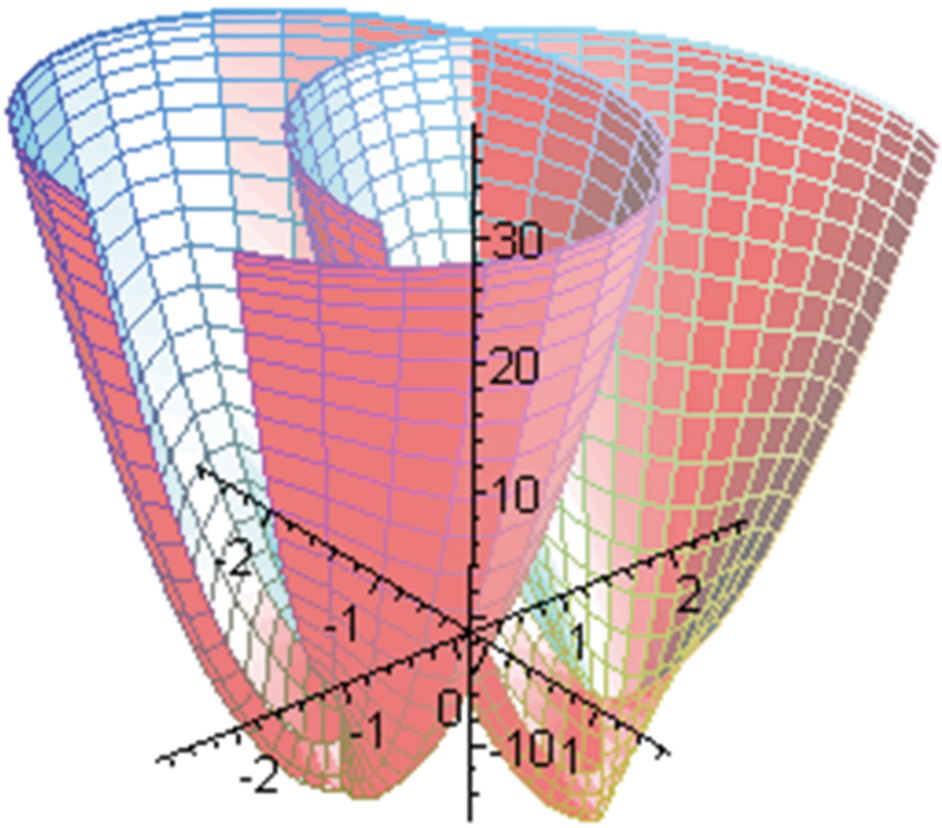
ISSN: 2150-4091 Volume 7, Number 1, January 2015

Scientific Research Publishing

Natural Science



ISSN: 2150-4091

Journal Editorial Board

ISSN: 2150-4091 (Print) ISSN: 2150-4105 (Online)
http://www.scirp.org/journal/ns

Natural Science, 2015, 7, 1-63
Published Online January 2015 in SciRes. http://www.scirp.org/journal/ns

Table of Contents

Volume 7 Number 1 January 2015

The figure on the front cover is from the article published in Natural Science, 2015, Vol. 7, No. 1, pp. 1-9 by S. M. Yakubenya.

Natural Science (NS)

Journal Information

SUBSCRIPTIONS

The *Natural Science* (Online at Scientific Research Publishing, www.SciRP.org) is published monthly by Scientific Research Publishing, Inc., USA.

Subscription rates:

Print: $89 per copy.
To subscribe, please contact Journals Subscriptions Department, E-mail: sub@scirp.org

SERVICES

Advertisements
Advertisement Sales Department, E-mail: service@scirp.org

Reprints (minimum quantity 100 copies)
Reprints Co-ordinator,Scientific Research Publishing, Inc., USA.
E-mail:sub@scirp.org

COPYRIGHT

PRODUCTION INFORMATION

For manuscripts that have been accepted for publication, please contact:
E-mail: ns@scirp.org

Natural Science, 2015, 7, 1-9
Published Online January 2015 in SciRes. http://www.scirp.org/journal/ns
http://dx.doi.org/10.4236/ns.2015.71001

A Manganese Ions Ground State in Mn_xSi_{1-x}: Negative-*U* Properties Centre?

S. M. Yakubenya

RRC "Kurchatov Institute", Moscow, Russia
Email: benyajamp@yahoo.com

Received 3 November 2014; revised 8 December 2014; accepted 26 December 2014

Abstract

In the paper, the properties of magnetic diluted and strong correlated systems of Mn_xSi_{1-x} systems are discussed. The double defects including manganese ion and silicon vacancy are the frame work of the our model introduced for the description of these systems properties. The role of the Jahn-Teller distortions of different symmetry types in MnSi system magnetic-properties formation is discussed. It has been established that the manganese related defect is the center with negative-*U* properties and Jahn-Teller's full symmetric vibration mode initiates change of a crystal-field value from intermediate to strong.

Keywords

Manganese Silicide, Silicon, Impurity, Strong Correlated System, Jahn-Teller Effect

1. Introduction

Intermetallic compounds based on the elements with incompletely filled 3d- or 4f-shell have attracted attention of researchers on an extent of more than 40 years [1]-[3]. See review [2] for details. Such interest was caused firstly by all their unusual magnetic and thermodynamic properties [4] [5], which directly connected with electronic structure of a magnetic ion formed corresponding to sublattice of intermetallic compound. In a case of strongly diluted systems, when magnetic ions are accompanying impurity, interest to their properties is connected to their key role in formation of practically important parameters, in particular free-carrier lifetime in the system [6] [7].

The problem of a manganese ions ground state in Mn_xSi_{1-x}, values of manganese ions magnetic moment and gyromagnetic ratio for it, is discussed in the present work. Two limiting cases are considered: 1) a concentration manganese ion is infinitesimal as a free carriers concentration $(x \to 0)$. Interaction between magnetic ions and free carrier is absent; 2) concentration of Mn is close to that of Si $(x \sim 0.5)$. The interaction of magnetic moments manganese ions with quasiparticals is a very important component in magnetic-property-compound for-

mation. In the first case, Mn ions can be treated as a doping impurity of Si crystal.

Both systems are crystallized in the cubic lattice type. Symmetry group T_d is realized for the magnetic deluted system [8] and B20 takes place in the case of strongly correlated system correspondingly [9]. Taken into account that interaction between atomic orbitals and crystal field is dominated with respect to other interactions in crystal, we shall look for the electronic configuration change of the manganese ions ground state as the increase of magnetic component-concentration in compound.

Magnetic moment value of the manganese ions in the systems in a wide temperature interval will be the main subject of our investigation. The model of the double defect, formulated early by us [10] [11], will be used for the description of the both systems magnetic properties. Short definition of the double defects model will be done below.

2. Model

Two different defect types in semiconductors irrelevant with each to other are discussed as a usual [8]. It is named by interstitial defects $\{X_i\}$ (impurity ions located in a crystal lattice pore) and substitution defects $\{X_s\}$ (impurity ions located in point of lattice).

In the same times, in the frame of the double defect model, the interstitial and substitution defects are considered as a two limiting cases of double defect. A silicon vacancy $\{V_{si}\}$ acts as a partner for the Mn impurity ions $\{\mathrm{Mn}_s\}$ in the suggested double defect.

At that when,

- distance between partners is infinity, then interstitial defect type is realized;
- distance between partners is an infinitesimal value, then we have substitution type of defect.

The account of silicon vacancy at interstitial defect case is necessary to consider all type of double defects from the common base. This will result to unique parameter—Fermi level position (E_F), which will define a degree of one-electronic orbital fillings of such kind of double defect. (E_F) is counted from vacuum level position in the case.

The intermediate case, when distance between partners is of the order of the lattice period, the only one valence bond saturated exists between partners. Such type of defects has been named by the pairing defects. A plenty of pairing defect different types may exist, depending on overlapping integral of wave functions partners and number of nonequivalent crystallographic positions in a lattice. Only single type of pairing defects is realized in silicon. Moreover, we can see electronic density redistribution between the central ion and ions of the nearest crystal environment as a result of temperature variation or other factors, in particular, a electronic configuration change of everyone components as result of Jahn-Teller distortions in the frame of the model. It is necessary to note, that electronic density redistribution can be realized and without a free carrier generation. The magnetic-moment average value located at manganese ion can be renormalized very strongly as free carriers appear in the system.

The considering of double defect model we intentionally limit to the consideration of single elementary cell, understanding thus, that significant electronic density delocalization take place due to hybridization effects and indemnification of a charge occurs much further, than the first coordination sphere. Here and further for a designation of such defects we shall use the notation for substitution $\{\mathrm{Mn}_s^m - V_{si}^n\}$ and $\{\mathrm{Mn}_i^m - V_{si}^n\}$, for pairing defects and interstitial defects $\{\mathrm{Mn}_i^m\} - \{V_{si}^n\}$. m and n indexes designate a charge of everyone components with respect to its nucleus. Take into account an electro neutrality condition of a crystal as a whole, we can write equation for the neutral charge state of double defect:

$$m + n \equiv 0 \tag{1}$$

where m and n can take different value, not only zero as it practices at the standard approach of the description of substitution defects behavior in semiconductors [8]. In other words, numbers of electrons, located on valence bonds between central ion and the nearest neighbors, may be differ from one to one for the impurity ion cases and ions of the basic lattice. Four electrons occupy silicon vacancy orbitals for the V_{si}^o case.

Now all are ready for proceed to discussion of a manganese ion ground state in intermetallic compound $\mathrm{Mn}_x\mathrm{Si}_{1-x}$ and in strongly diluted system silicon-manganese. We shall begin a discussion from the case of strongly diluted system: silicon-manganese.

3. Result and Discussion

3.1. Mn_xSi_{1-x} Strongly Diluted System ($x \to 0$)

The crystal lattice in such compound will mainly consist of silicon atoms and is described by point group of symmetry T_d [8]. Silicon demonstrates semiconductor properties and gap is in a one-partial excitation spectrum. The forbidden band value E_g is equal to 1.17 eV at temperatures close to absolute zero [12].

