal{F}}}$by $(\boldsymbol{q}, \boldsymbol{r}) \mapsto \boldsymbol{\psi_A}(\boldsymbol{q}, \boldsymbol{r})$where the level sets of $\boldsymbol{\psi_A}(\boldsymbol{q}, \boldsymbol{r})$are $[\boldsymbol{\psi_A}(\boldsymbol{q}, \boldsymbol{r})]^{\boldsymbol{\alpha}} = \{\boldsymbol{qa} + \boldsymbol{r}: \boldsymbol{a} \in [\boldsymbol{A}]^{\boldsymbol{\alpha}}\}$. For convenience, denote the fuzzy number $\boldsymbol{\psi_A}(\boldsymbol{q}, \boldsymbol{r})$by $\boldsymbol{qA} + \boldsymbol{r}$and the range of $\boldsymbol{\psi_A}(\boldsymbol{q}, \boldsymbol{r})$ by$\mathbb{R}_{\boldsymbol{\mathcal{F}(A)}}$.

Definition 2.1.2

A fuzzy number u said to be symmetric w.r.t. $x \in \mathbb{R}$ if $u(x-y) = u(x+y)$ for all $y \in \mathbb{R}$. The fuzzy number u is non-symmetric if there doesn't exist x such that u is symmetric.

From the results in [4], if $\boldsymbol{A} \in \mathbb{R}_{\boldsymbol{\mathcal{F}}}$is non-symmetric fuzzy number then the arithmetic operations on the space $\mathbb{R}_{\mathcal{F}(A)}$ such as addition, subtraction and scalar product are well-defined. Indeed, let us recall that for each $\boldsymbol{B}, \boldsymbol{C} \in \mathbb{R}_{\boldsymbol{\mathcal{F}(A)}}$ and $\boldsymbol{\lambda} \in \mathbb{R}$it yields

(i) $B +_A C = \psi_A(\psi_A^{-1}(B) + \psi_A^{-1}(C))$
(ii) $\lambda B = \psi_A(\lambda \psi_A^{-1}(B))$
(iii) $B -_A C = B +_A (-1)C = \psi_A(\psi_A^{-1}(B) + (-1)\psi_A^{-1}(C))$.

In addition, the distance between elements of $\mathbb{R}_{\mathcal{F}(A)}$ can be measured by the metric d_A of the fuzzy number space. In particular, with $u = \psi_A(q_u, r_u), v = \psi_A(q_v, r_v), (q_u, r_u), (q_v, r_v) \in \mathbb{R}^2$, we have $d_A(u, v) = |q_u - q_v| + |r_u - r_v|$for all $u, v \in \mathbb{R}_{\mathcal{F}(A)}$.

It is well-known that $(\mathbb{R}_{\mathcal{F}_{(A)}}, d_A)$is a complete metric space. Moreover, for each $u = \psi_A(q_u, r_u)$due to the fact that the space $\mathbb{R}_{\mathcal{F}(A)}$is isometric to the space $\mathbb{R}^2$ it implies that the space $(\mathbb{R}_{\mathcal{F}_{(A)}}, +_A, \cdot_A, \|.\|_{\psi_A})$ with the induced norm is a Banach space since $\mathbb{R}_{\mathcal{F}_{(A)}}$ is isometric to $\mathbb{R}^2$ with the norm $\|u\|_{\psi_A} = \|\psi_A^{-1}(u)\|_{\mathbb{R}^2} = |q_u| + |r_u|$.

Lemma 2.1.1

If A_1, A_2 be two non-symmetric fuzzy numbers such that $A_2 = kA_1$ with $k \in \mathbb{R}\backslash\{0\}$then two norms $\|\cdot\|_{A_1}$and $\|\cdot\|_{A_2}$ are equivalent.

Proof

Because $u \in \mathbb{R}_{\mathcal{F}(A_1)}$then u also belongs to the space $\mathbb{R}_{\mathcal{F}(A_2)}$, there exist two pairs $({}^1q_u, {}^1r_u), ({}^2q_u, {}^2r_u) \in \mathbb{R}^2$ such that $u = \psi_{A_1}({}^1q_u, {}^1r_u) = \psi_{A_2}({}^2q_u, {}^2r_u)$, which follows that$\psi_{A_1}({}^1q_u, {}^1r_u) = \psi_{A_1}(k^2q_u, {}^2r_u)$.

Our aim is to prove that two norms $\|\cdot\|_{A_1}$and $\|\cdot\|_{A_2}$are equivalent, which means that there exist $m, M > 0$ such that $m\|u\|_{A_1} \le \|u\|_{A_2} \le M\|u\|_{A_1}$for all $u \in \mathbb{R}_{\mathcal{F}(A_1)}$.

Indeed, since A_1 is a non-symmetric fuzzy number then ψ_{A_1} is an injection and hence, we directly obtain$\begin{cases} {}^1q_u = {}^2q_u k \\ {}^1r_u = {}^2r_u \end{cases}$.

By the definition of norm, we have

$$\begin{cases} \|u\|_{A_1} = |{}^1q_u| + |{}^1r_u|, \\ \|u\|_{A_2} = |{}^2q_u| + |{}^2r_u| = \left|\dfrac{{}^1q_u}{k}\right| + |{}^1r_u|. \end{cases}$$

Now, we consider some following cases:

Case 1: If $|k|>1$ then it implies that

$\frac{1}{|k|}(|{}^1q_u| + |{}^1r_u|) \le \|u\|_{A_2} \le |{}^1q_{u|} + |{}^1r_u|$ or equivalently $\frac{1}{|k|}\|u\|_{A_1} \le \|u\|_{A_2} \le \|u\|_{A_1}$which proves that $\|\cdot\|_{A_1} \sim \|\cdot\|_{A_2}$.

Case 2: If $|k|>1$then it implies that

$$|{}^1q_{u|} + |{}^1r_u| \le \|u\|_{A_2} \le \frac{1}{|k|}(|{}^1q_u| + |{}^1r_u|)$$

or equivalently, $\|u\|_{A_1} \le \|u\|_{A_2} \le \frac{1}{|k|}\|u\|_{A_1}$which proves that $\|\cdot\|_{A_1} \sim \|\cdot\|_{A_2}$.

Case 3: If $|k|=1$then it is obvious that $\|u\|_{A_1} = \|u\|_{A_2}$for all $u \in \mathbb{R}_{\mathcal{F}(A)}$

Hence, the proof is completed.

Remark 2.1.1

However, if A_1, A_2 are symmetry fuzzy numbers then the similar conclusion as in Lemma 2. 1 can't be obtained. Indeed, since $A_2 = kA_1$ if A_1 is a symmetric fuzzy number w.r.t. $x \in \mathbb{R}$ then so is A_2. Next, by using Theorem 3.15 in \cite{este}, our proof is divided into following cases:

Case 1: If ${}^1q_u = {}^2q_u k$ and ${}^1r_u = {}^2r_u$ then by similar arguments as in Lemma 2.1.1 , we have $\|\cdot\|_{A_1} \sim \|\cdot\|_{A_2}$.

Case 2: If ${}^1q_u = -{}^2q_u k$ and ${}^1r_u = 2{}^2q_u x + {}^2r_u$ then it implies that $\begin{cases} {}^2q_u = -\frac{{}^1q_u}{k} \\ {}^2r_u = 2\frac{{}^1q_u x}{k} + {}^1r_u. \end{cases}$

Therefore, we have

$$\|u\|_{A_2} = |{}^2q_u| + |{}^2r_u| \le \left|\frac{{}^1q_u}{k}\right| + \left|\frac{2x}{k}\right| |{}^1q_u| + |{}^1r_u| \le \left(\frac{1}{|k|} + \left|\frac{2x}{k}\right|\right) |{}^1q_u| + |{}^1r_u|.$$

Then, by denoting $M = \max\left\{\frac{1}{|k|} + \left|\frac{2x}{k}\right|; 1\right\}$, we directly get that $\|u\|_{A_2} \le M \|u\|_{A_1}$. For each $u \in \mathbb{R}_{\mathcal{F}(A_1)}$, we have

$$\begin{aligned} \mathrm{F}(\|u\|) &:= \|u\|_{A_2}^2 - m^2\|u\|_{A_1}^2 \\ &= \left(\left|\frac{{}^1q_u}{k}\right| + \left|\frac{2x^1}{k} q_u + {}^1r_u\right|\right)^2 - m^2(|{}^1q_u| + |{}^1r_u|)^2. \end{aligned} \tag{1}$$

Subcase 2.1: If ${}^1q_u {}^1r_u \ge 0$ and $\frac{{}^1q_u}{k}\left(\frac{2^1q_u x}{k} + r\right) \ge 0$ then the expression (1) becomes

$$\begin{aligned} &\|u\|_{A_2}^2 - m^2\|u\|_{A_1}^2 \\ &= \left[\left(\frac{2x+1}{k}\right)^2 - m^2\right]^1 q_u^2 + \left(\frac{4x^1r_u}{k} + \frac{2^1r_u}{k} - 2^1r_u m^2\right)^1 q_u + {}^1r_u^2(1 \\ &- m^2) \end{aligned}$$

If the above expression is known as a quadratic function of variable ${}^1q_u$ then the delta discriminant Δ' of the quadratic equation $\mathrm{F}(\|u\|) = 0$ is given by

$$\Delta' = {}^1r^2m^2\left(\frac{2x+1}{k} - 1\right)^2 \ge 0,$$

which means that the solution set of the quadratic equation $\mathrm{F}(\|u\|) = 0$ is nonempty.

Subcase 2.2: If ${}^1q_u {}^1r_u \ge 0$ and $\frac{{}^1q_u}{k}\left(\frac{2^1q_u x}{k} + r\right) \le 0$ then the expression (1) becomes

$$\| u \|_{A_2}^2 - m^2 \| u \|_{A_1}^2 = \left[\left(\frac{2x-1}{k} \right)^2 - m^2 \right]^1 q_u^2 + \left(\frac{4x^1 r_u}{k} - \frac{2^1 r_u}{k} - 2m^{21} r_u \right)^1 q_u + {}^1 r_u^2 (1-m^2)$$

If the above expression is known as a quadratic function of variable ${}^1 q_u$ then the delta discriminant Δ' of the quadratic equation $\mathrm{F}(\| u \|) = 0$ is given by

$$\Delta' = {}^1 r^2 m^2 \left(\frac{2x-1}{k} - 1 \right)^2 \geq 0$$

which means that the solution set of the quadratic equation $\mathrm{F}(\| u \|) = 0$ is nonempty.

Subcase 2.3 If ${}^1 q_u {}^1 r_u \leq 0$ and $\frac{{}^1 q_u}{k} \left(\frac{2^1 q_u x}{k} + r \right) \geq 0$ then the expression (1) becomes

$$\| u \|_{A_2}^2 - m^2 \| u \|_{A_1}^2 = \left[\left(\frac{2x+1}{k} \right)^2 - m^2 \right]^1 q_u^2 + \left(\frac{4x^1 r_u}{k} + \frac{2^1 r_u}{k} + 2m^{21} r_u \right)^1 q_u + {}^1 r_u^2 (1-m^2)$$

If the above expression is known as a quadratic function of ${}^1 q_u$ then the delta discriminant Δ' of the quadratic equation $\mathrm{F}(\| u \|) = 0$ is given by

$$\Delta' = {}^1 r^2 \left(\frac{2x+1}{k} + 1 \right)^2 \geq 0$$

which means that the solution set of the quadratic equation $\mathrm{F}(\| u \|) = 0$ is nonempty.

We can conclude that there always exists an element $u \in \mathbb{R}_{\mathcal{F}(A_1)}$ or equivalently, a pair $({}^1 q_u, {}^1 r_u) \in \mathbb{R}^2$ such that the inequality $\|u\|_{A_2}^2 \leq m^2 \|u\|_{A_1}^2$ doesn't hold. Therefore, it implies that two norms $\| \cdot \|_{A_1} and \| \cdot \|_{A_2}$ are not equivalent.

2.2.2. Fréchet-Caputo fractional derivative

Let E, F be normed spaces and denote $\mathcal{L}(E, F)$ by the space of all continuous mappings.

Definition 2.2.1

Let $A \in \mathbb{R}_{\mathcal{F}}$

and

$$f \in \mathcal{L}(J, \mathbb{R}_{\mathcal{F}(A)}), f(t) = \psi_A(q(t), r(t))$$

with

$$q(t), r(t) \in L^1(J, \mathbb{R}) \cap C(J, \mathbb{R})$$

.

Then the Riemann-Liouville (RL) fractional integral of order $p \in (0,1]$ of function f is defined by

$$_F^{RL}\mathcal{J}_{0^+}^p f(t) = \psi_A\left(I_{0^+}^p q(t), I_{0^+}^p r(t)\right),$$

where $I_{0^+}^p q(t) = \frac{1}{\Gamma(p)}\int_0^t (t-s)^{p-1} q(s)ds, I_{0^+}^p r(t) = \frac{1}{\Gamma(p)}\int_0^t (t-s)^{p-1} r(s)ds.$

Definition 2.2.2

Let $A \in \mathbb{R}_{\mathcal{F}}$

And $f \in \mathcal{L}(J, \mathbb{R}_{\mathcal{F}(A)}), f(t) = \psi_A(q(t), r(t))$

Let $A \in \mathbb{R}_{\mathcal{F}}$ and $f \in \mathcal{L}(J, \mathbb{R}_{\mathcal{F}(A)}), f(t) = \psi_A(q(t), r(t))$ with. The Fréchet-Caputo fractional derivative of order of the function $p \in (0,1]$ of function f is defined by

$$_F^{C}\mathcal{D}_{0^+}^p f(t) = {}_F^{RL}\mathcal{I}_{0^+}^{1-p} f_{\mathcal{F}'}(t) = \psi_A\left(I_{0^+}^{1-p} q'(t), I_{0^+}^{1-p} r'(t)\right).$$

Definition 2.2.3

Let $A \in \mathbb{R}_{\mathcal{F}}$ be non-symetric fuzzy number and $f \in \mathcal{L}(J, \mathbb{R}_{\mathcal{F}(A)}), f(t) = \psi_A(q(t), r(t))$ with $q(t), r(t) \in L^1(J, \mathbb{R}) \cap C(J, \mathbb{R})$. The Fréchet-Riemann-Liouville fractional derivative of order $p \in (0,1]$ of the function f is defined by

$$_F^{RL}\mathcal{D}_{0^+}^p f(t) = ({}_F^{RL}\mathcal{I}_{0^+}^{1-p} f)_{\mathcal{F}'}(t) = \psi_A\left((I_{0^+}^{1-p} q(t))', (I_{0^+}^{1-p} r(t))'\right).$$

Definition 2.2.4

Let $A \in \mathbb{R}_{\mathcal{F}}$ and $f \in \mathcal{L}(J, \mathbb{R}_{\mathcal{F}(A)}), f(t) = \psi_A(q(t), r(t))$ with $q(t), r(t) \in L^1(J, \mathbb{R}) \cap C(J, \mathbb{R})$. The Fréchet-Riemann-Liouville fractional derivative of order $p \in (0,1]$ of the function f is defined by

$$_F^{RL}\mathcal{D}_{0^+}^p f(t) = ({}_F^{RL}\mathcal{I}_{0^+}^{1-p} f)_{\mathcal{F}'}(t) = \psi_A\left((I_{0^+}^{1-p} q(t))', (I_{0^+}^{1-p} r(t))'\right).$$

2.2.3. Application of the Fréchet -Caputo derivative

Consider following fuzzy intitial value problem (FIVP) for fractional differential equation

$$\begin{cases} {}^{C}_{F}\mathcal{D}^{p}_{0+}\mu(t) = f(t,\mu(t)) \\ \mu(0) = \mu_0, \end{cases} \quad (1)$$

where ${}^{C}_{F}\mathcal{D}^{p}_{0+}\mu(.)$ is the Fréchet-Caputo derivative of order $p \in (0,1]$ of the continuous function $\mu(.), \mu_0 \in \mathbb{R}_{\mathcal{F}_{(A)}}$ with A is a fuzzy number and the function $f: J \times C(J, \mathbb{R}_{\mathcal{F}_{(A)}}) \to \mathbb{R}_{\mathcal{F}_{(A)}}$ is continous function with $J = [0;b]$.