Mono- and polycrystalline silicon used as a material for solar batteries, is presented up to 15 metal impurities at concentration above than 10^{12} ions/sm^3 and higher [13] [14]. Their number includes also manganese. Transition metal impurities located in interstitial position in crystal lattice of silicon, as a role [8] [13] [14]. Sample preparation, in which main part impurity manganese ions located in unit of crystal lattice, is not ordinary task and we know few works only [15]-[17], in which authors have solved the given problem. Concentration of the point defects due to manganese impurity ions located in unit of crystal lattice, does not exceed 10^{15} ions/cm^3 in the crystal and any magnetic interactions as a manganese—manganese as a free carrier—magnetic ion can be neglected.

But, crystal field effects and hybridization between impurity and band states are played a dominant role in the formation of the magnetic-ions electronic structure. Manganese atoms have half-filled 3d-shell and two electrons occupy an outside 4s-shell. An electronic configuration of the Mn^0 looks like $3d^54s^2$. Impurity ion 3d-orbitals are split by tetrahedral-symmetry crystal field on the t_2- and e_2-orbitals (three- and double-degenerated states). The electronic configuration looks like $\left(e_{2\uparrow}\right)^2\left(t_{2\uparrow}\right)^3$ (↑—spin projection up; ↓—spin projection down). The order of filling such orbitals is defined by a lot of parameters, including of a crystal field value (strong, intermediate, weak). The Fermi level position in the gap is defined by charging state of defects and their mutual concentration.

In early works directed to research the behavior of manganese ions in silicon [18] [19], key parameters of electronic structure of the defects connected to them have been determined, but interpretation of the received results was carried out within the framework of standard representations about substitution defects in semiconductors. Repeatedly, a question on electronic structure of substitution defects connected with manganese ions have returned only 30 years later, after from carrying out of the first experiments [15] [16]. Let's note, that researches of EPR (Electron Paramagnetic Resonance) spectra of Mn_xSi_{1-x} $(x \to 0)$ samples and simultaneously with it of spectra DLTS (Deep Level Transition Spectroscopy) [15] [16] have allowed to remove question on the magnetic moment value located on a magnetic ion (**Figure 1**, **Figure 2**).

Fermi level (E_F) is located near the valence band top in the investigated crystals at low temperatures. It is caused by presence of shallow acceptors at concentration exceeding the general concentration of manganese ions [8]. As the result, the dominate part of manganese ions, should be in charge state $\left\{Mn_s^{+1} - V_{si}^{o}\right\}$ and the magnetic moment of manganese ions has pure spin origin. It is defended by two electrons, which occupied e_2-orbitals (configuration $\left(e_{2\uparrow}\right)^2$). Moreover, it is expected, what other electrons of manganese and silicon participate in valence bond formation, as it is take place in perfect silicon crystal. There is very critical contradiction in last postulation.

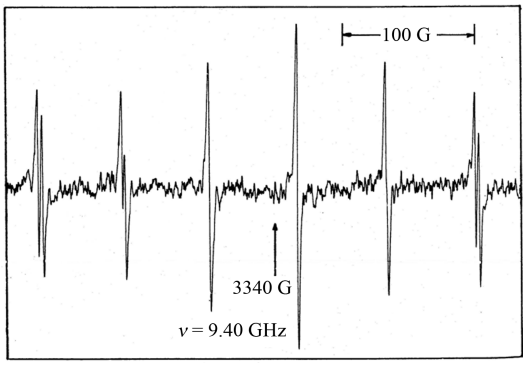


Figure 1. EPR first-derivative absorption signal from Mn_s^+ in arbitrary orientation [15]. (*g*-factor is pure spin value and it may be associated with manganese ion configuration $\left(e_{2\uparrow}\right)^2$ or $\left(e_{2\uparrow}\right)^2\left(t_{2\uparrow}\right)^3\left(t_{2\downarrow}\right)^3$. See text below for the details.)

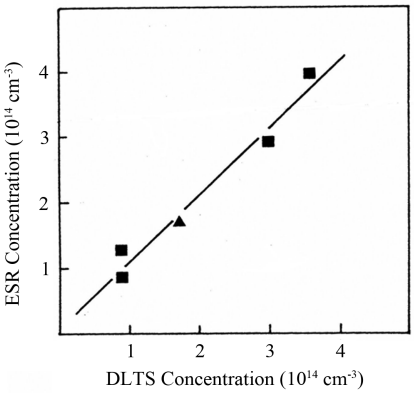


Figure 2. Correlation plot of 0.38 eV DLTS peak intensity $Mn_s^{+/0}$ almost and integrated ESR intensity of the Mn_s^+ signal [15].

As very well know, that V_{si}^{+2} is the greatest possible charging state silicon vacancy [20] [21]. In other words, two electrons occupy of 3s-orbitals of silicon at any Fermi level position in the gap. Energy ionization 3s-shell of silicon is ~35 eV [22]. In same times, defect $\left\{Mn_s^{+1}-V_{si}^{o}\right\}$, having four saturated valence bonds between surrounding and central ions, It can be realized if Mn^{4+} charge state take place. It is equivalent to postulate, that electrons with bonding energy ~100 eV, participate in valence bond formation. But it is not realistic suggestion.

In the same time EPR spectra of $\left\{Mn_s^{+1}-V_{si}^{o}\right\}$ defects is described by means of cubic spin-Hamiltonian very well [15] [23]. At the standard approach, when one considers only electrons located on a magnetic ion, it is not look possible to find the reasonable decision of the arising contradiction.

The similar situation took place in system GaAs:Mn and was a subject of discussions more than 30 years [24]-[26]. The decision of the problem within the framework of double defect model has been found some years ago [27].

Let's consider, that in valence bond formation between magnetic ion and surrounding only electrons with bonding energy in free atom less than 35 eV on absolute value are involved. Energy scale is counted from a vacuum level position. In other words, two from four valence bond is saturated in $\left\{Mn_s^{+1}-V_{si}^{o}\right\}$ defects. Two valence bond is broken. The defect in such configuration is unstable with respect to Jahn-Teller distortion. The most probable direction of distortions is directions of {110} type (tetragonal distortions of local symmetry—*E*-mode of vibration [28] [29]). Depending on a relation between phonon local energy $\left(\hbar\omega_{\text{loc}}\right)$ and depth of a potential pit $\left(E_{J\text{-}T}\right)$ Jahn-Teller effect can be static or dynamic (**Figure 3(a)**, **Figure 3(b)**).

$E_{J\text{-}T}$ and Δ-barrier value between potential energy minimums are defined as:

$$E_{J\text{-}T}=\left(B_1\right)^2/\left(2k_e-4B_2\right) \tag{2a}$$

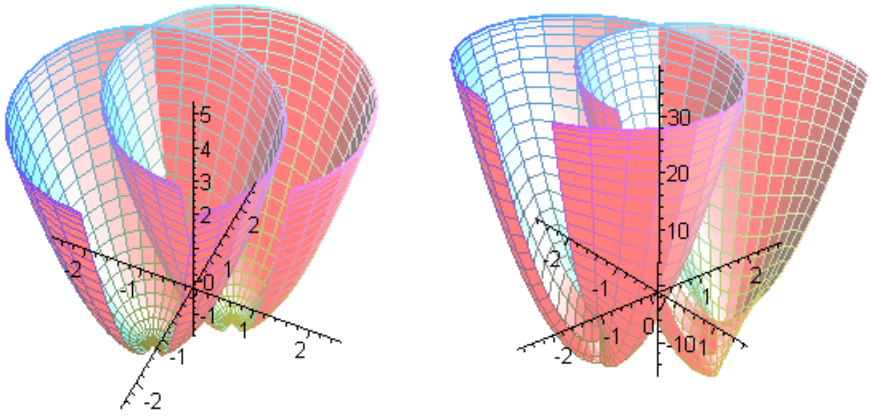
$$\Delta=4B_2*\frac{E_{J-T}}{\left(k_e-2B_2\right)} \tag{2b}$$

where B_1, B_2, k_e constant (for more detail see [28]).

A dynamic type of Jahn-Teller effect takes place for the manganese ions in such system (see EPR spectra). But, we shall have full moment *J*, located on manganese ion incoming in $\left\{Mn_s^{+1}-V_{si}^{o}\right\}$ defect, equal two at presence tetragonal Jahn-Teller distortions only. An electronic configuration of the magnetic electrons is $\left(\left(e_{2\uparrow}\right)^2\left(t_{2\uparrow}\right)^3\left(t_{2\downarrow}\right)^1\right)$. Here, t_2-orbitals is a gibrids of 3p-orbitals of silicon and 3d-orbitals of manganese.

Provided that simultaneously with tetragonal distortion it takes place also full symmetric Jahn-Teller distortions of local symmetry (A_1—mode of vibrations). The full magnetic moment (μ_{total}) located on an ion of manganese is equal $\left(\sqrt{2}\right)\mu_B$ (μ_B-Born magneton). Electronic configuration is $\left(e_{2\uparrow}\right)^2\left(t_{2\uparrow}\right)^3\left(t_{2\downarrow}\right)^3$. $\left(\mu_{\text{total}}\right)$ is connected with manganese spin $\left(S=1\right)$ by the following ratio:

$$\mu_{\text{total}}=\left(\sqrt{S*\left(S+1\right)}\right)\mu_B. \tag{3}$$



(a) (b)

Figure 3. (a) Potential energy surface for the dynamic Jahn-Teller effect case; (b) Potential energy for the static Jahn-Teller effect case (in generalized coordinates space).

As we see, very good consent between calculated and experimentally values of moment *J* takes place. In other words, the defect $\left\{\mathrm{Mn}_s^m - V_{si}^n\right\}$ can exist in two isoelectronic configurations:

$$m = +1 \quad n = 0 \tag{4a}$$

$$m = -1 \quad n = +2 \tag{4b}$$

but, transition between configurations take place without change of density of free carriers of a charge.

The defects with an electronic configuration (4b) are responsible for observable EPR spectra. In the same time, DLTS spectra are connected with (4a) configuration of the double defect.