2.3. Existence of solution

If A is a fuzzy number and $\mu \in \mathbb{R}_{\mathcal{F}_{(A)}}$ is a solution to the problem (2) then it satisfies the following integral equation

$$\mu(t) = \mu_0 +_A \ {}^{RL}_{\mathcal{F}}\mathcal{I}^{p}_{0+} f(t,\mu(t)), \quad t \in J. \quad (2)$$

Proof

Because $f(t,\mu(t)) \in \mathbb{R}_{\mathcal{F}_{(A)}}$, there exists $q_{f\mu}, r_{f\mu}: J \to \mathbb{R}$ such that

$$f(t,\mu(t)) = \psi_A(q_{f\mu}(t), r_{\mu t}(t)).$$

Assume that is a solution of (2).

From ${}^{C}_{\mathcal{F}}\mathcal{D}^{p}_{0+}\mu(t) = f(t,\mu(t))$, we have ${}^{RL}_{\mathcal{F}}\mathcal{I}^{1-p}_{0+}\mu'(t) = f(t,\mu(t))$ wher $\mu'(t) = \psi_A(q'_{f\mu}(t), r'_{f\mu}(t))$.

By taking fractional integration of order p of two sides, we obtain

$${}^{RL}_{\mathcal{F}}\mathcal{I}^{p}_{0^+}\,{}^{RL}_{\mathcal{F}}\mathcal{I}^{1-p}_{0^+}\mu'(t) = {}^{RL}_{\mathcal{F}}\mathcal{I}^{1-p}_{0^+} f(t,\mu(t)).$$

This imlies $\int_0^t \mu'(s)ds = {}^{RL}_{\mathcal{F}}\mathcal{I}^{p}_{0^+} f(t,\mu(t))$. Therefore $\mu(t) = \mu_0 +_{A\mathcal{F}} {}^{RL}\mathcal{I}^{p}_{0^+} f(t,\mu(t)), t \in J$.

Definition 3.1.1

$\mu(t) \in \mathbb{R}_{\mathcal{F}_{(A)}}$ is called an intergral solution of the problem (2) if it satifies integral equation $\mu(t) = \mu_0 +_{A\mathcal{F}} {}^{RL}\mathcal{I}^{p}_{0^+} f(t,\mu(t))$.

On space $C(J, \mathbb{R}_{\mathcal{F}_{(A)}})$ we defined the supremum metric $\mathcal{H}$ and weighted metric ${}^{r}d$ as follows $\mathcal{H}(\mu,\Upsilon) = \sup_{t\in J} d_A(\mu(t),\Upsilon(t))$ and ${}^{r}d = \sup_{t\in J}\{t^r d_A(\mu(t),\Upsilon(t))\}$.

Theorem 3.1.1

Assume that A is a fuzzy number, the fuzzy-valued function f is jointly continuous and satisfies the Lipschitz condition w.r.t. the last argument, i.e., there exists a constant $L > 0$ so that

$$d_A(f(t,\mu(t)),\, f(t,\Upsilon(t))) \le L d_A(\mu(t),\Upsilon(t)).$$

for *all* $t \in J, \mu, \Upsilon \in C\left(J, \mathbb{R}_{\mathcal{F}(A)}\right)$.

Then, the fuzzy initial value problem (2) has a unique integral solution defined on J.

Proof: Let us define an operator $\mathcal{G}: C\left(J, \mathbb{R}_{\mathcal{F}(A)}\right) \to C\left(J, \mathbb{R}_{\mathcal{F}(A)}\right)$ by

$\mathcal{G}[\mu](t) = \mu_0 + {}_{AF}^{RL}\mathcal{I}^p f(t,\mu(t))$.

Assume that $(q_{f\mu}(t), r_{f\mu}(t)) \in \mathbb{R}\times\mathbb{R}, (q_{f\Upsilon}(t), r_{f\Upsilon}(t)) \in \mathbb{R}\times\mathbb{R}$ such that

$$f(t,\mu(t)) = \psi_A(q_{f\mu}(t), r_{f\mu}(t)),\; f(t,\Upsilon(t)) = \psi_A(q_{f\Upsilon}(t), r_{f\Upsilon}(t)).$$

For each $t \in J$ *and* $\mu, \Upsilon \in C\left(J, \mathbb{R}_{\mathcal{F}(A)}\right)$, we have

$d_A\left(\mathcal{G}[\mu](t), \mathcal{G}[\Upsilon](t)\right)$	$= d_A({}_F^{RL}\mathcal{J}^p f(t,\mu(t)), {}_F^{RL}\mathcal{J}^p f(t,\Upsilon(t)))$
	$= \lvert I_{0+}^p(q_{fu}(t) - q_{fv}(t))\rvert + \lvert I_{0+}^p(r_{fu}(t) - r_{fv}(t))\rvert$
	$= \lvert I_{0+}^p(q_{f\mu}(t) - q_{f\Upsilon}(t))\rvert + \lvert I_{0+}^p(r_{f\mu}(t) - r_{f\Upsilon}(t))\rvert$
	$\le \frac{1}{\Gamma(p)}\int_0^t (t-s)^{p-1}\left(\lvert q_{fu}(s) - q_{fv}(s)\rvert + \lvert r_{fu}(s) - r_{fv}(s)\rvert\right)ds$
	$\le \frac{1}{\Gamma(p)}\int_0^t (t-\gamma)^{p-1} d_A(f(\gamma,\mu(\gamma)), f(\gamma,\Upsilon(\gamma)))d\gamma$
	$\le \frac{1}{\Gamma(p)}\int_0^t (t-\gamma)^{p-1}\left(\lvert q_{f\mu}(\gamma) - q_{f\Upsilon}(\gamma)\rvert + \lvert r_{f\mu}(s) - r_{f\Upsilon}(\gamma)\rvert\right)d\gamma$
	$\le \frac{L\Gamma(p)}{\Gamma(2p)} t^{2p-1\ 1-p} d_A(\mu,\Upsilon).$

This implies

$$t^{1-p} d_A(\mathcal{G}[\mu], \mathcal{G}[\Upsilon]) \le \frac{L t^p \Gamma(p)^{1-p}}{\Gamma(2p)}\, d_A(\mu,\Upsilon) \tag{3}$$

The operator $\mathcal{G}^n$ is defined by

$\mathcal{G}^n[\mu](t) = \mathcal{G}(\mathcal{G}^{n-1}[\mu](t))$ for all $n \in \mathbb{N}, t \in J$.

Then, we will show that $\mathcal{G}^n$ is a contraction mapping for a big enough $n \in \mathbb{N}$. To this end, by induction principle, we will show that the following estimation holds for all $\mu \in C\left(J, \mathbb{R}_{\mathcal{F}(A)}\right)$. We have

$$d_A(\mathcal{G}^n[\mu](t), \mathcal{G}^n[\Upsilon](t)) \leq \frac{L^n t^{np+p-1}\Gamma(p)}{\Gamma((n+1)p)}(\quad {}^{1-p}d_A(\mu,\Upsilon)). \tag{4}$$

Indeed, if $n=1$ then we gain the estimation (3) from the inequality (4). Moreover, let us assume that the estimation (3) is true for $n=k$ and then, we will prove that (3) holds for $n=k+1$. For each $\mu,\Upsilon \in C\left(J, \mathbb{R}_{\mathcal{F}(A)}\right)$ and $t \in J$, we have

$d_A(\mathcal{G}^{k+1}[u](t), \mathcal{G}^{k+1}[\Upsilon](t))$	$= d_A(\mathcal{G}(\mathcal{G}^k[u](t)), \mathcal{G}(\mathcal{G}^k[u](t)))$
	$\leq d_A\left(\quad {}_F^{RL}\mathcal{J}_{0^+}^p f(t, \mathcal{G}^k[\mu](t)), {}_F^{RL}\mathcal{J}_{0^+}^p f(t, \mathcal{G}^k[\Upsilon](t))\right)$
	$\leq \frac{1}{\Gamma(p)}\int_0^t (t-\gamma)^{p-1}\, d_A(f(s, \mathcal{G}^k[\mu](\gamma), f(s, [\mu](\gamma))d\gamma$
	$\leq \frac{L}{\Gamma(p)}\int_0^t (t-\gamma)^{p-1}\, d_A(\mathcal{G}^k[\mu](\gamma), \mathcal{G}^k[\Upsilon](\gamma))d\gamma.$

By employing the induction hypothesis, we have

$$\begin{aligned} d_A(\mathcal{G}^{k+1}[\mu](t), \mathcal{G}^{k+1}[\Upsilon](t)) &\leq \frac{L^{k+1}\quad {}^{1-p}d_A(\mu,\Upsilon)}{\Gamma((k+1)p)}\int_0^t (t-\gamma)^{p-1}\gamma^{kp+p-1}d\gamma \\ &\leq \frac{L^{k+1}t^{(k+2)p-1}}{\Gamma((k+2)p)}B(p,(k+1)p)(\quad {}^{1-p}d_A(\mu,\Upsilon)) \\ &\leq \frac{L^{k+1}t^{(k+2)p-1}\Gamma(p)}{\Gamma((k+2)p)}(\quad {}^{1-p}d_A(\mu,\Upsilon)), \end{aligned}$$

where $B(p,q)$ is Beta function [14]. Therefore, the inequality (3) holds for $n=k+1$. From the inequality (3), we have

$$t^{1-p}d_A(\mathcal{G}^n[\mu](t), \mathcal{G}^n[\Upsilon](t)) \leq \frac{L^n t^{np}\Gamma(p)}{\Gamma((n+1)p)}(\quad {}^{1-p}d_A(\mu,\Upsilon)).$$

Then, by taking supremum both sides, we obtain ${}^{1-p}d_A(\mathcal{G}^n[\mu], \mathcal{G}^n[\Upsilon]) \leq \frac{Lb^{np}\Gamma(p)}{\Gamma((n+1)p)}(\quad {}^{1-p}d_A(\mu,\Upsilon)) \to 0$ as $n \to \infty$.

It implies that the operator $\mathcal{G}^n$ is a contraction when $n \in \mathbb{N}$ is big enough. Applying contraction principle, we guarantee the unique existence of fixed point μ of the operator $\mathcal{G}^n$, that is the unique integral solution of the problem (2). The proof is complete.

2.4. Example

In this example, we consider the equation of the mass-Spring-DamperSystem as

$$\frac{m}{\sigma^{2(1-\gamma)}}{}_F^C\mathcal{D}_{0^+}^{2\gamma}\mu(t)+\frac{\beta}{\sigma^{1-\gamma}}+_A k\mu(t)=F(t) \qquad 0<\gamma\leq 1$$

where A is a fuzzy number, $\mu(t), F(t)\in\mathbb{R}_{\mathcal{F}_{(\mathcal{A})}}$ the mass is m, the damping coefficient is β the spring costant is k anf $F(t)$ represent the forcing function, an auxilary parameter σ is introduced into the fractional tenporal operator:

$$\frac{d}{dt}\to\frac{1}{\sigma^{1-\gamma}}.\frac{d^{\gamma}}{dt^{\gamma}}, \qquad m-1\leq\gamma\leq m, m\in M=1,2,3,\ldots$$

$$\frac{d^2}{dt^2}\to\frac{1}{\sigma^{2(1-\gamma)}}.\frac{d^{2\gamma}}{dt^{2\gamma}}, \qquad m-1\leq\gamma\leq m, m\in M=1,2,3,\ldots$$

Consider a constant source, $F(t)=f_0, \mu(0)=\mu_0, \frac{d\mu}{dt}(0)=0$, we have equation (2), may be writen as follow:

$${}_F^C\mathcal{D}_{0^+}^{2\gamma}\mu(t)=\frac{\eta^2}{k}f_0-_A\eta^2\mu(t)$$

where $\eta^2=\frac{k\sigma^{2(1-\gamma)}}{m}$.

We can see that $f(t,\mu(t))=\frac{\eta^2}{k}f_0-_A\eta^2\mu(t)$is jointly continous function and satisfies the Lipschitz condition. So, we obtain the solution of problem is

$$\mu(t)=\left(\mu_0-_A\frac{f_0}{k}\right).E_{2\gamma}\{-\eta^2t^2\gamma\}+_A\frac{f_0}{k}.$$

where the Mittag-Leffler function $E_a(t)$ is defined by a power as: $E_a(t)=\sum_{m=0}^{\infty}\frac{t^m}{\Gamma(am+1)}$.

3. CONCLUSION

Firstly, in this paper, we introduced a new concept of fractional differentiability for a class of linear correlated fuzzy valued function namely the Fréchet-Caputo fractional derivative.

After that, we studied fuzzy fractional PDEs under Fréchet- Caputo differentiability. By using the fixed point theorem, we have proved some new results on the existence and uniqueness of fuzzy solution for the fuzzy intitial value problem. Fréchet derivative makes this problems have only a unique solution instead of having two different types of solutions such as the usual gH derivative .