In this case it is easy to coordinate value of activation energy $\varepsilon_i^{\exp}$ obtained in experiment (DLTS), and $\varepsilon_i^{\mathrm{cal}}$ (calculated). Observable value will be well coordinated with the value, received by means of "spin marking method" [30]. Experimental value $\varepsilon_i^{\exp}$ is equal to 0.38 eV [15] and $\varepsilon_i^{\mathrm{cal}}$:

$$\varepsilon_i^{\mathrm{cal}} = 3\varepsilon_0 \tag{5}$$

where ε_0 is energy of one electron transition in silicon. It is equal to ~0.13 eV [20] [21].

The process of electron capture on the deep level is defined by:

$$\left\{\mathrm{Mn}_s^{-1} - V_{si}^{+2}\right\} + 2\varepsilon_0 \Rightarrow \left\{\mathrm{Mn}_s^{+1} - V_{si}^{0}\right\} + \varepsilon_0 \Rightarrow \left\{\mathrm{Mn}_s^{0} - V_{si}^{0}\right\} + h^+ \tag{6}$$

where h^+ hole, the free carrier of a charge.

Within the framework of the standard approach:

$$\varepsilon_i^{\mathrm{cal}} = \varepsilon_0 \tag{7}$$

$$\mathrm{Mn}_s^{+1} + \varepsilon_0 \Rightarrow \mathrm{Mn}_s^{0} + h^+. \tag{8}$$

Look discussion in [10] [11] for more detail.

In the conclusion we shall note, that full symmetric vibration mode initiates change of a crystal field value from intermediate (a configuration $\left\{\mathrm{Mn}_s^{+1} - V_{si}^{0}\right\}$) to strong (a configuration $\left\{\mathrm{Mn}_s^{-1} - V_{si}^{+2}\right\}$). Also $\left\{\mathrm{Mn}_s^{+1} - V_{si}^{0}\right\}$ defect demonstrate properties, characteristic for the centers with negative energy of coulomb correlation (centers with negative-*U* properties). In this connection, most likely, revision of experimental data interpretation on behavior Fe and Ni ions in silicon is necessary.

3.2. Mn_xSi_{1-x} Strongly Correlated System ($x \sim 0.5$)

Crystalline system MnSi, in which both components are approximately having equal concentrations, traditionally attributes to strong correlated systems [1]. As opposed to strongly diluted systems, considered above, manganese ions build up it's own sublattice and any shallow acceptors is absent in the system. Take into account, that electrical resistivity of the manganese silicite samples drop monotonically down to 20 mK [31], we can postulate

that conduction band states (local level in the gap in magnetic-diluted-system case) is not empty in the system. It is formed by mean of hybridization 3p-orbitals of silicon and 4s-3d-orbitals manganese [32].

At the same time, small distinction in types of a crystal lattice between magnetic diluted system (group of symmetry T_d—four equivalent atoms of silicon in the first coordination sphere—2.35 Å [8]) and crystals of manganese silicide (type of lattice B20—three equivalent atoms of silicon on distance 2.396 Å and one—2.313 Å [9] at [111] direction) takes place.

Displacement of atoms of the nearest environment in the opposite direction in lattice B20 in comparison with atoms position in tetrahedral lattice is powerful argument for the existence of the cooperative Jahn-Teller effect in such system. But, distances between the nearest ions of manganese in $Mn_{0.5}Si_{0.5}$, which do not form valence bonds with each other, practically coincides with such parameter for metallic γ-manganese [33].

Systematic measurements of the magnetic moment, located on the manganese ions, show presence of several temperature intervals in which magnetic ions demonstrate completely different properties [33] [34]. It is paramagnetic at temperature above 30 K. First-order transition takes place at 29.5 K [5] [35]. Two different types of magnetic ordering are realized at temperature below 29.5 K. Helical magnetic structure with a long-range order was observed at $H_{\text{ext}} \to 0$ (H_{ext}—external magnetic field) and the structure becomes progressively more conical at increase of H_{ext} value. Moreover, there is critical value of $H_{\text{ext}}^{0}(T)$ when ferromagnetic order type is realized in the system (**Figure 4**).

Within the framework of double defect model the given transformations can be treated in the following way. The value of the magnetic moment, located on the manganese ions, in a paramagnetic phase ($\mu_{\text{para}} \sim 1.4\mu_B$ [34] [36] [37], g-factor is not equal 2.0 but 2.81) practically coincides with those one for a case of magnetic diluted system. That is why we can postulate, that interaction between free carriers of charge and the magnetic moment, located on ions of manganese, can be neglected. Manganese ions precess around the [111] axes of the crystal. Cooperative Jahn-Teller effect takes place (see text above).

First-order transition take place at 29.5 K and transition $\left\{Mn_s^{+1} - V_{si}^{0}\right\} \Rightarrow \left\{Mn_s^{+1-d} - V_{si}^{0+d}\right\}$ is realized. Here parameter $d > 0$ and it depicts a redistribution electron density between partners in the double defect (see discussion for the magnetic diluted system). Let's note, that the amplitude of full symmetric vibrations in this case may be less then in magnetic diluted system. The system is closed in one of Jahn-Teller's minimum (**Figure 3(b)**). Moreover, we have got fist-order transition.

It is necessary to note, that very intricate case of Jahn-Teller distortions is realized here—trigonal + tetragonal + full symmetrical types of distortions take place simultaneously and we have got a problem with large quantity of independent parameters. As result, numbers of Jahn-Teller's minima can be equal to 24 [28] [29]. In the case,

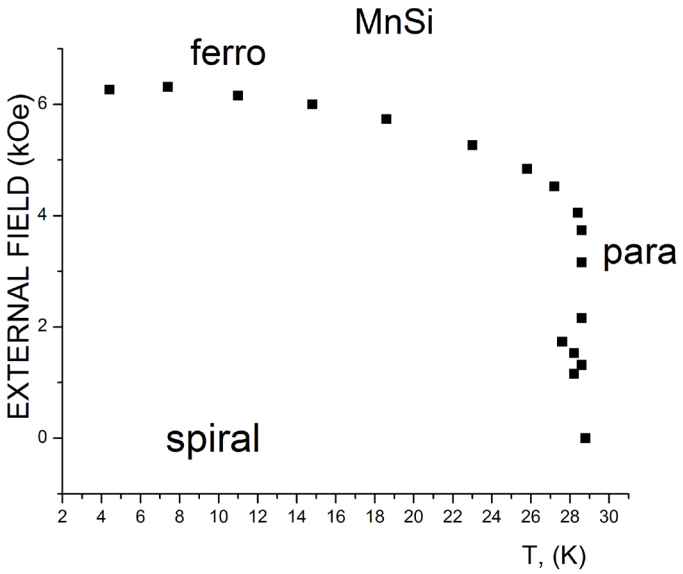


Figure 4. Magnetic phase diagram of MnSi. Inset show the projection of Mn atomic positions on a <100> plane. This atomic arrangement lacks a center of symmetry [34].

the spiral structure is connected with a jump between minima of potential energy. As result of applied external magnetic field, the transformation of a local minimum to absolute one is a case of the transition from the spiral structure to ferromagnetic order type in $Mn_{0.5}Si_{0.5}$. Parameter $\mu_{fer}H^0_{ext}(T\to 0)$ can be defined, approximately, as a potential barrier Δ [29] between different Jahn-Teller's minima:

$$\Delta = \mu_{fer}H^0_{ext}(T\to 0). \tag{9}$$

$\mu_{fer}\cong 0.4\mu_B$ and $H^0_{ext}(T\to 0)\cong 6.5$ kOe at temperatures below T_N [34]. The discussion about μ_{fer} value in frame of fluctuated Fermi-liquid model may be found in work.

Separately it is necessary to focus on the dependence of the phase-transition temperature from hydrostatic pressure value.

As we already discussed above, the first order phase transition, occurring at temperature 29.5 K at normal pressure, is connected with redistribution of electronic density between manganese and surrounding silicon atom. But 3p-orbitals of silicon is more delocalizated in comparison with 3d-orbitals of manganese [22]. As result:

$$\left(\partial E_{3p}/\partial P\right)_{Si} > \left(\partial E_{3d}/\partial P\right)_{Mn} \tag{10}$$

where P -hydrostatic pressure value.

But if ratio (10) is corrected, the phase transition temperature has to drop at the hydrostatic-pressure increase. Moreover, there is critical value $P_{critical}$ when transition is impossible. $\left\{Mn_s^{-1}-V_{si}^{+2}\right\}$ defect configuration is unfavorable and phase transition is not observed. It was confirmed experimentally [5] [31].

4. Conclusions

Thus, the model of double defect, stated above, can be successfully applied for the description of the magnetic properties of Mn_xSi_{1-x} system with different concentrations of manganese ions.

We found that the angular momentum (J) of a manganese ion is integer, but not fractional as assumed early.

It is assumed that local level, located in the forbidden band of silicon and connected with recharging of impurity manganese ions, transforms in the band with increase of manganese concentration up to $x\sim 0.5$ in the Mn_xSi_{1-x} system. We found correct descriptions for the level position in the gap of silicon at low manganese concentration in frame of double defect model. The transition from intermediate crystal field to strong one as the result of full symmetry of Jahn-Teller vibration mode has been detected.

It has been established that $\left\{Mn_s^{+1}-V_{si}^{0}\right\}$ defect is the center with negative-U properties.

The favorable difference of our model from the approach of weakly-fluctuated Fermi-liquid one, which had been used widely for the description of physical properties of manganese silicide in last years, is the opportunity of definition of original causes of those or other effects.

It is obvious that a number of experimental and theoretical studies perform with the aim of specification of our model parameters. First of all, it concerns measurement of inelastic neutrons scattering spectra on the MnSi single-crystalline sample and on amorphous a-MnSi, in which helical magnetic order is absent. It is fruitful to compare EPR spectra of Mn_xSi_{1-x} $(x\to 0)$ and a-MnSi in a wide temperature interval.

Useful addition to these experiments would be a comparison of the μ^+-mesons scattering experiment results and Knight's shift in spectra of nuclear magnetic resonance MnSi at temperatures $\sim T_N$.

Acknowledgements

The author is indebted to P. A. Alekseev and V. N. Luzukov for helpful discussion and critical reading of the manuscript.

References

[1] Izumov, U.A. and Kurmaev, E.Z. (2008) Materials with Strong Electron Correlations. *Uspekhi Fizicheskikh Nauk*, **178**, 25-60. http://dx.doi.org/10.3367/UFNr.0178.200801b.0025

[2] Stishov, S.M. and Petrova, A.E. (2011) Helicoidal Band Magnatic MnSi. *Uspekhi Fizicheskikh Nauk*, **181**, 1157-1170. http://dx.doi.org/10.3367/UFNr.0181.201111b.1157

[3] Karado, K., Matsuzaki, I., Yamazuki, I., Kreitzman, S.R. and Brewer, J.H. (1990) Spin Dynamics of the Itinerant Helimagnet MnSi Studied by Positive Muon Spin Relaxation. *Physical Review B*, **42**, 6515-6522.

http://dx.doi.org/10.1103/PhysRevB.42.6515

[4] Grigoriev, S.V., Maleyev, S.V., Okorokov, A.I., Chetverikov, Yu.O. and Eckerlebe, H. (2006) Field-Induced Reorientation of the Spin Helix in MnSi near *Tc*. *Physical Review B*, **73**, Article ID: 224440. http://dx.doi.org/10.1103/PhysRevB.73.224440

[5] Petrova, A.E., Bauer, E.D., Krasnorusski, V. and Stishov, S.M. (2006) Behavior of the Electrical Resistivity of MnSi at the Ferromagnetic Phase Transition. *Physical Review B*, **74**, Article ID: 092401. http://dx.doi.org/10.1103/PhysRevB.74.092401

[6] Macdonald, D., Rosenits, P. and Deenapanray, P.N.K. (2007) Recombination Activity of Manganese in p- and n-Type Crystalline Silicon. *Semiconductor Science and Technology*, **22**, 163-168. http://dx.doi.org/10.1088/0268-1242/22/2/028

[7] Roth, T., Rosenits, P., Deiz, S., Glunz, S.W., Macdonald, D., Beljakowa, S. and Pensl, G. (2007) Electronic Properties and Dopant Pairing Behavior of Manganese in Boron-Doped Silicon. *Journal of Applied Physics*, **102**, Article ID: 103716. http://dx.doi.org/10.1063/1.2812698

[8] Harrison, W.A. (1980) Electronic Structure and Properties. W.H. Freeman and Company, San Francisco.

[9] Hayes, T.M., Allen, J.W., Boyce, J.B. and Hauser, J.J. (1981) Interplay of Structure and Magnetic Properties in MnSi, a Concentrated Spin-Glass. *Physical Review B*, **23**, 4691-4699. http://dx.doi.org/10.1103/PhysRevB.23.4691

[10] Yakubenya, S.M. (1991) Manganese Ion Properties in Silicon: Interstitial Defects. *Physica Tverdogo Tela*, **33**, 1462-1469.

[11] Yakubenya, S.M. (1991) Manganese Ion Properties in Silicon: Substitution and Pairing Defects. *Physica Tverdogo Tela*, **33**, 1470-1477.

[12] Grigo'ev, I.S. and Meilikhov, E.Z. (1991) Handbook of Physical Values. Energoatomizdat, Moscow.

[13] Buonassisi, T., Heuer, M., Istratov, A.A., Pickett, M.D., Marcus, M.A., Lai, B., Cai, Z., Heald, S.M. and Weber, E.R. (2007) Transition Metal Co-Precipitation Mechanisms in Silicon. *Acta Materialia*, **55**, 6119-6126. http://dx.doi.org/10.1016/j.actamat.2007.07.030

[14] Istratov, A.A., Hieslmair, H. and Weber, E.R. (1999) Iron and Its Complexes in Silicon. *Applied Physics A*, **69**, 13-44. http://dx.doi.org/10.1007/s003390050968

[15] Czaputa, R., Feichtinger, H., Oswald, J., Sitter, H. and Haider, M. (1985) Energy Level of the 0 to + Charge Transition of Substitutional Manganese in Silicon. *Physical Review Letters*, **55**, 758-761. http://dx.doi.org/10.1103/PhysRevLett.55.758

[16] Haider, M., Sitter, H., Czaputa, R., Feichtinger, H. and Oswald, J. (1987) Experimental Identification of the Energy Level of Substitutional Manganese in Silicon. *Journal of Applied Physics*, **62**, 3785-3789. http://dx.doi.org/10.1063/1.339217

[17] Graff, K. (1995) Metal Impurities in Silicon-Device Fabrication. Springer, Berlin. http://dx.doi.org/10.1007/978-3-642-97593-6

[18] Carlson, R.O. (1956) Properties of Silicon Doped with Manganese. *Physical Review*, **104**, 937-940. http://dx.doi.org/10.1103/PhysRev.104.937

[19] Ludwig, G.W. and Woodbury, H.H. (1962) Electron Paramagnetic Resonance in Solid State Physics, Vol. 13, Academic Press, New York.

[20] Baraff, G.A., Kane, E.O. and Schlüter, M. (1980) Theory of the Silicon Vacancy: An Anderson Negative-U System. *Physical Review B*, **21**, 5662-5683. http://dx.doi.org/10.1103/PhysRevB.21.5662

[21] Newton, J.L., Chatterjee, A.P., Harris, R.D. and Watkins, G.D. (1983) Negative-U Properties of the Lattice Vacancy in Silicon. *Physica B + C*, **116**, 219-223. http://dx.doi.org/10.1016/0378-4363(83)90250-4

[22] Radtsig, A.A. and Smirnov, B.M. (1986) Handbook of Atom Characteristic and Atomic Ions. Energoatomizdat, Moscow.

[23] Abragam, A. and Bleaney, B. (1970) Electron Paramagnetic Resonance of Transition Ions. Clarendon Press, Oxford.

[24] Andrianov, D.G., Lazareva, G.V., Savel'ev, A.S. and Yakubenya, S.M. (1983) Magnetic Properties of GaAs:Mn. *Phizica I Tekhika Poluprovodnicov*, **17**, 813-817.

[25] Il'in, N.P., Masterov, V.F. and Vasil'ev, A.E. (1992) Impurity Centers with Incompletely Filling 3d-Shell in Binary Compound. *Phizica I Tekhika Poluprovodnicov*, **26**, 1866-1877.

[26] Schneider, J., Kaufmann, U., Wilkening, W., Baeumler, M. and Köhl, F. (1987) Electronic Structure of the Neutral Manganese Acceptor in Gallium Arsenide. *Physical Review Letters*, **59**, 240-244. http://dx.doi.org/10.1103/PhysRevLett.59.240

[27] Yakubenya, S.M. (1997) Paramagnetic Resonance in Manganese Impurity Centers in GaAs:Mn. *Journal of the Mos-*

cow Physical Society, **7**, 273-286.

[28] Bersuker, I.B. (1983) The Jahn-Teller Effect. Plenum, New York.

[29] Bersuker, I.B. and Polinger, V.Z. (1982) Vibronic Interactions and the Jahn-Teller Effect. In: Lowdin, P.-O., Ed., *Advances in Quantum Chemistry*, Vol. 15, Academic Press, New York, 85-160.

[30] Alexsandrov, P.A. and Yakubenya, S.M. (1990) Method of Spin Marking in Semiconductors. In: Sumino, K., Ed., *Proceeding of International Conference IC—STDCS*, Elsevier Science Publishers B.V., Amsterdam, 1605-1608.

[31] Pfleiderer, C., McMullan, G.J., Julian, S.R. and Lonzarich, G.G. (1997) Magnetic Quantum Phase Transition in MnSi under Hydrostatic Pressure. *Physical Review B*, **55**, 8330. http://dx.doi.org/10.1103/PhysRevB.55.8330

[32] Oh, S.J., Allen, J.W. and Lawrence, J.M. (1987) Electron Spectroscopy Study of FeSi and CoSi. *Physical Review B*, **35**, 2267-2276. http://dx.doi.org/10.1103/PhysRevB.35.2267

[33] Ishikawa, Y., Shirane, G., Tarvin, J.A. and Kohgi, M. (1977) Magnetic Excitations in the Weak Itinerant Ferromagnet MnSi. *Physical Review B*, **16**, 4956-4961. http://dx.doi.org/10.1103/PhysRevB.16.4956

[34] Shirane, G., Cowley, R., Majkrzak, C., Sokoloff, J.B., Pagonis, B., Perry, C.H. and Ishikawa, Y. (1983) Spiral Magnetic Correlation in Cubic MnSi. *Physical Review B*, **28**, 6251-6256. http://dx.doi.org/10.1103/PhysRevB.28.6251

[35] Stishov, S.M., Petrova, A.E., Khasanov, S., Panova, G.Kh., Shikov, A.A., Lashley, J.C., Wu, D. and Lograsso, T.A. (2007) Magnetic Phase Transition in the Itinerant Helimagnet MnSi: Thermodynamic and Transport Properties. *Physical Review B*, **76**, 52405-52412. http://dx.doi.org/10.1103/PhysRevB.76.052405

[36] Ziebeck, K.R.A. and Brown, P.J. (1980) Measurement of the Paramagnetic Response Function in the Weak Itinerant Magnetic Compound MnSi Using Polarised Neutron Scattering. *Journal of Physics F*: *Metal Physics*, **10**, 2015-2019. http://dx.doi.org/10.1088/0305-4608/10/9/017

[37] Ziebeck, K.R.A., Brown, P.J., Booth, J.G. and Bland, J.A.C. (1981) Observation of Spatial Magnetic Correlations in the Paramagnetic Phase of Weak Ferromagnets; Short-Range Magnetic Order in MnSi at 20 T_N. *Journal of Physics F*: *Metal Physics*, **11**, L127-L131. http://dx.doi.org/10.1088/0305-4608/11/6/002

Natural Science, 2015, 7, 10-17
Published Online January 2015 in SciRes. http://www.scirp.org/journal/ns
http://dx.doi.org/10.4236/ns.2015.71002

Fabrication of Thorium and Thorium Dioxide

Balakrishna Palanki

(Retired) Nuclear Fuel Complex, Hyderabad, India
Email: palankibalakrishna@yahoo.com