REFERENCES

1. B. Bede, L. Stefanini (2013), Generalized differentiability of fuzzy-valued functions, Fuzzy Sets Syst. 230, 119-141.
2. B. Bede, S.G. Gal (2005), Generalizations of the differentiability of fuzzy number valued functions with applications to fuzzy differential equations, Fuzzy Sets Syst. 151, 581–599.
3. T.G. Bhaskar, V. Laskshmikantham, S. Leela (2009), Fractional differential equations with a Krasnoseski-Krein type condition, Nonlinear Analysis Hybrid Systems (3).
4. E.Esmi, F.S. Pedro, L.C. Barros (2018), Fréchet derivative for linearly correclated fuzzy function, Information Sciences, 435, 150-160.
5. Daryoush Behmardi, Encyeh Dehghan Nayeri (2008), Applied Mathematical Sciences, Introduction of Fr´echet and Gateaux Derivative, Vol 2, no. 20, 975 – 980.
6. S. Das (2008), Functional Fractional Calculus for System Identification and Controls, Springer-Verlag, Berlin, Heidelberg.
7. S. Das (2011), Functional Fractional Calculus, Springer-Verlag, Berlin, Heidelberg.
8. Daryoush Behmardi, Encyeh Dehghan Nayeri (2008), Applied Mathematical Sciences, Introduction of Fréchet and Gateaux Derivative, Vol 2(20), 975-980.
9. H. Garg (2014), A hybrid GA-GSA algorithm for optimizing the performance of an industrial system by utilizing uncertain data, in Handbook of Research on Artificial Intelligence Techniques and Algorithms, P. Vasant (Ed.), IGI Global, USA, Ch. 20, pp. 625 – 659.
10. M. Hukuhara (1967), Integration des applications mesurables dont la valeur est un compact convex, Funkc. Ekvacioj 10, 205–229.
11. M. Mazandarani, M. Najariyan (2014), Type-2 fuzzy fractional derivatives, Commun. Nonlinear Sci. Numer. Simul. 19, 2354-2372.
12. O. Kaleva (1987), Fuzzy differential equations, Fuzzy Sets Syst, 24, 301–317.
13. A.A. Kilbas, H.M. Srivastava, J.J. Trujillo (2006), Theory and Applications of Fractional Differential Equations, Elsevier Science B.V, Amsterdam.
14. Kilbas, AA, Srivastava, HM, Trujillo (2006), Theory and Applications of Fractional Differential Equations. Elsevier, Amesterdam.
15. A. Ya. Kruger (2006), On Fréchet Subdifferentials, Journal of Mathematical Sciences, Vlume 116(3),3325–3358.
16. V. Lakshmikantham, J.J. Nieto (2003), Differential equations in metric spaces: An introduction and an application to fuzzy differential equations, Dyn. Contin., Discrete Impuls. Syst. Ser. A: Math. Anal,10, 991–1000.
17. J. Li, A. Zhao, J. Yan (2012), The Cauchy problem of fuzzy differential equations under generalized differentiability, Fuzzy Sets Syst. 200, 1–24.
18. H.V. Long, J.J. Nieto, N.T.K. Son, New approach for studying nonlocal problems related to differential systems and partial differential equations in generalized fuzzy metric space, Vietnam Journal of Mathematics 46 (3), 531–555
19. H.V. Long, N.T.K. Son, N.T.M. Ha, L.H. Son (2014), The existence and uniqueness of fuzzy solutions for hyperbolic partial differential equations, Fuzzy Optim. Decis. Making 13, 435-462.
20. H.V. Long, N.T.K. Son, N.T.M. Ha, Jen-Chih Yao (2017) , Ulam stability for fractional partial integro diferential equation with uncertainty, Acta mathematica Vietnamica, 4, 675–700.
21. H.V. Long, N.T.K. Son, H.T.T. Tam (2017), The solvability of fuzzy fractional partial differential equations under Caputo gH-differentiability,Fuzzy set and system, 35-63.

22. R. Magin (2010), Fractional calculus model of complex dynamics in biological tissues, Computers and Mathematics with Application,59, 586-1593.
23. R. Magin, M. Ortigueira, I.Podlubny and J.J.Trujillo (2011), On the fractional signals and systems, Signal Processing ,91, 350-371.
24. M. Mandal, and M., Maiti (1999), Inventory of damagable items with variable replenishment rate, stock-dependent demand and some units in hand, Applied Mathematical Modelling, 23, 799-807.
25. Merala, T.J. Roystona and R. Magin (2010), Fractional calculus in viscoelasticity: An experimental study, Communications in Nonlinear Science and Numerical Simulation 15, 939-945.
26. K.S. Miller, B. Ross (1993), An introduction to the fractional calculus and differential equations, John Wiley, New York.
27. R. Moore, W. Lodwick (2003), Interval analysis and fuzzy set theory, Fuzzy Sets Syst. 135 (1), 5–9.
28. Oswaldo González-Gaxiola (2009), A Note on the Derivation of Fréchet and Gâteaux, Applied Mathematical Sciences, Vol.3 (19), 941 - 947.
29. J.J. Nieto and R.Rodríguez-Lospez (2005), Contractive mapping theorems in partially ordered sets and applications to ordinary differential equation, Order, 22, 223-239.
30. I. Podlubny (1999), Fractional Differential Equations, vol. 198 of Mathematics in Science and Engineering, Academic Press, New York, NY, USA.
31. M. Puri, D. Ralescu (1999), Differentials of fuzzy functions, J. Math. Anal. Appl. 91,(1983), 552–558.
32. L. Podlubny, Fractional differential equation, San Diego: Academic Press.
33. S.G.Samko, A.A. Kilbas, O.I. Marichev, (1993) Fractional Integrals and Derivatives: Theory and Applications, Gordon and Breach Science Publishers, Switzerland.
34. L. Stefanini, A generalization of Hukuhara difference, in: D. Dubois, M.A. Lubiano, H. Prade, M.A. Gil, P. Grzegorzewski, O. Hryniewicz (Eds.) (2008), Soft Methods for Handling Variability and Imprecision, Series on Advances in Soft Computing, vol.48.
35. L. Stefanini, B. Bede (2009), Generalized Hukuhara differentiability of interval-valued functions and interval differential equations, Nonlinear Anal, 71 ,1311-1328.
36. L. Stefanini (2010), A generalization of Hukuhara difference and division for interval and fuzzy arithmetic, Fuzzy Sets Syst, 161, 1564–1584.

BÀI TOÁN VỀ GIÁ TRỊ CỦA BIẾN MỜ CHO CÁC PHƯƠNG TRÌNH VI PHÂN PHÂN SỐ TRÊN KHÔNG GIAN TUYẾN TÍNH

***Tóm tắt**: Bài báo tập trung vào không gian hàm mờ tương quan tuyến tính với các tính chất đặc biệt của nó $\mathbb{R}_{\mathcal{F}_{(A)}}$ Đầu tiên, chúng tôi hiển thị một số thuộc tính đặc biệt của toán tử chuẩn trên hai không gian với $\mathbb{R}_{\mathcal{F}_{(A_1)}}, \mathbb{R}_{\mathcal{F}_{(A2)}}$ with$A_2 = kA_1, A_1, A_2$ mang lại các giá trị mờ. Bên cạnh đó, chúng tôi nghiên cứu phương trình vi phân phân số trong không gian $\mathbb{R}_{\mathcal{F}_{(A)}}$bằng cách kết hợp các tính chất đặc biệt của không gian với các phép tính bậc phân số và định lý điểm cố định, chúng tôi đã xây dựng các giả thiết để đảm bảo rằng bài toán có nghiệm duy nhất.*

***Từ khoá**: Phương trình vi phân, đạo hàm Fréchet, đạo hàm Caputo-Fréchet.*

SMARTPHONE BASED OPTICAL SENSOR FOR ENVIRONMENTAL APPLICATION

Mai Hong Hanh

University of Science, Vietnam National University

Abstact: *In the recent years there has been a huge development on the detection devices using smartphone that are reliable, easy-to-use, and low cost. In this work, a smartphone based optical sensor is constructed by implementing an external light source, a collimating lens, a diffraction grating, and a CMOS chip of a smartphone as a detector. The construction allows the device to function with an optical bandwidth of 300 nm (from 400 to 700 nm) and its resolution of 0.26 nm/pixel. It can be used both for measuring the absorption, transmission emission spectrum. As a proof of concept, the optical sensor using smartphone is then applied to investigate concentration of methylene blue (MB), a reactive dye in wastewater from textile industry. Despite of its cost-effectiveness, the sensor exhibits reliable results, which can be considerably comparable with that of laboratory instrument.*

Keywords: *Optical sensor, smartphone, absorption, methylene blue.*

Received: 11 March 2020

Accepted for publication: 21 April 2020

Email: hanhhongmai@hus.edu.vn

1. INTRODUCTION

In recent years, there has been an increased interest in the development of simple, light-weight, low-cost, portable and rapid detection devices for applications related to clinical diagnosis, health care and environmental monitoring [1–3]. The combinations of portable mobile devices with internet connectivity, touch screen displays, high resolution cameras, and high-performance CPUs have facilitated the development of a new device generation. These kinds of devices are not only suited for scientific research but also for daily work which normally does not require dedicated instruments and laboratory conditions for sensing, detection and analysis. Since smart phones are ubiquitous, thus, integrating detecting smart phones into devices is a promising approach for the creation of a detection device for public health and environmental protection. As a result, many research groups

have actively engaged in developing or converting smartphones into optical sensing devices such as optical microscopes [4], spectroscopy [5–7], surface plasmon resonance biosensors [8,9], crystal integrated label-free biosensors [10,11], blood glucose monitors [12], or pH sensors [13,14].. These devices have been applied for food quality analysis, for diagnosing disease, for monitoring of nutritional status and water quality, or for determining the presence of environmental contaminants.

Methylene blue (MB) is one of the most widely used substances for dyeing cotton, wood and silk. It is also commonly used in textile industry. The release of MB into the water is a concern due to the toxicity, mutagenicity and carcinogenicity of the MB and its biotransformation products [15–17]. Several procedures have been reported for measurement of MB in different matrices, including UV-VIS analysis, liquid-liquid extraction and solid phase extraction and final analysis by high performance liquid chromatography [18,19]. UV-VIS analysis is a relatively simple method that is the most widely used in various areas.

In the present study, a low-cost, portable, sensitive smartphone based optical sensor for environmental applications is reported. The optical sensor is based on UV-VIS method includes an entrance slit, a single lens, a diffraction grating, and a smartphone. This CMOS camera is a wavelength-independent photon collector that can function as a detector. The images captured by the sensor are then converted into intensity distribution plots versus wavelength. The cradle held the smartphone and the optical elements were made by 3D printing. As a proof of concept, the variation in absorbance of MB with different concentrations was measured. The obtained results were then compared with that of a conventional laboratory spectrometer to determine the accuracy, and the sensitivity of the sensor.

2. CONTENT

2.1. Material and methods

Methylene blue (MB) was purchased from Merk. In this work, six solution samples corresponding with six concentrations of 0.5 mM, 1 mM, 1.5 mM, 2 mM, 2.5 mM, and 3 mM were prepared.

2.2. Optical setup

The smartphone-based sensor was designed to interface with an iPhone 5s smartphone by the Apple Inc., of which the camera can function as a digital light detector. A polylactic acid (PLA) plastic cradle was printed by a 3D imprint machine with the resolution of 0.05 mm. It was installed to hold all the optics including a cuvette holder, an entrance slit (50 µm, Thorlab), a collimating lens (focal length of 50 mm, Thorlab), and a diffraction grating (1300 grooves/mm, Edmund Optics). It is robust and can exclude light from external source. On top of the cradle, the smartphone camera was fixed firmly. For absorption measurement, a LED with the wavelength from 400nm to 700nm was used to illuminate the sample cuvette. Only transmitted light collected was allowed to pass through the entrance slit. Light that

entered the optical chamber was then collimated by the collimating lens, and guided to the diffraction grating. The diffraction grating, aligned at an angle of ~ 47 degrees with respect to the fluorescent beam so as only the first-order diffracted light was directed onto the CMOS camera (8 MP, 3264×2448 pixels) of the smartphone. The captured images were then analysed and converted into intensity distribution plots versus wavelength using ImageJ software. Due to the spectral responsibility of the Si-based sensor and internal infrared cut-off filters within the camera optics, the sensor can function in a wavelength range from 400 nm to 700 nm. A schematic diagram of the designed detection system is shown in Figure 1.

Prior to studying the sensor characteristics, the pixel information of the captured fluorescent spectrum was calibrated in wavelength scale. To do this, the established procedure for calibrating smartphone based optical devices was applied in this work [5]. In detail, the entrance slit was illuminated by two lasers: HeCd and HeNe laser with known wavelength of 442 nm and 532.8 nm respectively. The known wavelengths of the two lasers and their wavelength separation were used to set the wavelength span corresponding to the pixel scale along the illumination direction. In this work, the wavelength span of 190.8 nm corresponds to 571 pixels, leads to a spectrum/pixel resolution of 0.334 nm/pixel. In comparison with other works on smartphone based devices, the smartphone based sensor exhibits compatible spectrum/pixel resolution [14115].

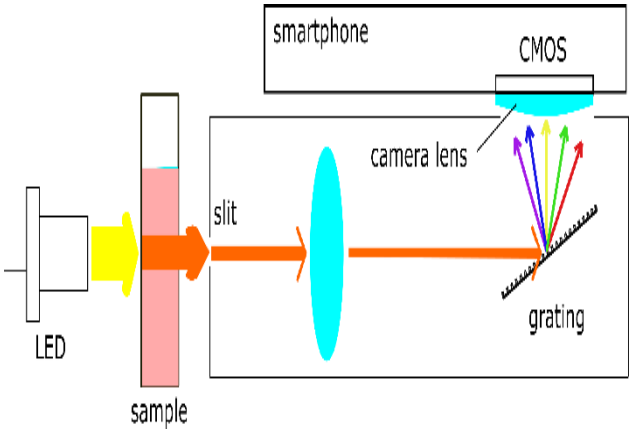


Figure 1. *Schematic of the smartphone based optical sensor*

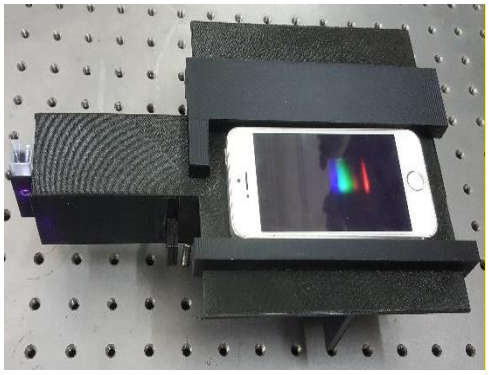


Figure 2. *Final assembly of the smartphone-based sensor*

2.3. Results and discussion

The absorption spectra of MB with different concentrations measured by a laboratory UV VIS spectrometer (AvaSpec-ULS2048, Avantes) are presented in Figure 3. The six samples were illuminated by a LED. The laboratory spectrometer's detector is a CCD having a spectral resolution of 0.5 nm. Similar to previous published reports, absorption spectrum of MB has a strong absorption peak at 663 nm[20]. The absorption spectra of the same MB with different concentrations from the smartphone-based sensor are presented in Figure 4.