Received 10 November 2014; revised 9 December 2014; accepted 28 December 2014

Abstract

Thorium based nuclear fuel is of immense interest to India by virtue of the abundance of Thorium and relative shortage of Uranium. Thorium metal tubes were being cold drawn using copper as cladding to prevent die seizure. After cold drawing, the copper was removed by dissolution in nitric acid. Thorium does not dissolve being passivated by nitric acid. Initially the copper cladding was carried out by inserting copper tubes inside and outside the thorium metal tube. In an innovative development, the mechanical cladding with copper was replaced by electroplated copper with a remarkable improvement in thorium tube acceptance rates. Oxalate derived thoria powder was found to require lower compaction pressures compared to ammonium diuranate derived urania powders to attain the same green compact density. However, the green pellets of thoria were fragile and chipped during handling. The strength improved after introducing a ball milling step before compaction and maintaining the green density above the specified value. Alternatively, binders were used later for greater handling strength. Magnesia was conventionally being used as dopant to enhance the sinterability of thoria. The normal sintering temperature for magnesia doped thoria was 1600˚C - 1700˚C, which was achieved in electrically heated molybdenum element sintering furnaces with reducing atmosphere. 0.25 mole percent addition of niobia to the thoria was found to bring down the sintering temperature to 1150˚C. Sintering could be done in ordinary furnaces in air atmosphere using silicon carbide or Kanthal heating elements. Electrical conductivity was measured for both magnesia and niobia doped sintered thoria and used in interpreting differences in sintering behavior.

Keywords

Thorium, Electroplating, Thorium Dioxide, Pelletizing, Sintering, Coring, Electrical Conductivity

How to cite this paper: Palanki, B. (2015) Fabrication of Thorium and Thorium Dioxide. *Natural Science*, **7**, 10-17.
http://dx.doi.org/10.4236/ns.2015.71002

1. Introduction

The work reported in this paper was carried out by this author intermittently during the period 1973 to 2007. Most of it was documented as internal reports and some of it was published [1]-[10]. Several other valuable publications are also available from literature [11]-[52]. The importance of thorium for the Indian Nuclear Program became clear when it was found that the country was endowed with large mineral resources of thorium. Early development and production work on thorium metal and thorium oxide was carried out at Bhabha Atomic Research Center, Mumbai. Later, thoria pellets for the research reactors CIRUS and DHRUVA, the Fast Breeder Test Reactor (FBTR) as well as the Pressurized Heavy Water Reactor (PHWR) were made at Nuclear Fuel Complex, Hyderabad. The first core of a 220 MWe PHWR requires 385 depleted uranium bundles for flux flattening. It is possible to use thoria bundles instead of depleted uranium oxide bundles. In such a case only 35 thoria bundles are needed [53]. A few export orders were also executed at Hyderabad such as thoria crucibles for emf measurements in Germany and thoria pellets for fusion fission hybrid research facility LOTUS in Switzerland. Some information is presented in the following sections.

2. Thorium Metal Production

Thorium oxide powder received from Indian Rare Earths Ltd. was re-calcined at 900˚C. The calcined powder was mixed with calcium granules and reduced at 900˚C under argon atmosphere. The reacted mass was soaked in water, followed by acid leaching, washing and drying. The powder thus produced was highly pyrophoric and contained impurities carbon, oxygen and nitrogen affecting adversely ductility of the bulk metal. Alternatively, thorium oxalate was chlorinated in the presence of carbon. The crude chloride was purified by vacuum sublimation. The purified chloride was reduced with magnesium to get Th-Mg alloy. The alloy was given a pyro vacuum treatment to separate out the magnesium. The thorium sponge thus obtained was less pyrophoric than the powder. Thorium metal powder was compacted at 300 MPa and sintered in vacuum at 1300˚C for one hour. 98% TD could be obtained. The sintered material could be cold rolled to more than 90% reduction without intermediate annealing [13].

3. Electroplating Copper on Thorium

Sintered thorium metal was copper jacketed and hot extruded into tubes. The tubes were then cold drawn to meet surface and dimensional requirements in the specifications. Thorium can not be easily drawn as it galls or seizes the drawing die and hence thorium metal tubes need to be clad before drawing. Copper was chosen as the clad material by virtue of its excellent deformability and compatibility with the die during the cold drawing process. The copper clad thorium tubes were drawn and the copper cladding eliminated by dissolution in nitric acid. Thorium metal itself does not dissolve in nitric acid due to quick passivation. Tubes of ID as small as 9.5 mm and wall thickness 0.7 mm were produced. The process of mechanical cladding yielded tubes with wrinkles on the OD which was a cause for rejection. This author had experience in electro less plating and electroplating in the making of Microwave Integrated Circuits which required deposition of copper on alumina substrates [54]. Hence it was decided to try copper deposition on to the thorium tube as an alternative to mechanical cladding. Electro less deposition yields very small thicknesses and used when the cathode is nonconductive. It also requires a number of chemicals for the pre-treatment and plating processes. Electroplating was chosen at the time as a suitable rectifier was readily available.

Copper was deposited in acidic bath using copper sulphate at room temperature. Two anodes were required, a pipe for covering the OD of the thorium tube and a copper rod for ID. The two anodes and the cathode were positioned concentrically maintaining appropriate distances by using a Perspex fixture to hold them in place. Too low a current density yielded coarse copper deposit while a too high a current density yielded a dark deposit. Once these limits are avoided, a large window of current densities is available from plating literature. Separate power supplies were used for the internal anode and external anode to enable independent control of current densities. One would normally expect an increase in plating rate with an increase in current density. But this was not the case with respect to plating on the inner diameter of the thorium tube. The anodic current density for the internal anode was maintained at about half of that of the external anode. This ensured copper deposition of uniform thickness on the inside and outside of the thorium tube. This also prevented passivation of the copper rod and co-deposition of unwanted elements from the bath on to the thorium. The details are shown schemati-

cally in **Figure 1**. A similar method was developed later at Hyderabad by this author to deposit copper on zirconium alloy billets before hot extrusion [55].

Thorium is difficult to plate on. It is due to a thin naturally forming oxide surface film that is often difficult to remove and that reforms quickly when a cleaned surface is exposed to air or water. As a result, adherent electrodeposits are obtained only when either: 1) the oxide film is removed for sufficient time to permit an initial deposit; 2) the film is replaced with another that does not interfere with adhesion; 3) the film is incorporated into the deposit in a compatible manner; or 4) the surface is severely etched to allow mechanical keying between the substrate and deposit [56]. Successful procedures that rely on one or more of the above principles have been documented. However, for our purpose of cold drawing, maximum adhesion has not been aimed at.

4. Thorium Dioxide Pelletizing

Uranium oxide powder was being compacted and sintered on tonnage scale in India since the early seventies. When the requirement of thoria pellets was received, the process parameters of uranium oxide were used for pressing and sintering thorium oxide [15]. However, this led to problems. For example, while a green density of 5.6 to 5.8 g/cc worked fine with uranium oxide, the thorium oxide pellets pressed to the same green density yielded pellets that chipped readily on handling. The thorium oxide powder was supplied by India Rare Earths and it was oxalate derived. It was found that the as received powders were not suitable for compacting directly. Therefore a ball milling step was introduced. Ball milled powder yielded green pellets of density over 6.4 g/cc and these had better green strength than pellets made from unmilled powders and pressed to 5.8 g/cc. The pellets on sintering yielded high sintered densities [3] [4].

However it became difficult to control the air radio-activity levels in the milling process in spite of containment. The separation of the balls from the milled powder too posed a problem. Hence the milling process was

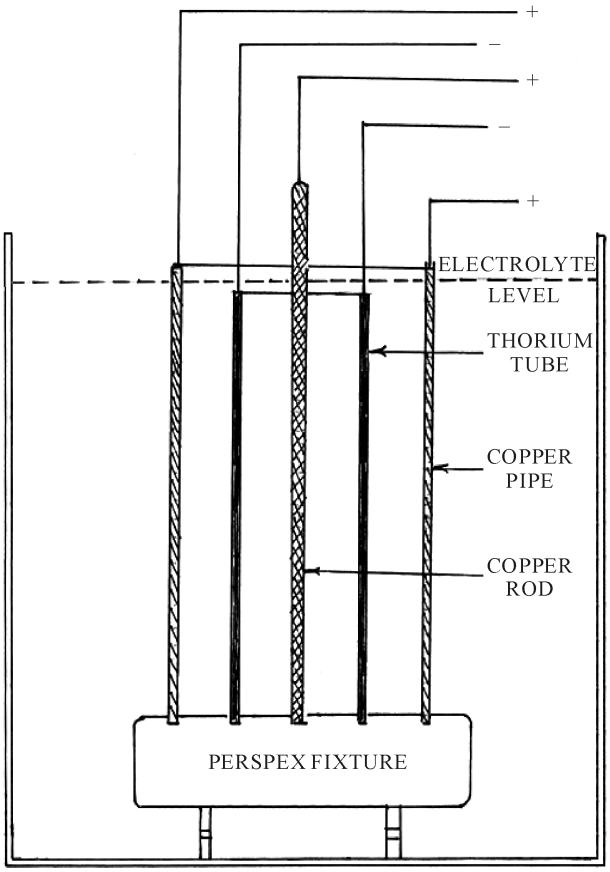


Figure 1. Fixture for holding thorium metal tube, copper rod (internal anode) and copper pipe (external anode).

discontinued. 1% by weight Zinc behenate binder was admixed into the thoria powder to yield strong green pellets and finally dense sintered pellets.

5. Thorium Dioxide Sintering

5.1. Additive Effects on Sinterability

In our early work, it was found that sintering of thoria needed a sintering aid. MgO was being used as sintering aid. $MgSO_4$ was being added to thorium nitrate solution and co-precipitated as oxalate. The oxalate was calcined to yield thorium oxide with about 250 ppm of Mg.