As seen in the figure, the sensor exhibited similar absorption responses in comparison with that of the laboratory spectrometer. A clear increase at 663 nm with respect to the increase of MB concentrations observed in both cases demonstrates the potential of using the sensor for MB detection.

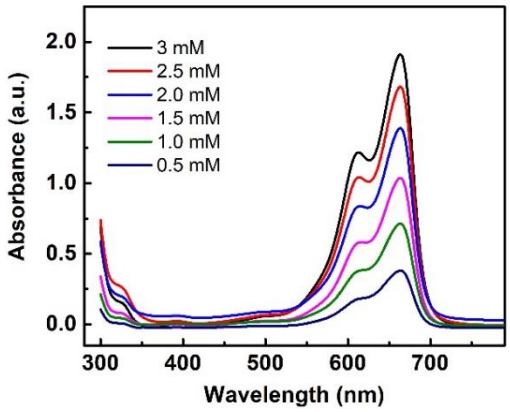


Figure 3. *Spectra absorption of MB measured by a laboratory UV VIS Avantes spectrometer*

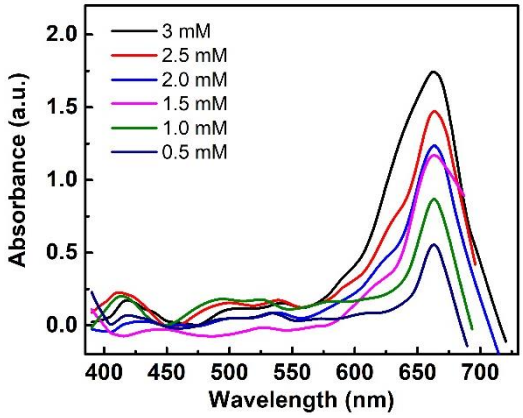


Figure 4. *Spectra absorption of MB measured by smartphone-based optical sensor*

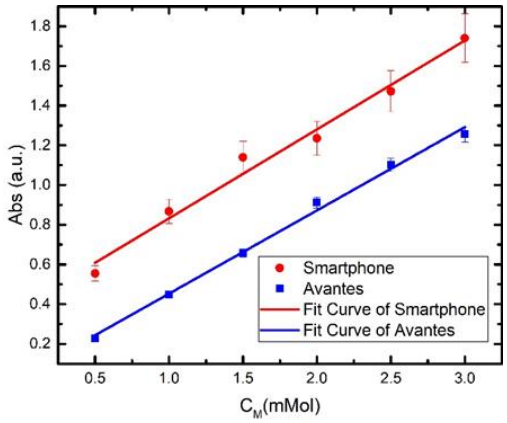


Figure 5. *Calibration curves obtained from smartphone based optical sensor and laboratory UV-VIS spectrometer results.*

In order to evaluate the accuracy and the sensitivity of the smartphone-based sensor, calibration curves which demonstrate the dependence of peak intensities at 663 nm as a function of the MB concentrations were taken into account (Figure 5). The slopes of the two calibration curves which demonstrate the sensitivity of the two sensors are of 0.44 for the smartphone-based sensor and of 0.41 for the Avantes laboratory spectrometer. The linear correlations (R^2) are relatively relevant of 0.97 and of 0.99. The error of the laboratory sensor is of 3%, while the error of the smartphone-based sensor is 7%. These positive agreements confirmed the sensitivity, the accuracy of the smartphone-based sensor. In addition, the

smartphone-based sensor was made from inexpensive components. It is also compact and portable, making it suitable for safety food inspection and in-field testing.

3. CONCLUSIONS

A low-cost, accurate, and portable smartphone-based sensor for environmental applications was developed. The sensor uses a LED as the light source, a CMOS camera of a smartphone as the detector, and a grating as the dispersive unit. The spectrum/pixel resolution is 0.334 nm/pixel. As a proof of concept, the smartphone-based sensor was applied to measure the concentrations of MB. Experiment results showed that if the percentage of MB increased, a linear increase in absorbance intensity at wavelength of 663 nm occurred. The smartphone-based sensor showed a comparable sensitivity with that of a laboratory UV VIS spectrometer. With the noticeable advantages such as small, low-cost, portability, and high accuracy, the smartphone-based sensors can be effectively used in field trip for fast detection of toxic dyes in waist of textile industry.

Acknowledgments

This research was supported by the International Foundation for Science (IFS), Stockholm, Sweden, and by the Organization for the Prohibition of Chemical Weapons (OPCW), through a grant to Dr. Hanh Hong Mai. Grant NO. I-2-W-6258-1.

REFERENCES

1. S.S.S. Ng, T.-O. Chan, K.-W. To, J. Ngai, A. Tung, F.W.S. Ko, D.S.C. Hui (2010), "Validation of Embletta portable diagnostic system for identifying patients with suspected obstructive sleep apnoea syndrome (OSAS)", Respirology. 15 336–342.
2. K. Dingli, E.L. Coleman, M. Vennelle, S.P. Finch, P.K. Wraith, T.W. Mackay, N.J. Douglas (2003), "Evaluation of a portable device for diagnosing the sleep apnoea/hypopnoea syndrome", Eur. Respir. J. 21 253–259.
3. J.G.E. Gardeniers, A. van den Berg (2004), "Lab-on-a-chip systems for biomedical and environmental monitoring", Anal. Bioanal. Chem. 378 1700–1703.
4. I. Navruz, A.F. Coskun, J. Wong, S. Mohammad, D. Tseng, R. Nagi, S. Phillips, A. Ozcan (2013), "Smart-phone based computational microscopy using multi-frame contact imaging on a fiber-optic array", Lab Chip. 13 4015–4023.
5. H. Yu, Y. Tan, B.T. Cunningham (2014), "Smartphone fluorescence spectroscopy", Anal. Chem. 86 8805–8813.
6. M.A. Hossain, J. Canning, K. Cook, A. Jamalipthe (2015), "Smartphone laser beam spatial profiler", Opt. Lett. 40 5156.
7. Y. Wang, X. Liu, P. Chen, N.T. Tran, J. Zhang, W.S. Chia, S. Boujday, B. Liedberg (2016), "Smartphone spectrometer for colorimetric biosensing", Analyst. 141 3233–3238.
8. Y. Liu, Q. Liu, S. Chen, F. Cheng, H. Wang, W. Peng (2015), "Surface Plasmon Resonance Biosensor Based on Smart Phone Platforms", Sci. Rep. 5 12864.
9. K. Bremer, B. Roth (2015), "Fibre optic surface plasmon resonance sensor system designed for smartphones", Opt. Express. 23 17179.
10. Y. Zhuo, B.T. Cunningham (2015), "Label-free biosensor imaging on photonic crystal surfaces", Sensors (Switzerland). 15 21613–21635
11. D. Gallegos, K.D. Long, H. Yu, P.P. Clark, Y. Lin, S. George, P. Nath, B.T. Cunningham (2013),

"Label-free biodetection using a smartphone", Lab Chip. 13 2124.
12. V. Dantu, J. Vempati, S. Srivilliputhur (2014), "Non-invasive blood glucose monitor based on spectroscopy using a smartphone", Conf Proc IEEE Eng Med Biol Soc. 3695–3698.
13. M.A. Hossain, J. Canning, S. Ast, K. Cook, P.J. Rutledge, A. Jamalipthe (2015), "Combined "dual" absorption and fluorescence smartphone spectrometers", Opt. Lett. 40 1737–1740.
14. S. Dutta, D. Sarma, P. Nath (2015), "Ground and river water quality monitoring using a smartphone-based pH sensor", AIP Adv. 5 1–10.
15. S.S. Auerbach, D.W. Bristol, J.C. Peckham, G.S. Travlos, C.D. Hébert, R.S. Chhabra (2010), "Toxicity and carcinogenicity studies of methylene blue trihydrate in F344N rats and B6C3F1 mice", Food Chem. Toxicol. 48 169–177.
16. A.F. Hassan, A.M. Abdel-Mohsen, M.M.G. Fouda (2014), "Comparative study of calcium alginate, activated carbon, and their composite beads on methylene blue adsorption", Carbohydr. Polym. 102 192–198.
17. A. Petzer, B.H. Harvey, G. Wegener, J.P. Petzer, Azure B (2012), "A metabolite of methylene blue, is a high-potency, reversible inhibitor of monoamine oxidase", Toxicol. Appl. Pharmacol. 258 403–409.
18. L.-X. Sun, A.M. Reddy, N. Matsuda, A. Takatsu, K. Kato, T. Okada (2003), "Simultaneous determination of methylene blue and new methylene blue by slab optical waveguide spectroscopy and artificial neural networks", Anal. Chim. Acta. 487 109–116.
19. N. Bélaz-David, L.A. Decosterd, M. Appenzeller, Y.A. Ruetsch, R. Chioléro, T. Buclin, J. Biollaz (1997), "Spectrophotometric determination of methylene blue in biological fluids after ion-pair extraction and evidence of its adsorption on plastic polymers," Eur. J. Pharm. Sci. 5 335–345.
20. D. Melgoza, A. Hernández-Ramírez, J.M. Peralta-Hernández (1999), "Comparative efficiencies of the decoltheisation of Methylene Blue using Fenton's and photo-Fenton{'}s reactions", Photochem. Photobiol. Sci. 8 596–599.

PHÁT TRIỂN CẢM BIẾN QUANG HỌC SỬ DỤNG ĐIỆN THOẠI THÔNG MINH ỨNG DỤNG TRONG XỬ LÝ NƯỚC THẢI MÔI TRƯỜNG

Tóm tắt: *Trong những năm gần đây, các thiết bị sử dụng điện thoại thông minh đang được phát triển mạnh mẽ. Trong công trình này, cảm biến quang học sử dụng điện thoại thông minh đã được nghiên cứu và phát triển. Cảm biến bao gồm một nguồn sáng, một thấu kính chuẩn trực, một cách tử nhiễu xạ và sử dụng chip CMOS của điện thoại thông minh thay cho detector. Cảm biến có thể đo được phổ hấp thụ, và phổ truyền qua trong dải bước sóng khoảng 300 nm (từ 400 đến 700nm) và với độ phân giải 0,26nm/pixel. Thiết bị cũng được sử dụng để đo nồng độ xanh methylene (MB), một loại thuốc nhuộm trong nước thải từ ngành dệt may. Mặc dù có cấu tạo đơn giản và được chế tạo với chi phí thấp nhưng cảm biến cho thấy kết quả đáng tin cậy, có thể so sánh được với thiết bị trong phòng thí nghiệm.*

Từ khoá: *Cảm biến quang học, điện thoại thông minh, phổ hấp thụ, xanh methylene.*

ABOUT A MULTICLASS TRAFFIC FORECASTING MODEL OF HANOI

Pham Xuan Hinh

Hanoi Metropolitan University

Abstact: *Multiclass traffic forecasting problem is one of the essential problems in programming and developing traveling system today. In this paper, we represent one multiclass traffic forecasting model using to forecast transportation of many classes of objects travelling. Studying this model maybe useful in setting up and organizing the interprovince traveling system in the Hanoi.*

Keywords: *Forecasting model using to forecast transportation.*

Received: 18 March 2020

Accepted for publication: 20 April 2020

Email: pxhinh@daihocthudo.edu.vn

1. INTRODUCTION

Hanoi city has now been expanded into a large enough area with a large population. Transport infrastructure connecting urban districts and suburban districts has been relatively completed. The problem that we need to construct a usable transport network for Hanoi city leading to the development of the city. In order to have a good transport network, in addition to invest in consolidating, upgrading and renewing traffic routes as well as purchasing equipment and facilities,... an important step is to come up with a network organization model which is a optimal interdistrict and inter-provincial transportation network to serve the city's socioeconomic development needs. In order to build an optimal transportation network, we have to conduct surveys on roads, traffic and goods, etc. at the central points (districts , towns, townships). This is a relatively difficult work, requiring a lot of time, effort and money. Due to the growing demands of the transportation network, in order to have a good transport network and can be used for a long time, we must predict the flow of passengers, goods and other means at central points and traffic flows of every area. Then we can make plan developing transportation network for the next years, so that it meets the needs and demands of current and future economic and social development of each region. In recent years, when researching on traffic, people are often interested in choosing the mode of transportation of the participants and dividing these objects into classes. Road

users have the right to choose their own route, or follow a predetermined route in the most efficient way. Thus, objects in traffic can travel in two ways, or according to a predetermined, pre-determined route; or follow an arbitrary route. We divide the objects of traffic in classes, depending on purposes and travel needs of those objects. For simplicity, we divide those objects into 2 classes, the first class is the objects participating in traffic according to the scheduled time (for example: Officials to work, students, students coming to school,...), the second class is the other travel and truck flows. On each class of objects participating in traffic, there are different ways to go. When traffic participants find their own unique route, they themselves form the flow of traffic on the network. This transportation network will be optimized according to the *Wardrop* 's variational principle. If we can predict the size of these objects between two points which are the source and target, predict form, which class, which route these objects follow, then we can make a plan and organize the optimal transport network of Hanoi city for the present and the future. Therefore, in this paper we present a mathematical model that predicts multi-class travel of traffic participants, combining the starting point (source) - the ending point (target) and the way to choose the mean of travel, the path of those objects. The solution to the problem helps us to plan as well as create an effective transport model, meeting the needs and requirements for the development of Hanoi city.

2. CONTENT

2.1. Mathematical model

Symbols:

l is the class of objects participating in traffic.

m is the way (type of travel) of objects participating in traffic.

d^l_{pqm} is the flow of people from starting area p to area ending q by way m of class l.

O^l_p is the total flow of guest of class l originating from area p.

D^l_q is the total flow of guest of class l go to from area p.

R_{pq} is the set of splicing routes go from area p to area q.

A is the way of the object to join the traffic without being bound to the route.

T is the way of the road participants according to a predetermined route.