The thoria pellets were sintered in the same furnaces that were being used for sintering UO_2 pellets, namely, electrically heated molybdenum element furnaces that operated at 1600˚C - 1700˚C in a reducing atmosphere. Later it was found by this author that thorium oxide could be sintered in air at 1150˚C using 0.25 mole percent Nb_2O_5 [3]-[6]. Pellets sintered in a reducing atmosphere ranged in color from bluish gray to white. Those sintered in air ranged from brownish white to white in color. Sintering at 1700˚C required round the clock shift operation while sintering at 1150˚C in air could be done in two shifts.

5.2. Black Spots in Sintered Thoria

It has been reported that on prolonged heating to 1800˚C to 1900˚C in vacuum, thoria blackens with loss of oxygen, although the loss is insufficient to be reflected in chemical analysis or lattice parameter measurements. On reheating in air to 1200˚C or 1300˚C, the white color is restored [57]. Thoria can also become oxygen deficient due to lower valence cation impurities which leave oxygen sites in the crystal lattice vacant. Black spots were noticed by this author in the thoria pellets sintered in hydrogen at 1700˚C. On analysis it was found to have been caused by iron contamination from the wear out of steel balls in the milling step. This pointed to the need for use of wear resistant balls such as those made of agate or tungsten carbide.

In one of the experiments, a one mm diameter stainless steel wire piece was introduced in a green thoria pellet. On sintering in hydrogen, a 3 mm diameter black patch resulted. Spectrographic analysis of the black patch revealed the presence of iron, chromium and nickel, the ingredients of stainless steel. The black color is believed to be due to oxygen deficiency in the thoria lattice caused by substitution of Th sites by lower valent Fe, Cr or Ni. Niobia added thoria pellets sintered in hydrogen at 1700˚C were totally black in color. The niobium in reducing atmosphere might have acted as a lower valence additive to thoria.

5.3. The Phenomenon of Coring

Thoria pellets could be readily thermal etched to be able to see well defined grain boundaries. A Laser etching technique has been developed for etching thoria for micro structural investigation in reprocessing [24]. Sometimes a cored structure is visible in fractured sintered ceramic pellets. Two types of coring have been mentioned in literature. In one type, reported in thoria and called “picture framing”, the grain size in the core was smaller than that in the rim [39]. In another type, reported in alumina, the grain size in the core is much larger than that in the rim [58]. Some of our sintered thoria pellets upon fracturing exhibited a cored structure. The section of the pellet showed a white rim around a glassy inside. In a pellet of 25 mm diameter, the rim thickness varied from pellet to pellet from 1 mm to 3 mm. Zawdzki and Apte [59] studied coring in UO_2 sintered pellets. They found that a cored structure would result when the heating rate was 400˚C/hr and when the UO_2 had sulphur impurity of 20 ppm or above. There was no coring when the heating rate was lower, at 250˚C/hr for the same material. Coring in Mg doped Alumina was attributed to grain growth inhibition within the cored region. The Mg was thought to have escaped from the rimmed region leading to discontinuous grain growth.

The cored structure appears to be caused by either additives or impurities. Additives are meant for inhibiting grain growth, as MgO in Al_2O_3 and ThO_2. The absence of the additive in the rimmed region of the pellet causes discontinuous grain growth leading to the formation of large grains in the rim [58]. Impurities such as Sulphur are capable of causing discontinuous grain growth and their presence in the core region of the pellet causes discontinuous grain growth and large grains in the core [59]. An analysis of the core region of our sintered thoria pellets pointed to the presence of sulphur content above 500 ppm. The origin of the sulphur may be in the $MgSO_4$ added to thorium nitrate to provide Mg as sintering aid. The Sulphur appears to have caused discontinuous grain growth leading to larger grain sizes in the cored region. Recalcination of the thoria powder batches

that had shown coring on sintering earlier did not show coring again. This author recommends mechanical addition of MgO to ThO_2 powder instead of $MgSO_4$ to $Th(NO_3)_4$ and co-precipitation.

The green density in a compact in uniaxial or biaxial pressing is usually not uniform unlike in isostatic pressing. The rim is of a higher green density than the inside. At fast rates of heating, and for larger diameter pellets, there is a chance that the surface temperature is higher than the core temperature. The skin with a larger number of particle to particle contacts is likely to densify first leading to subsequent constrained sintering of the core. The situation is further aggravated by the absence of grain growth inhibitors or presence of discontinuous grain growth promoters in certain regions of the pellet.

5.4. Evaporation and Condensation of Thorium Dioxide

Two types of sintering furnaces were used for sintering thoria. One type was the pusher type continuous sintering furnace. The other type was a batch sintering furnace. In some sintering runs, the bottom of the water cooled stainless steel lid of the batch sintering furnace was found to have got coated with a white powdery layer, which was identified as thorium oxide. ThO_2 seems to have evaporated in the reducing atmosphere from the pellets being sintered and deposited on the underside of the lid.

6. Electrical Conductivity of Thorium Dioxide with Additives

Electrical conductivity of sintered thoria was determined by this author at the Plasma Physics Division of BARC [8]. Holes were drilled in the thoria disks prepared for conductivity measurements. Initially, the pellets could not be drilled using even diamond drills. Then it was realized that the drill bits were getting damaged due to quick overheating. The thoria disks fully immersed in water could be readily drilled using drill bits of High Speed Steel. For fixing the contact wires to the sintered thoria disc, platinum paste was used. The electrical conductivity of niobia doped thoria was found to be lower than that of magnesia doped thoria in the temperature range 900˚C to 1250˚C. From defect chemistry considerations, it can be expected that the effect of addition of higher valence Nb_2O_5 to ThO_2 would be to enhance the diffusion of thorium in ThO_2 and to decrease electrical conductivity. The effect of addition of (lower valence) MgO would be to decrease diffusion and enhance the electrical conductivity. The higher diffusivities of otherwise slowly moving cation mean that a lower sintering temperature would suffice for niobium doped ThO_2 than for magnesium doped ThO_2. This has indeed been found to be the case, the sintering temperatures being 1150˚C and 1700˚C, respectively [5].

7. Dissolution of Thorium Dioxide

It is generally known that it is very difficult to dissolve sintered thoria rejects for recycling in the pelletizing plant as well as in reprocessing after irradiation. At Hyderabad it was found that 1 g of sintered thoria readily dissolved in 10 ml of hot concentrated nitric acid with one drop of hydrofluoric acid. In cases where the HF content was larger, the dissolution either slowed down or stopped. It appears that in higher HF concentrations, a coating of ThF_4 over the ThO_2 might be hindering further dissolution. Niobia doped air sintered thoria required longer time to dissolve compared to hydrogen sintered magnesia doped thoria [9]. Dissolution concerns may have to be addressed while reprocessing after irradiation.

8. Summary

Electroplating of copper instead of mechanical cladding resulted in greater acceptance rates of cold drawn thorium metal tubes. The compaction and sintering parameters for thorium oxide were found to be somewhat different from those of uranium oxide. While the conventional additive MgO required a sintering temperature of 1700˚C, achievable in high temperature sintering furnaces with molybdenum elements in a reducing atmosphere, Nb_2O_5 brought down the sintering temperature to 1150˚C which could be achieved in ordinary furnaces with air atmosphere. Electrical conductivity measurements on thoria with Mg and Nb additives were made and interpreted with respect to sintering behavior. Coring, that is, the existence of non uniform grain size between the rim of the sintered thoria pellet and the inside is likely when the pellet diameter is large (25 mm) as in the case of CIRUS research reactor pellets, when the heating rates are rapid (400˚C/hr) and when there is impurity such as sulphur. Mechanical addition of MgO to ThO_2 powder instead of $MgSO_4$ to $Th(NO_3)_4$ for co-precipitation is recommended.

Acknowledgements

The author is grateful to Mr. R. N. Jayaraj, Chief Executive, Nuclear Fuel Complex for granting permission to publish this paper.

References

[1] Balakrishna, P., Varma, B.P., Krishnan, T.S., Mohan, T.R.R. and Ramakrishnan, P. (1986) The Role of Point Defects in the Sintering of UO_2 and ThO_2. *Transactions of the Powder Metallurgy Association of India*, **13**, 117-125.

[2] Balakrishna, P., Varma, B.P., Rajendran, R. and Balaramamoorthy, K. (1987) Thorium Oxide Blanket Fabrication for Indian Fast Breeder Test Reactor. In: *Proceedings of TCM, Utilization of Thorium Based Nuclear Fuel*: *Current Status and Perspectives*, **7**, 73-88.

[3] Balakrishna, P., Varma, B.P., Krishnan, T.S., Mohan, T.R.R. and Ramakrishnan, P. (1987) Developments in Thoria Calcination, Compaction and Sintering. *Transactions of the Powder Metallurgy Association of India*, **14**, 76-82.

[4] Balakrishna, P., Varma, B.P., Krishnan, T.S., Mohan, T.R.R. and Ramakrishnan, P. (1988) Thorium Oxide: Calcination, Compaction and Sintering. *Journal of Nuclear Materials*, **160**, 88-94. http://dx.doi.org/10.1016/0022-3115(88)90012-8

[5] Balakrishna, P., Varma, B.P., Krishnan, T.S., Mohan, T.R.R. and Ramakrishnan, P. (1988) Low Temperature Sintering of Thoria. *Journal of Materials Science Letters*, **7**, 657-660. http://dx.doi.org/10.1007/BF01730326

[6] Balakrishna, P., Varma, B.P., Rajendran, R., Krishnan, T.S., Balaramamoorthy, K., Mohan, T.R.R. and Ramakrishnan, P. (1989) Activated Sintering of Thoria. *Engineering Materials*, **29**, 669-676.

[7] Balakrishna, P., Somayajulu, G.V.S.R.K., Krishnan, T.S., Mohan, T.R.R. and Ramakrishnan, P. (1991) Developments in Low Temperature Sintering of Thoria. In: Vincenzini, P., Ed., *Ceramics Today—Tomorrow's Ceramics*, *Materials Science Monograph* 66*D*, Elsevier, Amsterdam, 2995-3001.