$h^l_r, r \in R_{pq}$ is the flow of *(A)* of class l from area p to area q on route r of the network.

K_{pq} The fixed truck flow *(T)* from area p to area q is represented by h^k_r .

v^l is the number of people per vehicle, it denotes the relationship between the flow of people and the flow *(A).*

In this article we only consider two ways of going (A) and (T) and they are considered

to be operating independently on the network, so the total cost corresponding to each way is separate. The cost of taking *T* is the cost on the routes connecting the area *p* and area *q which* is fixed and shown by the timetable and fare schedule.

We consider the road network to be a set of points *(N)* and roads *(I)*; The Northern Delta region is divided into several provinces, denoted by *(Z).* The total cost of online travel is a weighted, linear function for the time passengers stay in the car, the time passengers spend outside the car of travel, the cost of money as well as the length of the routes (A) . The time passengers sit in the vehicle of each ramp of the road network is an increasing function of the total flow through itself. The time passengers spend outside the vehicle on the road network is the time of entry and exit at the starting and ending points. The monetary costs incurred on the road network are the ramp and parking fees at the end. The vehicle operating cost for a ramp is assumed to be a linear function of the time traveled on the ramp (in minutes) and the length of the ramp (in kilometers). Travel time and length of the distance are variables of the target function, they affect both: individual travel and operating costs related to (*A)* , the impact of these variables on The objective function is represented by the rating factor. The monetary cost of *(T)* is the fare of *(T).* We define the variables related to cost as follows:

$t_a(f_a)$ is the time in the vehicle when the vehicle is traveling on the connecting road *a* of (*A), which* is the function of the total flow f_a *(minutes).*

k_a the fee for each vehicle (A) on the ramp a or parking fees at the end ramp *(thousand VND).*

s_a is the length of the connecting road a (Km).

$\delta_r^a = 1$ if road a is on route r and is zero in other cases.

$w_{pq,au}$ is the time to travel outside the vehicle by (A) from area p to area q (minutes).

$t_{pq,tr}$ is the time in vehicle of (T) from area p to area q (minutes).

$k_{pq,tr}$ the fare (T) from the area p to the area q (thousand VND).

$w_{pq,tr}$ is the time for going outside of vehicle (T) from area p to area q (minutes).

And other symbols:

$\gamma_1^l = 1$ is the coefficient corresponding to the travel time in (A), of class l.

γ_2^l is the coefficient corresponding to the monetary cost of (A), of class l.

γ_3^l is the coefficient corresponding to the time spent outside the car of (A), of class l.

γ_4^l is the coefficient corresponding to the travel distance of (A), of class l.

γ_5^l is the coefficient corresponding to the travel time in (T), of class l.

γ_6^l is the coefficient corresponding to the fare of (T), of class l.

γ_7^l is the coefficient corresponding to the time spent outside the car of (T), of class l.

γ_8^l is the coefficient corresponding to the waiting time for the transportation of class I

participants .

$\mu^l$is the cost-sensitive parameter of class l.

We have the following **mathematical model:**

$$minT(d,h) = \sum_a \int_0^{f_a} t_a(x)dx + \sum_{la} \left(\gamma_2^l f_a^l k_a + \gamma_4^l f_a^l s_a\right) + \sum_{lpq} \frac{\gamma_2^l}{v^l}. d_{pq,au}^l w_{pq,au}$$
$$+ \sum_{lpq} \left(\frac{d_{pq,tr}^l}{v^l}(\gamma_5^l t_{pq,tr} + \gamma_5^l t_{pq,tr} + \gamma_5^l t_{pq,tr} + \gamma_5^l)\right) + \sum_{lpqm} \frac{1}{\mu^1 v^1} d_{pqm}^l (ln d_{pqm}^l - 1)$$

With binding conditions:

$$\sum_{r \in R_{pq}} h_r^l = \frac{d_{pqau}^l}{v^l}; p, q \in Z; l \in L$$

$$\sum_{r \in R_{pq}} h_r^k = K_{pq}; p, q \in Z$$

$$\sum_{qm} d_{pqm}^l = \mathrm{O}_p^l; p \in Z; l \in L$$

$$\sum_{pm} d_{pqm}^l = D_q^l; q \in Z; l \in L$$

$$h_r^l \geq 0, r \in R_{pq}, p, q \in Z; l \in \mathrm{L}$$

$$h_r^k \geq 0, r \in R_{pq}, p, q \in Z$$

Here

$$f_a = \sum_l f_a^l + f_a^k = \sum_{lr} h_r^l \delta_r^a + \sum_{lr} h_r^k \delta_r^a \ , a \in A$$

The objective of the problem is to minimize the total cost of transport on the network. From the solution of the problem, we will know how to determine the capacity of the objects in traffic at both source and target points, from which the route and the type of route can be selected.

2.2. Preliminary solution

We construct the *Lagrange* function as follows:

$$L(d,h) = T(d,h) - \sum_{lpq} u_{pq}^l \left(\sum_{r \in R_{pq}} h_r^l - \frac{d_{pq,au}^l}{v^l}\right) - \sum_{pq} u_{pq}^k \left(\sum_{r \in R_{pq}} h_r^k - K_{pq}\right)$$
$$- \sum_{lp} \alpha_p^l \left(\sum_{qm} d_{pqm}^l - \mathrm{O}_q^l\right) - \sum_{lq} \beta_q^l \left(\sum_{pm} d_{pqm}^l - D_q^l\right.$$

Here $u_{pq}^l, u_{pq}^k, \alpha_p^l, \beta_q^l$ are the *Lagrange* agents with corresponding constraints.

Taking partial derivative according to the flow variable of *(A)* of *Lagrange* function, we achieve the following optimal conditions:

$$\sum_a t_a(f_a)\delta_r^a + \gamma_2^1 \sum_a k_a \delta_r^a + \gamma_4^1 \sum_a s_a \delta_r^a - u_{pq}^l \geq 0, pq \in Z; l \in L$$

$$h_r^l(\sum_a t_a(f_a)\delta_r^a + \gamma_2^1 \sum_a k_a \delta_r^a + \gamma_4^1 \sum_a s_a \delta_r^a - u_{pq}^l) = 0, pq \in Z; l \in L$$

Similarly, we also achieve the optimal condition for the flow (T).

We define the total cost of *(A)* of route *r* of class *l* as:

$$c_r^l = \left(\sum_a t_a(f_a)\delta_r^a + \gamma_2^1 \sum_a k_a \delta_r^a + \gamma_4^1 \sum_a s_a \delta_r^a\right)$$

If $h_r^l \geq 0, r \in R_{pq}, c_r^l = u_{pq}^l$. That means, if all the flow from p to q is positive flow of class l then the travel cost of the total flow of (A) is equal.

If $h_r^l = 0, r \in R_{pq}, c_r^l \geq u_{pq}^l$. That means, if the flow from p to q is flow 0 in class l, then the total travel cost of (A) is not less than the other flows. (These conditions derive from *Wardrop* 's variational principle).

The unit of overall cost of *(A)* is calculated in minutes of the vehicle on the road. Its corresponding coefficient is the number of minutes in the vehicle on the road per corresponding unit of variable (cents, number minutes outside the car, km,...). The monetary cost factor of *(A)* takes into account the occupancy, the length of the travel that adversely affects travel and the monetary cost of activity.

The coefficient of the time in the vehicle of the travel on the road in the target function is equal to 1.

The time spent on the ramp depends on the total coupling flow, applied for two ways *(A)* and (*T*), then take the partial derivative of the O-D variable of the *Lagrange* function, we obtain:

(A): $$\left(\frac{\gamma_3^l}{v^l}\right) w_{pq,au} + \left(\frac{1}{\mu^l v^l}\right) ln\ d_{pq,au}^l + \left(\frac{u_{pq}^l}{v^l}\right) - \alpha_p^l - \beta_q^l = 0, pq \in Z;\ l \in L$$

(T): $$\frac{1}{v^l}\gamma_5^l t_{pq,tr} + \gamma_6^l k_{pq,tr} + \gamma_7^l w_{pq,tr} + \gamma_8^l + \left(\frac{1}{u^l v^l}\right) ln\ d_{pq,au}^l - \alpha_p^l - \beta_q^l = 0, pq \in Z;\ l \in L$$

From the above two equations we have:

$$d_{pq,au}^l = \exp \mu^l v^l\ \alpha_p^l + \beta_q^l - \mu^l\ \gamma_3^l w_{pq,au} + u_{pq}^l$$

$$d_{pq,tr}^l = \exp \mu^l v^l\ \alpha_p^l + \beta_q^l - \mu^l\ \gamma_5^l t_{pq,tr} + \gamma_6^l k_{pq,tr} + \gamma_7^l w_{pq,tr} + \gamma_8^l$$

We put

$$c_{pq,au}^l = u_{pq}^l + \gamma_3^l w_{pq,au}^l$$

$$c_{pq,tr}^l = \gamma_5^l t_{pq,tr}^l + \gamma_6^l k_{pq,tr}^l + \gamma_5^l w_{pq,tr}^l + \gamma_8^l$$

Thence inferred:

$$d_{pq,au}^{l} = A_p^l O_p^l B_q^l D_q^l \exp -\mu^l c_{pq,au}^l$$
$$d_{pq,tr}^{l} = A_p^l O_p^l B_q^l D_q^l \exp -\mu^l c_{pq,tr}^l$$

The *Lagrange* factors are α_p^l, β_q^l; moreover, A_p^l, B_q^l are defined as:

$$1/\ A_p^l = \sum_q B_q^l D_q^l . \exp -\mu^l c_{pq,au}^l + \exp -\mu^l c_{pq,tr}^l$$
$$1/\ B_q^l = \sum_p A_p^l O_p^l . \exp -\mu^l c_{pq,au}^l + \exp -\mu^l c_{pq,tr}^l$$

3. CONCLUSION

The above text is a traffic forecasting problem model that can be applied to Hanoi City. Of course, there are other models related to this problem. Choosing which model is suitable for our processing ability and practical situation can only be solved on the basis of specific calculations. With the above model, this is a nonlinear planning problem with a very large number of variables, thoroughly solving this problem is still difficult. However, with the current development of computer technology, we hope to solve this problem thoroughly, and can create a good software to implement this model for traffic of Hanoi city.

REFERENCES

1. David Boyce , Hillel Bar-Gera (2003), Validation of multiclass urban travel forecasting models combining origin - destination, mode and route choices, *Journal of Regional Science*, 43.
2. Hillel Bar-Gera, David Boyce (2003), *Origin - based Algorithms for Combined Travel Forecasting Models*, Transportation Research Part B, 37, p. 405-422.

VỀ MỘT MÔ HÌNH DỰ BÁO GIAO THÔNG ĐA THÀNH PHẦN CHO THÀNH PHỐ HÀ NỘI

Tóm tắt: *Để giúp cho việc hoạch định chiến lược phát triển mạng lưới giao thông nội đô và các vùng ngoại thành được tốt thì chúng ta phải dự báo được các thành phần tham gia giao thông của thành phố Hà Nội. Trong bài báo này, chúng tôi trình bày một mô hình dự báo giao thông đa thành phần, dự báo sự đi lại của nhiều lớp đối tượng tham gia giao thông. Việc nghiên cứu mô hình này có thể sẽ là hữu ích cho việc thiết lập và tổ chức mạng lưới giao thông liên quận huyện cho thành phố Hà Nội.*

Từ khóa: *Mô hình dự báo giao thông đa thành phần*

DEVELOPING RSA AND RABIN SIGNATURE SCHEMES IN CASE OF EXPONENT E=3

Hoang Thi Mai, Le Chi Chung, Ngo Van Than

Hanoi Metropolitan University, Hanoi University of Culture

Abstract: *The RSA and Rabin signature schemes are both developed based on the difficulty of the factorizing problem. While the exponent e in the RSA scheme has to satisfy gcd(e,ϕ(n)) = 1, in the Rabin scheme, e=2 and is always the divisor of ϕ(n). On solving the problem of constructing a signature scheme with low signature-verifying cost for digital transaction that require authentication of signature validity in a great deal, this study suggests a signature schemes base on the graphical model in case of exponent e= 3 and 3 is the divisor of ϕ(n). This scheme is similar to the Rabin scheme, with e=3 as the divisor of (p-1) and (q-1). With exponent e=3, the schemes have low signature-verifying cost, which meet the requirement of the problem above.*

Keywords: RSA Signature Scheme, Digital Signature Scheme, Rabin Signature Scheme, Cube Root Signature Scheme

Received: 19 March 2020

Accepted for publication: 20 April 2020

Email: htmai@hnmu.edu.vn

1. INTRODUCTION

In digital transactions, there are many activities that require authentication of signature validity in a great deal, such as the profile admission of public administration service; activities in which authentication is considered compulsory, such as inspecting digital certification. Therefore, in order to use digital signature efficiently, signature-verifying algorithm consuming little time needs to be applied.

Evaluating the digital signature systems that have been included as standard RSA (Rivest–Shamir–Adleman-the first public key cryptosystems), discrete logarithms (DL) and elliptic curves (EC) with input parameters of the same level of safety given in Figure 1 (cited from FIP 186-2 [1], NIST 800-56 [2] and presented by Darrel Hankerson in [3]), Lenstra and Verheul turned out that in case of a small exponent (e = 3), RSA was more efficient than EC and DL systems [4], [5].

According to the RSA signature scheme with small exponent, in publication [6] in 1979, M.O Rabin proposed a key signature scheme with an exponent e = 2. The choice of e = 2 has brought the outstanding advantage that Rabin's signature-verifying algorithm only requires a modulo squared operation. In the RSA scheme, with $n = p \times q$ (p, q are primes), we have $gcd(e, \phi(n)) = 1$ ($with\ \phi(n) = (p-1) \times (q-1)$). In the Rabin scheme, by choosing e=2, we clearly have e as the division of $\phi(n)$, which means the parameters p, q satisfy $p \equiv q \equiv 1\ (mod\ 2)$. When proposing a signature scheme [6], Rabin noted that if there is $p \equiv q \equiv 1\ (mod\ 3)$, it is possible to replace the congruent quadratic equation with the congruent cubic equation to produce a signature scheme with safety ensured by the factorizing problem.

	Security level (bits)				
	80 (SKIPJACK)	112 (Triple-DES)	128 (AES-Small)	192 (AES-Medium)	256 (AES-Large)
DL parameter *q* EC parameter *n*	160	224	256	384	512
RSA modulus *n* DL modulus *p*	1024	2048	3072	8192	15360

Figure 1. *Security level*

In 1980, Williams improved the version of the Rabin scheme, abbriaviated to RW [7]. This scheme only required one Jacobi symbol calculation while the Rabin scheme required 4 Jacobi symbol calculations for creating the signatures. With such ultimate feature, it was brought into ISO/IEC 9796 standard in 2002 [8]. In 1986, based on the improvements made on the Rabin scheme in the RW scheme [7], Williams proposed a RW-based signature scheme with an exponent e = 3.

Apart from the research of Williams [7] [9] [10], there were numerous other research on improving the RSA scheme with small exponent, such as the authors J. H. Loxton, David S. P. Khoo, Gregory J. Bird and Jennifer Seberry [11], Scheidler [12],... These results constructed a version of the RSA scheme with e=3 and larger class of primes *p* and *q*.

In [13], we proposed a deterministic signature scheme combining RSA and Rabin for the case where *p - 1* is a multiple of 3 and *q - 1* is primate with 3. The proposed scheme cost little verification time as a 3 exponent modulo.

On solving the problem of constructing a signature scheme with low signature-verifying cost for digital transaction that require authentication of signature validity in a great deal, in this paper, the we propose a probabilistic signature scheme based on RSA and Rabin with the exponent e = 3. The organization of this paper is organized as follows. In section II, the paper presents the mathematical basis of the signature scheme as the open problem the square root on Z_p with p is the prime number larger than 3. In section III and IV, a probabilistic

signature scheme and the correctness and safety of the proposed scheme are presented. Finally, section V summarizes obtained results and future research.

2. CONTENT

2.1. Mathametic base

2.1.1. Symbols

With all a $\in \mathbb{Z}_n$ corresponding only with $(a_p, a_q) \in \mathbb{Z}_p \times \mathbb{Z}_q$ with $a_p = a \bmod p$ and $a_q = a \bmod q$ and reverse mapping, denoted as *CRT*, is determined by the formula:

$$CRT(u,v) = (q.(q^{-1} \bmod p).u + p.(p^{-1} \bmod p).v) \bmod n \quad (1)$$

- Mapping on the preservation of multiplication means:

$$CRT(u.x \bmod p, v.y \bmod q) = CRT(u,v).\ CRT(x,y) \bmod n \quad (2)$$

2.1.2. Some additive results.

a) Lemma 1: With the prime $\boldsymbol{p = t.k^s + 1}$ with gcd(t,k) = 1, denoted as :

$$u = -t^{-1} \bmod k; \quad (3)$$

Then d defined by

$$d = \frac{u(p-1) + k^s}{k^{s+1}}$$

is an integer.

Proof

As $p = t.k^s + 1$, we have:

$$d = \frac{u(p-1) + k^s}{k^{s+1}} = \frac{u.t.k^s + k^s}{k^{s+1}} = \frac{u.t + 1}{k}$$

According to (3):

$$ut = -t.t^{-1} \bmod k = -1 \bmod k$$

then

$$ut = xk - 1 \text{ with random integer x.}$$

Then, we have:

$$d = \frac{xk - 1 + 1}{k} = x$$

So, d is an integer.

b) The value p and d when k=3

In the case of s=1, then p = t.3 + 1 with gcd(t,3) = 1

We have t ≡ 1, 2 (mod 3).

More specifically, we have:

– With t ≡ 1 (mod 3) then p ≡ 4 (mod 9). According to (3), we have u ≡ 2 (mod 3)
– With t ≡ 2 (mod 3) then p ≡ 7 (mod 9). According to (3), we have u ≡ 1 (mod 3)

Then d is obtained by using the help of (4).

$$d = \begin{cases} \dfrac{2p+1}{9} \; with\; p \equiv 4 \pmod 9 \\ \dfrac{p+2}{9} with\; p \equiv 7 \pmod 9 \end{cases}$$

c) Definition 1 (Function *CR*, where the letters CR stand for "Cube Root")

Given p ≠ 1 (mod 9) as an odd prime, we have:

$$d = \begin{cases} 3^{-1} \bmod (p-1) & \text{nếu } p \neq 1 \pmod 3 \\ \frac{2p+1}{9} & \text{nếu } p \equiv 4 \pmod 9 \\ \frac{p+2}{9} & \text{nếu } p \equiv 7 \pmod 9 \end{cases} \quad (5)$$

Function CR (., p): GF(p) → GF(p) is determined by the following formula:

$$CR(a,p) = a^d \bmod p. \quad (6)$$

with GF(p), where the letters GF stand for "Galois field", is a finite field that is given by the integers mod p when p is a prime number.

Then, we have :

Lemma 2. With p ≠ 1 (mod 9) as an odd prime, then with a ∈ GF*(p) we have :

If $p \neq 1 \pmod 3$ then

$$CR(a,p)^3 \equiv a \pmod p. \quad (7)$$

If $p \equiv 4 \pmod 9$ then

$$CR(a,p)^3 \equiv a.\left(a^{\frac{p-1}{3}}\right)^2 \pmod p. \quad (8)$$

If $p \equiv 7 \pmod 9$ then

$$CR(a,p)^3 \equiv a.a^{\frac{p-1}{3}} \pmod p. \quad (9)$$

Lemma 3.

Considering the equation below with a ∈ $\mathbb{Z}_n$.

$$x^3 \equiv a \pmod{n}. \quad (10)$$

We have results as follows.

Conditions needed and sufficient for (10) to have a solution:

$$a^{\frac{p-1}{3}} \bmod p = 1 \quad (11)$$

Then, a solution of (10) is given by the following formula:

$$x = \text{CRT}(\text{CR}(a \bmod p, p), \text{CR}(a \bmod q, q)). \quad (12)$$

Corollary 1. If n can be analyzed into factors p and q, then equation (10) always be solved.

To the best of authors knowledge, there has not been any publication that indicates a solution of (10) can be found without knowing the analysis of n. In contrast, there has not been a claim that n can be analyzed if the equation (8) can be solved. Here, this paper would give a possibly closest result that can solve the opposite problem as follows.

Clause 1. If two different solutions of equation (10) are found, then n can be analyzed.

Time cost for arthmetic operation on Z_n.

The cost of running time some algorithms performing arithmetic operations:

(1) The cost of adding or subtracting two k-bits is O (k) [14, pp. 30-31].

(2) Multiplying two 2.k-bits integers by the method of Karatsuba-Ofman requires three k-bit double multiplications [14, p. 51]. At this time the multiplication cost, denoted by M, and the cost for squaring two large numbers, denoted by S, are approximately equal (M ≈ S). Moreover, we also get the cost for multiplication:

$$M = O\left(k^{ln3/ln2}\right).$$

Formula 13. *M (k) is the computation cost to perform a multiplication of two k-bit integers. Then with all positive integers h, k we have:*

$$M(h) \approx 3^t . M(k), \text{with } t = \log_2\left(\frac{h}{k}\right). \quad (13)$$

A 2.k-bit truncation algorithm in one modulo k-bit using Barrett's math requires two multiplication of k-bit numbers [14, p. 36]. Infer the value $t = \log_2\left(\frac{h}{k}\right) = 1$.. Therefore, the result is:

Formula 14: *The cost of implementing a reduced multiplication in modulo n is approximately* (14)

(3) Jacobi symbol-calculating algorithm for a number modulo k-bit base on the rule of reciprocal square has complexity as $O(k^2)$ [15, p. 98].

(4) The calculation cost of the inverse of a number in modulo k-bit, denoted by I, and the cost of dividing a number by modulo k-bits, denoted by D, according to Sorenson's algorithm has complexity O (k ^ 2 / lnk) [15, pp. 463-465].

(5) According to formula (1), performing the CRT function requires two inverse modulo p and a division modulo n. According to the results shown in Table 2.1, the cost of the inverse modulo k - bit is equal to the cost of a division modulo k - bit and that cost is k^2/lnk. So the time cost for a CRT function calculation is $3.k^2/lnk$.

The time cost for arithmetic operations on Zn is summarized in the following table.

Table 1. The cost of runtime of arithmetic operations on Z_n

Operation	Complexity	Algorithm
The cost of adding or subtracting two k-bit integers	O(k)	[14, pp. 30-31]
The cost for multiplying two k-bit integers	$M(k) = O\left(k^{ln3/ln2}\right)$	[14, p. 51].
The cost for ashortened multiplication in modulo n	3.M(len(n))	[14, p. 36]
Calculating Jacobi symbol of a number in modulo k-bit according to the law of reciprocal squares	$O(k^2)$	[15, p. 98].
The calculation cost of the inverse of number modulo k-bits	$O(k^2/lnk)$	[15, pp. 463-465].
The cost for divide a number by modulo k-bits	$O(k^2/lnk)$	[15, pp. 463-465].

2.2. Signature scheme PCRS

As stated in the introduction, this paper proposes signature schemes that have a low cost for verifying algorithms for use in one-stop transactions. The Rabin and RSA schemes, with exponent e as small as possible, have the feature above. In this section, the paper proposes a probabilistic signature scheme, namely PCRS. Similar to the Rabin scheme, the parameters p, q of PCRS satisfy the condition $p \equiv q \equiv 1\ (mod\ 3)$.

2.2.1. Systematic parameter

System parameter for signature schemes includes:

– Integer n = p.q with p, q are two primes so that:

$$\begin{aligned} &– \quad p = 3.t + 1 \text{ with } gcd(t,3) = 1 \\ &– \quad and\ q = 3.k + 1 \text{ with } gcd(k,3) = 1 \end{aligned} \tag{15}$$

– *Hash* Function: $\{0,1\}^{\infty} \rightarrow \{0,1\}^{h}$ satisfies security requirements for codes.

– Secret parameter d_p, d_q can be defined as follows:

$$d_p = \begin{cases} \frac{2p+1}{9} \; if \; p \equiv 4 \pmod 9 \\ \frac{p+2}{9} \; if \; p \equiv 7 \pmod 9 \end{cases};$$

$$- \quad d_q = \begin{cases} \frac{2q+1}{9} \; if \; q \equiv 4 \pmod 9 \\ \frac{q+2}{9} \; if \; q \equiv 7 \pmod 9 \end{cases}; \tag{16}$$

2.2.2. Signing message.

Algorithm 1

Input: M ∈ $\{0,1\}^{\infty}$ is the message to be signed.

Output: (R,s) ∈ $\{0,1\}^{k} \times \mathbb{Z}_n$ is the signature in M.

1. Repeat

   R = Random($\{0,1\}^{k}$);

   h = Hash(R||M); (17)

   $$t = h^{\frac{p-1}{3}} \bmod p; u = h^{\frac{q-1}{3}} mod \; q;$$

   until (t==1) and (u==1)

2. $h_p = h \bmod p$; $h_q = h \bmod q$; (18)

4. $s_p = h_p^{d_p} \bmod p$; $s_q = h_q^{d_q} \bmod q$; (19)

5. $s = CRT(s_p, s_q)$; (20)

6. return (R, s); (21)

2.2.3. Verifying Signature.

Algorithm 2

Input: M ∈ $\{0,1\}^{\infty}$; (R,s) ∈ $\{0,1\}^{k} \times \mathbb{Z}_n$ is the signature in M.

Output: *Accept* ∈{0,1} only accept the validity of the signature if and only if *Accept* = 1.

1. h = Hash(R||M);
2. $t = s^3 \bmod n$;
3. Accept = (t==h);

   return Accept;

2.2.4. The correctness of the signature scheme

The correctness of the PCRS scheme is given by the following result:

Clause 2. *All signatures (R, s) on text M created from algorithm 1 have an output value of 1 according to algorithm 2.*

Proof:

According to step 1 of algorithm 1 that creates a signature, we have t = 1 and u = 1 so from the formula (17) we have:

$$- \quad h^{\frac{p-1}{3}} \bmod p = 1 \tag{22}$$

Since h satisfies (22), the equation: $s^3 \equiv h \pmod n$ always has solution $s = CRT(s_p, s_q)$

According to formula (19) we have:

$$s_q^3 \bmod q = (h_q^{d_q})^3 \bmod q = (h_q^{3d_q}) \bmod q$$

And :

$$s_p^3 \ mod\ p = (h_p^{d_p})^3 mod\ p = (h_p^{3d_p}) mod\ p$$

– If p ≡ 4 (mod 9) then $3d_p = \frac{2p+1}{3} = 1 + 2\frac{p-1}{3}$, so:

$$s_p^3 \ mod\ p = (h_p^{3d_p}) mod\ p = h_p^{\left(1+2\frac{p-1}{3}\right)} mod\ p$$
$$= h_p . h_p^{2\frac{p-1}{3}} mod\ p = h_p . \left(h_p^{\frac{p-1}{3}}\right)^2 \bmod p = h_p$$

– If p ≡ 7 (mod 9) then $3d_p = \frac{p+2}{3} = 1 + \frac{p-1}{3}$, so:

$$s_p^3 \bmod p = (h_p^{3d_p}) \bmod p = h_p^{\left(1+\frac{p-1}{3}\right)} \bmod p$$
$$= h_p . h_p^{\frac{p-1}{3}} \bmod p = h_p$$

Similarly, we have: $s_q^3 \bmod q = h_q$

So:

$$\begin{cases} s_p^3 \bmod p = h_p \\ s_q^3 \bmod q = h_q \end{cases}$$

From formula (20) we have:

$$s^3 = \left(CRT(s_p, s_q)\right)^3 = CRT(s_p^3 \ mod\ p, s_q^3 \bmod q)$$
$$= CRT(h_p, h_q) = h \tag{23}$$

From the result gained from (23) we have a comparison (t == h) in step 3 of the PCSR1 signature-verifying algorithm that always returns a result of 1 (true). Therefore, *Accept =1*. This is what needs to be proven.

2.3. Time cost to run the PCRS schemes

2.3.1. Calculation cost of PCRS scheme

The cost of the signing algorithm is calculated based on the steps to conduct the algorithm and the size of the input parameters. Therefore, with size of modulo p guaranteeing security given in figure 1, we consider the size of parameters d_p, d_q based on formula (19)

According to formula (19), we have:

– With $p \equiv 4 \pmod 9$ then $d_p = \frac{2p+1}{9} < \frac{p}{4}$. Thus, the size of d_p is smaller than the size of modulo p at least 2 bits, which means $len(d_p) \leq \text{len}(p) - 2$.

– With $p \equiv 7 \pmod 9$ then $d_p = \frac{p+2}{9} < \frac{p}{8}$. Thus, the size of d_p is smaller than the size of modulo p at least 3 bits, which means $\text{len}(d_p) \leq \text{len}(p) - 3$

– We have equivalent result with parameter d_q.