[8] Balakrishna, P., Ananthapadmanabhan, V. and Ramakrishnan, P. (1994) Electrical Conductivity of Sintered Niobia Doped and Magnesia Doped Thoria. *Journal of Materials Science Letters*, **13**, 86-88. http://dx.doi.org/10.1007/BF00416807

[9] Balakrishna, P. (1994) Characterization and Sintering of Thorium Dioxide. Ph.D. Thesis, Indian Institute of Technology, Bombay.

[10] Balakrishna, P. (2012) ThO_2 and $(U,Th)O_2$ Processing—A Review. *Natural Science*, **4**, 943-949. http://dx.doi.org/10.4236/ns.2012.431123

[11] Kantan, S.K., Raghavan, R.V. and Tendolkar, G.S. (1958) Sintering of Thorium and Thoria. *Proceedings of International Conference on Peaceful Uses of Atomic Energy*, Geneva, 1-13 September 1958.

[12] Bichakal, S.S., Kini, V.S., Raghavan, R.V. and Garg, R.K. (1974) Preparation of Sinterable Grade Thoria Powder. *Proceedings of the Indo-German Seminar on Utilization of Thorium*, Bombay, 18-22 November 1974.

[13] Gurumurthy, K.R., Paibhale, S.V., Singh, S.P. and Raghavan, R.V. (1976) Fabrication of Thorium Shapes by Powder Metallurgy. *Transactions of Powder Metallurgy Association of India*, **3**, 47-49.

[14] Rodriguez, P. and Sundaram, C.V. (1981) Nuclear and Materials Aspects of the Thorium Fuel Cycle. *Journal of Nuclear Materials*, **100**, 227-249. http://dx.doi.org/10.1016/0022-3115(81)90534-1

[15] Das, M. and Prasad, P.N. (1984) Thorium Dioxide Pellets for Reactor Fuel Elements. Specification PP-M-784.

[16] Chandramouli, V., Anthonysamy, S., Vasudeva Rao, P.R., Divakar, R. and Sudararaman, R. (1996) PVA Aided Microwave Synthesis: A Novel Route for the Production of Nanocrystalline Thoria Powder. *Journal of Nuclear Materials*, **231**, 213-220. http://dx.doi.org/10.1016/0022-3115(96)00368-6

[17] Joseph, K., Sridharan, R. and Gnanasekaran, T. (2000) Kinetics of Thermal Decomposition of $Th(C_2O_4)_2 \cdot 6H_2O$. *Journal of Nuclear Materials*, **281**, 129-139. http://dx.doi.org/10.1016/S0022-3115(00)00241-5

[18] Purohit, R.D., Saha, S. and Tyagi, A.K. (2001) Nanocrystalline Thoria Powders via Glycine-Nitrate Combination. *Journal of Nuclear Materials*, **288**, 7-10. http://dx.doi.org/10.1016/S0022-3115(00)00717-0

[19] Ananthasivan, K., Anthonysamy, S., Sudha, C., Terrance, A.L.E. and Vasudeva Rao, P.R. (2002) Thoria Doped with Cations of Group VB—Synthesis and Sintering. *Journal of Nuclear Materials*, **300**, 217-229. http://dx.doi.org/10.1016/S0022-3115(01)00736-X

[20] Ananthasivan, K., Anthonysamy, S., Singh, A. and Vasudeva Rao, P.R. (2002) De-Agglomeration of Thorium Oxalate—A Method for the Synthesis of Sinter Active Thoria. *Journal of Nuclear Materials*, **306**, 1-9. http://dx.doi.org/10.1016/S0022-3115(02)01229-1

[21] Balakrishnan, K., Majumdar, S., Ramanujam, A. and Kakodkar, A. (2002) The Indian Perspective on Thorium Fuel

Cycles. Thorium Fuel Cycle: Options and Trends, IAEA TECDOC-1319, 257-265.

[22] Dash, S., Singh, A., Ajikumar, P.K., Subramaniam, H., Rajalakshmi, M., Tyagi, A.K., Arora, A.K., Narasimhan, S.V. and Raj, B. (2002) Synthesis and Characterization of Nano Crystalline Thoria Obtained from Thermally Decomposed Thorium Carbonate. *Journal of Nuclear Materials*, **303**, 156-168. http://dx.doi.org/10.1016/S0022-3115(02)00816-4

[23] Kutty, T.R.G., Khan, K.B., Hegde, P.V., Banerjee, J., Sengupta, A.K., Majumdar, S. and Kamath, H.S. (2004) Development of a Master Sintering Curve for ThO_2. *Journal of Nuclear Materials*, **327**, 211-219. http://dx.doi.org/10.1016/j.jnucmat.2004.02.007

[24] Sinha, S., Ramadasan, E., Jatha, V.P., Dasgupta, K., Sahoo, K.C. and Gantayet, L.M. (2007) Laser Etching of Thoria Pellets for Microstructural Investigations. *Applied Science*, **253**, 4404-4408.

[25] Anantharaman, K., Shivakumar, V. and Saha, D. (2008) Utilization of Thorium in Reactors. *Journal of Nuclear Materials*, **383**, 119-121. http://dx.doi.org/10.1016/j.jnucmat.2008.08.042

[26] Banerjee, J., Kutty, T.R.G., Kumar, A., Kamath, H.S. and Banerjee S. (2011) Densification Behavior and Sintering Kinetics of ThO_2-4%UO_2 Pellet. *Journal of Nuclear Materials*, **408**, 224-230. http://dx.doi.org/10.1016/j.jnucmat.2010.11.029

[27] Banerjee, S. and Govindan Kutty, T.R. (2012) Functional Materials. *Nuclear Fuels*, **10**, 387-466.

[28] Saller, H.A. and Keeler, J.R. (1953) The Fabrication of Thorium Tubes. BMI-805, Batelle Memorial Institute.

[29] Johnson, J.R. and Curtis, C.E. (1954) Properties of Thorium Oxide Ceramics. Interim Report ORNL 1809.

[30] Johnson, J.R. and Curtis, C.E. (1954) Note on Sintering of Thoria. *Journal of the American Ceramic Society*, **37**, 611. http://dx.doi.org/10.1111/j.1151-2916.1954.tb13996.x

[31] Beach, J.G. and Schaer, G.R. (1959) Electroplating on Thorium. *Journal of the Electrochemical Society*, **106**, 392-393. http://dx.doi.org/10.1149/1.2427366

[32] Harada, Y., Baskin, Y. and Handwerk, J.H. (1962) Calcination and Sintering Study of Thoria. *American Ceramic Society*, **45**, 253-257. http://dx.doi.org/10.1111/j.1151-2916.1962.tb11139.x

[33] (1966) Utilization of Thorium in Power Reactors. Technical Reports Series No. 52, IAEA, Vienna.

[34] Matzke, H.J. (1967) Xenon Migration and Trapping in Doped ThO_2. *Journal of Nuclear Materials*, **21**, 190-198. http://dx.doi.org/10.1016/0022-3115(67)90149-3

[35] Pope, J.M. and Radford, K.C. (1974) Physical Properties of Some Thoria Powders and Their Influence on Sinterability. *Journal of Nuclear Materials*, **52**, 241-254. http://dx.doi.org/10.1016/0022-3115(74)90171-8

[36] Weinreich, A.W., Britton, W.H., Hutchison, C.R., Johnson, R.G.R. and Burke, T.J. (1977) Fabrication of High Density, High Integrity Thorium Based Fuel Pellets. *Transactions of the American Nuclear Society*, **27**, 305-307.

[37] (1979) Status and Prospects of Thermal Breeders and Their Effect on Fuel Utilization, Technical Reports Series No. 195, IAEA, Vienna.

[38] White, G.D., Bray, L.A. and Hart, P.E. (1981) Optimization of Thorium Oxalate Precipitation Conditions Relative to Derived Oxide Sinterability. *Journal of Nuclear Materials*, **96**, 305-313. http://dx.doi.org/10.1016/0022-3115(81)90574-2

[39] Belle, J. and Berman, R.M. (1981) Thorium Dioxide—Properties and Nuclear Applications. DOE/NE-0060. http://dx.doi.org/10.2172/5986642

[40] (1987) Utilization of Thorium-Based Nuclear Fuel: Current Status and Perspectives. *Proceedings of TCM*, IAEA-TECDOC-412, IAEA, Vienna.

[41] Tetsuo Shiratori, K. and Fukuda, K. (1993) Fabrication of Very High Density Fuel Pellets of Thorium Dioxide. *Journal of Nuclear Materials*, **202**, 98-103. http://dx.doi.org/10.1016/0022-3115(93)90033-U

[42] Altas, Y., Eral, M. and Tel, H. (1997) Preparation of Homogeneous $(Th_{0.8}U_{0.2})O_2$ Pellets via Coprecipitation of $(Th,U)(C_2O_4)_2 \cdot nH_2O$ Powders. *Journal of Nuclear Materials*, **249**, 46-51. http://dx.doi.org/10.1016/S0022-3115(97)00185-2

[43] (2000) IAEA-TECDOC-1155. IAEA, Vienna.

[44] (2002) IAEA-TECDOC-1319. IAEA, Vienna.

[45] MacDonald, P.E. (2002) Advanced Proliferation Resistant, Lower Cost, Uranium—Thorium Dioxide Fuels for Light Water Reactors, Idaho National Engineering and Environmental Laboratory INEEL/EXT-02-01411.

[46] IAEA (2005) Thorium Fuel Cycle—Potential Benefits and Challenges. IAEA, Vienna.

[47] Grosse, K.-H., Hrovat, M. and Seemann, R. (2009) Manufacturing Technology for Thorium Based Fuel Elements. *Proceedings of Conference on Characterization and Quality Control of Nuclear Fuels*, Hyderabad, 18-20 February 2009.

[48] Hingant, N., Clavier, N., Dacheux, N., Barre, N., Hubert, S., Obbade, S., Taborda, F. and Abraham, F. (2009) Preparation, Sintering and Leaching of Optimized Uranium Thorium Dioxides. *Journal of Nuclear Materials*, **385**, 400-406. http://dx.doi.org/10.1016/j.jnucmat.2008.12.011

[49] Hingant, N., Clavier, N., Dacheux, N., Hubert, S., Barré, N., Podor, R. and Aranda, L. (2011) Preparation of Morphology Controlled $Th_{(1-x)}U_xO_2$ Sintered Pellets from Low Temperature Precursors. *Powder Technology*, **208**, 454-460. http://dx.doi.org/10.1016/j.powtec.2010.08.042

[50] (2012) Role of Thorium to Supplement Fuel Cycles of Future Nuclear Energy Systems. STI/PUB/1540, IAEA, Vienna.