– So, the size of d_p and d_q is 2 to 3 bits smaller than the size of p, q. To simplify the calculation while guarantee the validity of the calculation of time cost for running algorithm, we can choose the case of maximum length of d_p and d_q , which is equal to $\text{len}(p)$.

Considering the calculation cost of the signature-creating algorithm (algorithm 1)

– In step 1, the algorithm executes the loop with the stop condition t = 1 and u = 1 with the power of h being the cube root of the unit. Since the probability of finding a third-degree surplus is $\frac{1}{3}$, so in loop 1 of algorithm 1 we need to do it in the $3^2 = 9$ times of the exponentiation. We denote t_{exp} as the time cost for a power calculation, then the time cost for step 1 is approximately $9.\, t_{exp}$

– In step 2, two modulo are operated so the cost is $2.\ln(\text{len}(p))$.

– In step 3, two exponentiations are operated so the cost is $2.\, t_{exp}$

– In step 4, a CRT function calculation is needed, we denote it as t_{CRT}.

Considering the computational cost of the signature-verifying algorithm (algorithm 2)

– In step 2, a power of 3 on $\mathbb{Z}_n$ is required (by 2 multiplications). We denote t_m as the time cost of performing a multiplication, then the time cost of step 2 is $2.\, t_m$.

From the above analysis, we obtain the time cost of the PCRS scheme as follows:

The cost of the signature creation algorithm, denoted as T_1, and the test algorithm, denoted as T_2, in the PCRS sheme are given by the following formula:

$$T_1 = 11 \times t_{exp} + 2.\ln\big(len(p)\big) + t_{CRT} \qquad (24)$$

$$T_2 = 2 \times t_m \quad (25)$$

According to the square-multiplication method, the average time cost for u exponentiation is calculated by multiplication $t_{exp} = 1.5\ len(u).t_m$.

Moreover, according to the results given in table 1, the cost of multiplying two integers of which lengths is k - bit is $k^{ln3/ln2}$. Therefore $t_m = k^{ln3/ln2}$.

According to formula (1), performing the CRT function requires two inverses of modulo p and a division by modulo n. According to the results shown in table 1, the cost of the inverse in modulo k - bit is equal to the cost of a division by modulo k - bit and that cost is k^2/lnk. So, the time cost for a CRT function calculation is $3.k^2/lnk$.

With the security-guaranteed size modulo p given in figure 1 and the parameter d_p given by the formula (19), we have the same size parameter d_p of size modulo p.

Clause 4. *The cost of the signature-creation algorithm, denoted as T_1,, and the verifying algorithm, denoted as T_2, in the PCRS scheme are given by the following formula:*

$$T_1 = 16.5 \times \text{len}(p) \times \text{len}(p)^{\frac{ln3}{ln2}} + 2.\ln\big(\text{len}(p)\big) + \frac{3\text{len}(p)}{\ln \text{len}(p)}$$

Then:

$$T_1 = 16.5 \times \text{len(p)}^{(\frac{\ln 3}{\ln 2}+1)} + 2.\ln\big(\text{len}(p)\big) + \frac{3\text{len}(p)}{\ln \text{len}(p)} \quad (26)$$

And:

$$T_2 = 2 \times \text{len}(p)^{\frac{ln3}{ln2}} \quad (27)$$

2.3.2. The effectiveness of the proposed schemes

The PCRS scheme is probabilistic, combined with the principles of the RSA and Rabin schemes in case of e=3.

In loop 1 of algorithm 1, we need to perform averagely $3^2 = 9$ times (because the probability to find a third-degree surplus is $\frac{1}{3}$). In general, with a similar development of Rabin for verifying exponent that is prime e, the complexity of the signature-creating algorithm will require averagely e^2 times in the loop to find an e-degree surplus. Thus, the PCRS scheme's signature-creating algorithm has greater complexity than Rabin's because the Rabin scheme needs to perform averagely four times in the loop.

2.4. Security of the schemes PCRS

To create a valid signature onto document M depending on the scheme PCRS, a counterfeiter has two choices:

Firstly, from the public parameter n, the counterfeiter has to find two prime numbers p, q with n=p.q. With p and q known, dp and dq are found, therefore counterfeit signature can be created. To do this, the counterfeiter has to solve the factorizing problem.

Secondly, if there is no secret parameter (p,q), the counterfeiter has to find a root of the cubic congruent equation depending on modulo n. The finding of a root of equation (10) is called "CRP-Cube Root Problem".

According to Consequence 1, if n if analyzed into factors p and q, then the cubic congruent equation can always be solved depending on modulo n. And according to Clause 1, if two different roots of the equation (10) are found, then n can be analyzed.

So, the security of the signature schemes is guaranteed by the difficulty of the factorizing problem and CRP problem. The factorizing problem was proven by digital theory to be difficult and is the mathematic base for a lot of the encryption systems. The CRP problem was proven by Consequence 1 to be led from the factorizing problem. This is why the signature scheme PCRS is safe.

3. CONCLUSION

In this paper, we have suggested a signature scheme based on the probabilistic model that is developed on the RSA and the Rabin scheme in case of e=3, in which the PCRS scheme is similar to the Rabin but with e=3. The PCRS is a combination between RSA and Rabin when e=3 is the divisor of p-1 but coprime with q-1. The significant features of the suggested schemes in this paper are algorithm creating signatures without calculating Jacobi symbol and signature-testing algorithm based on the Rabin scheme (which leads to its low cost). These diagrams have security based on the difficulty of analyzing n in the sense that if n can be analyzed, the schemes can be broken. So far, there is no effective algorithm (polynomial time) that breaks one of the signature diagrams mentioned above.

The signature schemes suggested in this paper guarantee safety and have low verifying cost. This is one of the vital features so that the digital signature schemes can be implemented in digital administrative and commercial services.

REFERENCES

1. FIP 186-2 (2000), Digital Signature Standard (DSS). Federal Information Processing Standards Publication 186-2. National Institute of Standards and Technolog.

2. NIST 800-56, NIST Special Publication 800-56: Recommendation on key establishment schemes, Draft 2.0.

3. Darrel Hankerson, Alfred Menezes, and Scott Vanstone (2004), *Guide to El-liptic Curve Cryptography*, Springer-Verlag.

4. A. LENSTRA (2001), "Unbelievable security—matching AES security using public key systems.," *Advances in Cryptology— A SIACRYPT 2001 (LNCS 2248),* vol. 67, pp. 67-86.

5. Arjen K. Lenstra and Eric R. Verheul (2001), "Selecting Cryptographic Key Sizes," *Journal of Cryptology.*, vol. 14, pp. 255-293.

6. M. O. Rabin (1979), "Digitalized signatures and Public-Key Functions as Intractable as Factorization," *Foundations of Secure Computations, Academic Press New York.*
7. H. C. Williams (1980), "A modification of the RSA public-key encryption procedure"," *IEEE Transactions on Information Theory ISSN 0018–9448,* vol. 26, pp. 726-729.
8. ISO/IEC 9796, "International Standard ISO/IEC 9796 Information technology – Security techniques- Information Security risk management".
9. ISO/IEC 9796, International Standard ISO/IEC 9796 Information technology – Security techniques- Information Security risk management.
10. H. C. Williams (1985), "Some public-key crypto-function as intractable as factorization," *Cryptologia,* vol. 9, no. 3, pp. 223-237.
11. H. Loxton, David S. P. Khoo, Gregory J. Bird and Jennifer Seberry (1992), "A Cubic RSA Code Equivalent to Factorization," *Journal of Cryptology,* vol. 5, pp. 139-150.
12. R. Scheidler (1998), "A Public-Key Cryptosystem Using Purely Cubic Fields," *Journal of Cryptology,* vol. 11, p. 109–124.
13. H. T. Mai (2018), "The Signature Scheme In Combination With RSA And RABIN," *Journal of Military Science and Technology,* vol. 53, pp. 143-148.
14. Darrel Hankerson, Alfred Menezes, and Scott Vanstone (2004), Guide to El-liptic Curve Cryptography, Springer-Verlag.
15. Richard Crandall, Carl Pomerance (2005), Prime Number A Computational Perspective, Springer.

PHÁT TRIỂN CHỮ KÝ SỐ RSA VÀ RABIN VỚI SỐ MŨ E=

Tóm tắt: *Sơ đồ chữ ký số RSA và sơ đồ chữ ký số Rabin đều là các lược đồ chữ ký được xây dựng trên cơ sở tính khó giải của bài toán phân tích số. Nếu như số mũ xác thực trong chữ ký số RSA là e phải thỏa mãn gcd(e,ϕ(n)) = 1 thì trong hệ Rabin e=2 và luôn là ước của ϕ(n). Theo hướng kết hợp giữa RSA và Rabin, bài báo đề xuất lược đồ chữ ký theo mô hình xác suất cho trường hợp số mũ xác thực e= 3 và 3 là ước của ϕ(n).*

Từ khóa**:** *RSA Signature Scheme, Digital Signature Scheme, Rabin Signature Scheme, Cube Root Signature Scheme.*

AN INTUITIVE SOFTWARE FOR TEACHING AND LEARNING CYBER ATTACKS

Tong Anh Tuan

People's Police University of Technology and Logistics

Abstract: *Currently, cybersecurity issues are of great concern, and at the same time, this major is being increasingly trained at universities in the country. In this article, we outline some common types of cyber-attacks such as SQL Injection, Cross Site-Scripting, Denial of Service, and introduce a software that simulates those attacks through objects, processes, and illustrations. This software supports teachers as well as students in the process of teaching and learning.*

Keywords: *Teaching tool, SQL Injection, Cross Site-Scripting, Denial of Service.*

Received: 26 March 2020

Accepted for publication: 21 April 2020

Email: tuanqb92@gmail.com

1. INTRODUCTION

At present, the information security field is increasingly interested in training at universities. This area covers the security of network and information infrastructure, computer security, data, and application software, as well as research and development of public products and technological solutions. According to the project, "Project of training and developing human resources for information security until 2020" sets out the task of training human resources at 08 key training institutions on information security and producing more than 2,000 high-quality information security personnel who graduate from universities or higher education institutions [1].

About information security teaching, students need to understand cyber-attack techniques in general and each stage and specific principles of each technique in particular. Typically, some common cyberattacks nowadays include SQL Injection (SQLi), Cross Site-Scripting (XSS), Denial of Service (DOS) [2]. To show the principles and stages of these techniques, the conventional method used is to present the theory, interpret steps through diagrams, drawings, or practice specific Lab exercises.

In this article, we introduce the Cyber Demonstration software, teaching, and learning support tool. Accordingly, the software will simulate and illustrate the principles and steps

of common attack techniques such as SQLi, XSS, or DOS. The structure of the paper is as follows: Part II, we summarize some of the current cyber attack techniques; In Part III, we introduce the Cyber Demonstration software with functions and interfaces and finally, Conclusion.

2. CONTENT

2.1. Some common types of cyber-attacks

2.1.1. SQL Injection attack

SQL Injection (SQLi) is a cyber-attack technique that targets databases. Attackers often manipulate middleware with the ability to interact with the database, taking advantage of vulnerabilities that do not fully control the input data on computer software to execute SQL statements on the basis. database [3]. These SQL statements can allow a hacker to bypass login, collect and destroy all data. This is an attack technique that disrupts the translation and can also expose or lose all data.

Figure 1. Illustrating the basic steps in the SQLi attack. Accordingly, hackers will interact with the database through the functions available on the website and execute dangerous SQL statements to exploit the system.

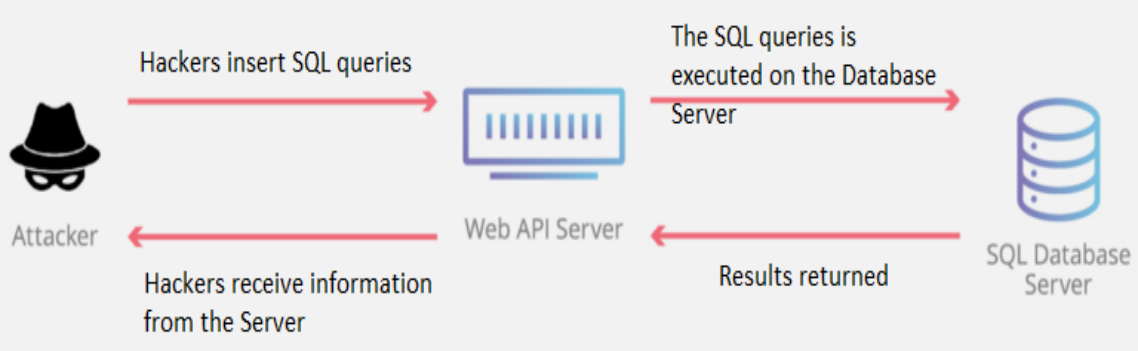


Figure 1. *Illustrating the basic steps in the SQLi attack*

SQLi attacks are classified into 4 primary groups as follows:

- Attack passing log;
- Attack using SELECT statement;
- Attack using INSERT statement;
- Attack using Stored Procedures.

SQLi vulnerabilities can be prevented by examining input data conditions, removing unusual characters and keywords, and limiting feedback from database servers.

2.1.2. Cross-Site Scripting Attack

Cross-Site Scripting (XSS) attack is a technique of inserting HTML tags or scripts that can be dangerous for users into websites. These include potential risks where users who may

have their Cookies stolen, their information typed or interacted on a fake Webpage [4]. Figure 2 depicts the basic steps in an XSS attack.

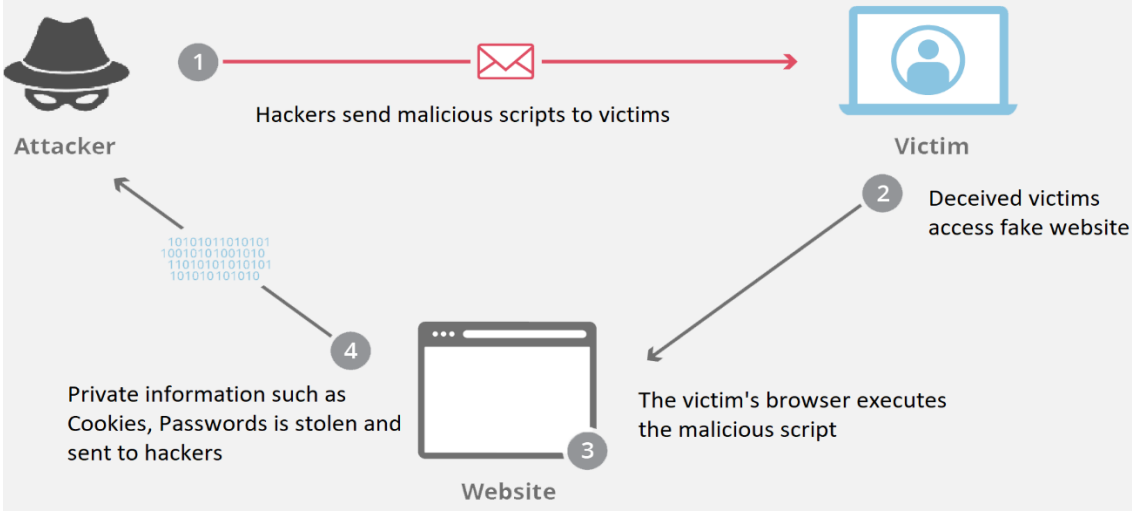


Figure 2. *Basic steps in XSS*

Unlike the SQLi vulnerability that affects database servers, the XSS vulnerability primarily affects Web users.

2.1.3. Password attacks

In most applications, information systems today, passwords are commonly used as an authentication method. Although more reliable authentication methods such as biometric authentication, asymmetric keys, OTP codes are being developed and applied more and more, password authentication still accounts for the majority.

Hackers are always looking for a user's password to gain access to their accounts. We have three techniques for discovering passwords:

- Brute Force: This is a technique that hackers will access the login interface of users, in turn, use the Username / Password pairs to check the probability of success [5]. In particular, the password is the sequence of characters that can be generated in turn in the password space. Hackers try every possible password in turn and stop when they find the right password. This technique has the advantage of undoubtedly yielding results; however, in the case of large password spaces, it takes much time to check, resulting in a lack of practical feasibility.

- Dictionary: This is a technique that tries passwords in turn, similar to Brute Force, but with a more limited password space. This space is made up of information about name, date of birth, personal information, or list of common passwords. Most users tend to set fairly common passwords or passwords related to personal information. Dictionary technology takes advantage of these trends to predict and find passwords faster than Brute Force. However, this technique does not always yield results.

- Keylogger: In this technique, hackers try to find the password from the data obtained from the user's keyboard [6]. To eavesdrop on keyboard data, the spyware will be secretly installed on the victim's device. When all keyboard data is recorded and sent to hackers, they have a high chance of filtering out passwords from this information.

2.1.4. Denial of service attacks

Denial of service is an attack in which hackers take advantage of the design of packets and protocols to execute unusual queries in large numbers, causing the recipient of that query to have resources exhausted, resulting in a usual inability to provide services.

Some techniques of denial of service attacks include SYN Flood, Smurft Attack, Teardrop, or ICMP Flooding [7].

- SYN Flood: In this technique, hackers take advantage of the loophole of the "three-step handshake" procedure in TCP communication to cause system flooding. Hackers will continuously send SYN packets to the receiving machine, receive SYN / ACK packets but do not respond with ACK packets as usual. This causes the victim machine to consistently allocate resources for these processes, resulting in exhaustion of system resources. 0illustrates the steps to attack the SYN Flood.

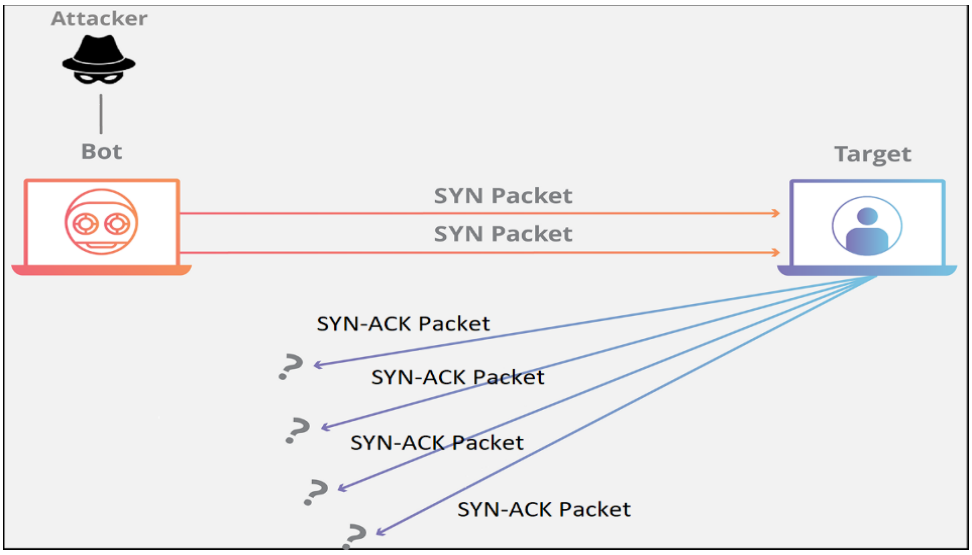


Figure 3. *Illustrates the steps to attack the SYN Flood*

- Smurf Attack: Hackers continuously send a large number of ICMP packets with the source IP address as the address of the victim machine to a network using a broadcast address, other computers in the network receiving ICMP messages will send back the response to the machine with the source IP address as the address of the victim machine; In the case of a vast number of computers in the network, the victim machine will be overloaded, resulting in the inability to operate normally.

- Teardrop attack: All data transmitted on the network to reach the target system must go through two processes: separation at source and reassembling at the destination. In the source system, the data will be divided into packets, each of which has a specific offset value to determine its position in the data. When these packets arrive at the destination, the system will rely on the value of offset to organize the packets to recreate the original data. In a Teardrop attack, hackers will send a series of packets with abnormal and non-renewable offset value to the target system. The destination system will not be able to rearrange the packets after receiving the packet; the system will spend much time to resolve, which results in flooding or errors.

In addition to the classic denial of service techniques, hackers now often use Botnets or servers on the Internet to increase the size of the attack, which is large enough to defeat the large systems in the network for a long time. DDOS -distributed denial of service and DRDOS- reflection denial of service are characteristic of these attacks.

2.2. Introducing the software supporting the training of cyber attack techniques

Besides traditional teaching methods, using whiteboard tools, chalk slides nowadays, new methods, using new tools are always encouraged to be applied to support learning, teaching, improve the vividness and visualization of the lessons.

In this section, we introduce the Cyber Demonstration software, a teaching aid tool built on the Windows operating system. Accordingly, the software will visually and vividly show the fundamental processes of a network attack technique such as SQLi, XSS ..., through scripts, objects, and motion effects designed based on the basis of the theory of that technique.

2.2.1. Platform and development process

Cyber Demonstration software is developed based on C # programming language, Microsoft .NET Framework 4.7.2 platform, Microsoft SQL Server 19 Database management system, and UI platform are Windows Presentation Foundation [8].

Figure 4. *The main steps in building Cyber Demonstration software*

The software development process follows a waterfall model [9], with the main stages described as 0

- Demand survey: Collecting, understanding users' needs;

- Software requirements specification: Defining the functions, environment and operating conditions of the software;

- Scenario design: Based on the theory of network attack techniques such as SQLi, XSS, DOS to build a suitable illustration scenario.

- Software programming: Using C # language, the .Net Framework platforms to design and build software;

- Testing, optimization: Testing and refining, optimizing content based on the opinions of students and teachers.

The software is designed and programmed in the 3-layer model, including GUI Layer, Business Layer, and Data Access Layer, with relationships shown in Figure 5.

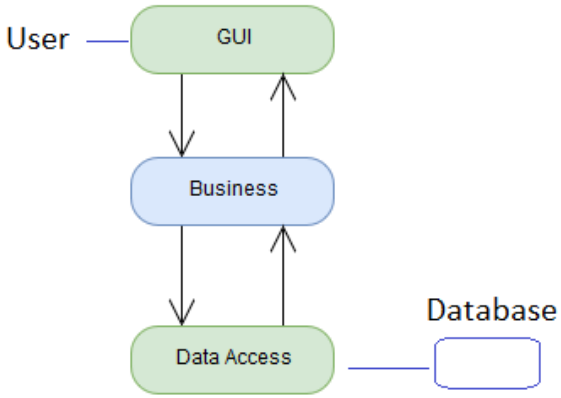


Figure 5. *The 3-layer model in software design*

- GUI Layer: A user interface layer, providing information display functions, providing objects for users to interact with the system.

- Business Layer: This layer receives requests from the GUI layer and accesses the Data Access layer to retrieve information and return to the GUI.

- Data Access Layer: This layer performs the access function to the database, serving queries sent from the Business layer.

2.2.2. The function of the software

Cyber Attack software is installed and operated on Windows operating system environments such as Windows 7, Windows 8.1, and Windows 10. With scenarios simulating standard network attack techniques today, including SQL Attack Injection, Cross-Site Scripting, Denial of Service, Keylogger, Password Detection, Distributed Service Denial, and Denial of Reflection Service. Details of the illustrative techniques are listed in Figure 6.

Each illustrated scenario includes the following components:

(1): Subjects: Including subjects representing objects in an attack such as Hackers (Attacker), Users (Users), Victims (Victim, Server), Control Server (C&C) Server), packets (Packet).

(2) Processes: Include behaviors, interactions between objects to illustrate each stage in a cyberattack.

(3) Descriptions and supporting tools for practice.

Cyber Demonstration	SQL Injection	Bypass
		SELECT
		INSERT
		Procedure Injection
	Cross-Site Scripting	Reflected XSS
		Stored XSS
		DOM-based XSS
	DOS	Tear Drop
		Smurf Attack
		Ping of Death
		SYN Flood
		ICMP Flood
	Password	Brute Force
		Dictionary
	Others	DDOS
		DRDOS
		Keylogger

Figure 6. *The techniques are presented in the software*

2.2.3. Software interface

The main interface of the software is divided into three primary columns. The left column is the list of attack techniques; The right column shows the attack steps, and the middle column is the illustrated scenario. The list of attack techniques presented by the software is listed in Section II.2.

An example of simulating the SYN Flood attack technique with users, hackers, victims, and packets is illustrated in Figure 7.

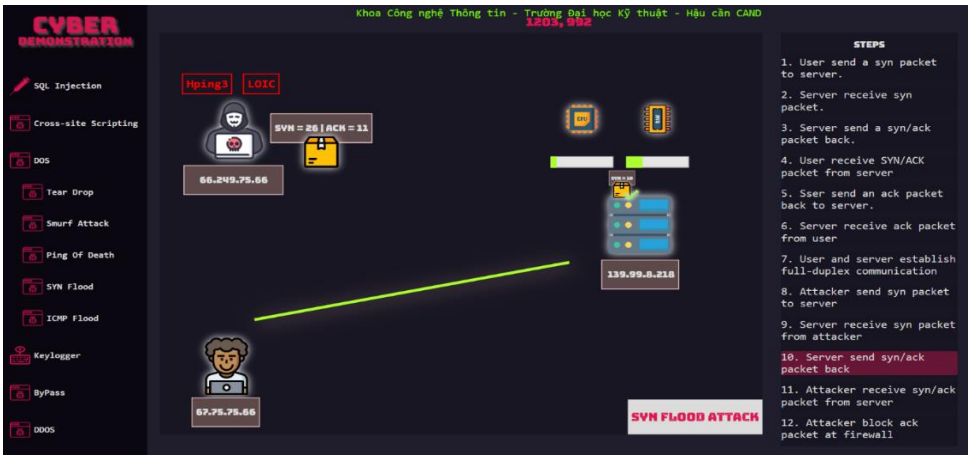


Figure 7. *Illustration interface of the SYN Flood attack technique*

Cyber Demonstration software can be installed simply through the graphical interface, stable operation on two popular operating systems today, Windows 7, and Windows 10. The software is compatible and stably operated with the following utilities: basic presentation tools such as projectors and projection pens.

3. CONCLUSION

In this article, we have summarized some common types of cyber-attacks, such as SQL Injection, Cross Site-Scripting, or Denial of Service. At the same time, we have introduced Cyber Demonstration software to support teachers and students in teaching and to learn about these offensive techniques. The objects, actions, events, and principles in an illustrated, simulated attack help learners gain a new approach when learning about cyber attacking techniques. In the coming time, the author group will continue to develop to upgrade and add new functions and techniques to make the software more useful for users.

REFERENCES

1. "Thủ tướng chính phủ, Quyết định số 99/QĐ-TTg, (2014), Phê duyệt Đề án 'Đào tạo và phát triển nguồn nhân lực an toàn, an ninh thông tin đến năm 2020', ngày 14/01/2014."
2. S. Hansman and R. Hunt, "A taxonomy of network and computer attacks," *Comput. Secur.*, vol. 24, no. 1, pp. 31–43, 2005.
3. J. Viegas and A. Orso (2006), "A classification of SQL-injection attacks and countermeasures," *Int'l Symp. Secur. Softw.*
4. J. Bozic and F. Wotawa (2013), "XSS pattern for attack modeling in testing," *2013 8th Int. Work. Autom. Softw. Test, AST 2013 - Proc.*, pp. 71–74.
5. M. J. H. Heule and O. Kullmann (2019), "The Science of Brute Force," *Best Writ. Math. 2018*, pp. 46–66.
6. M. Wazid et al (2013)., "A framework for detection and prevention of novel keylogger spyware attacks," in *7th International Conference on Intelligent Systems and Control, ISCO 2013*, pp. 433–

438.
7. E. Hugues-Salas *et al (2018).*, "Experimental demonstration of DDoS mitigation over a quantum key distribution (QKD) network using software defined networking (SDN)," *2018 Opt. Fiber Commun. Conf. Expo. OFC 2018 - Proc.*, pp. 1–3.
8. N. A. Roslan, R., & Mohd Zin, (2016) "Dot Net (. NET) Windows form application with Visual Studio.
9. Y. Bassil (2012), "A Simulation Model for the Waterfall Software Development Life Cycle" .

MỘT PHẦN MỀM HỮU ÍCH HỖ TRỢ CHO HỌC TẬP VÀ GIẢNG DẠY VỀ TẤN CÔNG MẠNG

Tóm tắt: *Hiện nay, vấn đề an ninh mạng đang nhận được nhiều quan tâm từ mọi người. Đồng thời, chuyên ngành An toàn thông tin cũng đang được đào tạo ngày càng nhiều ở các trường đại học trong cả nước. Trong bài báo này, chúng tôi trình bày một số kỹ thuật tấn công mạng phổ biến như SQL Injection, Cross Site-Scipting, Denial of Service... và giới thiệu một phần mềm mô phỏng các kỹ thuật đó thông qua các đối tượng, tiến trình và dữ liệu minh họa. Phần mềm hy vọng sẽ là một công cụ hữu ích hỗ trợ giáo viên, học sinh trong quá trình dạy và học.*

Từ khóa: *Công cụ giảng dạy, SQL Injection, Cross Site-Scripting, Denial of Service.*