[51] Csom, Gy., Reiss, T., Fehér, S. and Czifrus, Sz. (2012) Thorium. *Annals of Nuclear Energy*, **41**, 67-78.

[52] Das, D. and Bharadwaj, S.R. (2013) Thoria Based Nuclear Fuels: Thermophysical and Thermodynamic Properties, Fabrication, Reprocessing and Waste Management. Google eBook: Business & Economics Springer Science & Business Media.

[53] Balakrishnan, M.R. (1992) India Prepares for Its First Load of Thorium Fuel. *Nuclear Engineering International*, **37**, 14.

[54] Balakrishna, P. and Mahapatra, S. (1973) Electroless Copper on Alumina and Other Substrates. *Plating*, **60**, 636-639.

[55] Balakrishna, P., Murty, B.N., Jha, S.K., Krishnan, P.K. and Saibaba, N. (2005) Electroplating of Copper on Annular Zirconium Alloy Billets. *Plating and Surface Finishing*, **92**, 56-61.

[56] Dini, J.W. and Johnson, H.R. (1980) Plating on Some Difficult to Plate Metals and Alloys. SAND 79-8069, Sandia Laboratories.

[57] Peterson, S., Adams, R.E. and Douglas Jr., D.A. (1965) Properties of Thorium, Its Alloys and Its Compounds. ORNL-TM-1144. http://web.ornl.gov/info/reports/1965/3445601336962.pdf

[58] Burke, J.E., Lay, K.W. and Prochazka, S. (1980) The Effect of MgO on the Mobility of Grain Boundaries and Pores in Aluminum Oxide. In: Kuczynski, G.C., Ed., *Sintering Processes*, Plenum Press, London.

[59] Zawidzki, T.W., Apte, P.S. and Hoare, M.R. (1984) Effect of Sulphur on Grain Growth in UO_2 Pellets. *Journal of the American Ceramic Society*, **67**, 361-364. http://dx.doi.org/10.1111/j.1151-2916.1984.tb19537.x

Natural Science, 2015, 7, 18-31
Published Online January 2015 in SciRes. http://www.scirp.org/journal/ns
http://dx.doi.org/10.4236/ns.2015.71003

Earthquake Barcode from a Single-Degree-of-Freedom System

Chein-Shan Liu[1], Chih-Wen Chang[2]

[1]Department of Civil Engineering, National Taiwan University, Taipei, Taiwan
[2]Cloud Computing and System Integration Division, National Center for High-Performance Computing, Taichung, Taiwan
Email: liucs@ntu.edu.tw

Received 4 November 2014; revised 12 December 2014; accepted 29 December 2014

Abstract

Earthquake is a violent and irregular ground motion that can severely damage structures. In this paper we subject a single-degree-of-freedom system, consisting of spring and damper, to an earthquake excitation, and meanwhile investigate the response behavior from a novel theory about the dynamical system $\dot{\mathbf{x}}=\mathbf{f}(\mathbf{x},t)$, by viewing the time-varying signum function of $\|f\|^2\|x\|^2-2(f\cdot x)^2$. It can reflect the characteristic property of each earthquake through $(x(t),\dot{x}(t))$ and the second component $f_2=-\ddot{u}_g(t)-2\zeta\omega_n\dot{x}(t)-\omega_n^2x(t)$ of f, where $\ddot{u}_g(t)$ is a time-sampling record of the acceleration of a ground motion. The barcode is formed by plotting $\mathrm{Sign}:=\mathrm{sign}\left(\|\mathbf{f}\|^2\|\mathbf{x}\|^2-2(\mathbf{f}\cdot\mathbf{x})^2\right)$ with respect to time. We analyze the complex jumping behavior in a barcode and an essential property of a high percentage occupation of the first set of dis-connectivity in the barcode from four strong earthquake records: 1940 El Centro earthquake, 1989 Loma earthquake, and two records of 1999 Chi-Chi earthquake. Through the comparisons of four earthquakes, we can observe that strong earthquake leads to large percentage of the first set of dis-connectivity.

Keywords

Earthquakes, Single-Degree-of-Freedom System, Signum Function, Barcode, Jumping Behavior, The First Set of Dis-Connectivity, Scale Invariance

1. Introduction

Earthquakes occur around the world frequently. The most are so small that they cannot be susceptible, and only

a few are large enough to cause a noticeable damage of building. When a structure is subjected to an earthquake motion, its base foundation tends to move with the ground.

One of the most important applications of structural dynamics is by analyzing the response of structures to ground shaking caused by an earthquake, whose data are useful for establishing the structural response spectra. The identification and evaluation of ground motion parameters require access to the measurements of strong ground motions in an actual earthquake [1]-[3].

The basic element of an accelerograph is a transducer element as described in [4] [5], which in its simplest form is a single-degree-of-freedom mass-damper-spring system. Typically, the transducer has natural frequency f_n = 25 Hz and damping ratio ζ = 60% for modern analog accelerographs, and f_n = 50 Hz and ζ = 70% for modern digit accelerographs [4].

A simple stationary representation of earthquake-induced ground acceleration was proposed by Kanai [6] and Tajimi [7], based on the study of frequency content of a number of strong ground motion records. They suggested that the ground acceleration of the earth surface layer could be approximated by the absolute acceleration of a simple oscillator with a concentrated mass supported by a linear spring and a dashpot. The spectrum property of an earthquake is reflected in its interaction with the structures which have different damping ratios and natural periods. In general, one can use the response spectrum, power spectrum or the Fourier spectrum to characterize the spectrum property of an earthquake. In this paper we propose a novel method to characterize the earthquake property when it is exerted on a simple one-degree-of-freedom system:

$$\ddot{x}(t)+2\zeta\omega_n\dot{x}(t)+\omega_n^2 x(t)=p(t), \tag{1}$$

where $p(t)=-\ddot{u}_g(t)$ is the ground acceleration due to an earthquake; ζ is the damping ratio; and ω_n is the natural frequency.

The ground motion produced by earthquake can be very complicated. The exact form of earthquakes is subjected to a high degree of uncertainty and complexity. There are several ways to estimate the ground motion parameters, which are essential for describing the important characteristics of strong ground motion in a compact quantitative form. Because of the complexity of the earthquake ground motion, the identification of a single parameter that can accurately describe all important ground motion characteristics is impossible. In this paper we try to use the barcode and the percentage of the first set of dis-connectivity in the barcode to reflect the characteristic property of an earthquake ground motion. It is known that a barcode is an optical machine-readable representation of data relating to the object to which it is attached. A main feature of the barcode is the intervened black lines and white lines with varying spacings and widths. Barcode is ubiquitously used in the identification and classification of human made products. Here we will use the barcode to identify the property of earthquake with the input $p(t)=-\ddot{u}_g(t)$ in Equation (1).

2. A System Formulation

To facilitate the formulation we write a system of first-order ordinary differential equations (ODEs):

$$\dot{\mathbf{x}}=\mathbf{f}(\mathbf{x},t),\quad t\in\mathbb{R},\quad \mathbf{x}\in\mathbb{R}^n. \tag{2}$$

Here n = 2, $\mathbf{x}=(x,y)=(x,\dot{x})$, and two components of $\mathbf{f}$ are $f_1=y$ and $f_2=p(t)-\omega_n^2 x-2\zeta\omega_n y$, respectively.

As that done by Liu [8], for Equation (2) we can define a unit orientation vector:

$$\mathbf{n}:=\frac{\mathbf{x}}{\|\mathbf{x}\|}, \tag{3}$$

where $\|\mathbf{x}\|:=\sqrt{\mathbf{x}\cdot\mathbf{x}}>0$ is the Euclidean norm of $\mathbf{x}$, and the dot between two vectors, say $\mathbf{x}\cdot\mathbf{y}$, denotes the inner product of $\mathbf{x}$ and $\mathbf{y}$. Then we can derive [9]-[11]

$$\dot{\mathbf{x}}=\frac{\mathbf{f}\cdot\mathbf{x}}{\|\mathbf{x}\|^2}\mathbf{x}+\left[\frac{\mathbf{f}}{\|\mathbf{x}\|}\otimes\frac{\mathbf{x}}{\|\mathbf{x}\|}-\frac{\mathbf{x}}{\|\mathbf{x}\|}\otimes\frac{\mathbf{f}}{\|\mathbf{x}\|}\right]\mathbf{x}, \tag{4}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{x}\|=\mathbf{f}\cdot\frac{\mathbf{x}}{\|\mathbf{x}\|}. \tag{5}$$

Equations (4) and (5) can be put together as

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\mathbf{x}\\ \|\mathbf{x}\|\end{bmatrix}=\begin{bmatrix}\dfrac{\mathbf{f}}{\|\mathbf{x}\|}\otimes\dfrac{\mathbf{x}}{\|\mathbf{x}\|}-\dfrac{\mathbf{x}}{\|\mathbf{x}\|}\otimes\dfrac{\mathbf{f}}{\|\mathbf{x}\|} & \dfrac{(\mathbf{f}\cdot\mathbf{x})\mathbf{x}}{\|\mathbf{x}\|^3}\\ \dfrac{(\mathbf{f}\cdot\mathbf{x})\mathbf{x}^{\mathrm{T}}}{\|\mathbf{x}\|^3} & 0\end{bmatrix}\begin{bmatrix}\mathbf{x}\\ \|\mathbf{x}\|\end{bmatrix}, \tag{6}$$

which is a Lie-type system endowing with the Lie-group symmetry $SO_o(n, 1)$ in the Minkowski space.

3. Two Branches Solutions

In order to develop a numerical scheme from Equation (6), we suppose that the coefficient matrix is constant with the pair

$$(\mathbf{a},\mathbf{b})=\left(\frac{\mathbf{f}}{\|\mathbf{x}\|},\frac{\mathbf{x}}{\|\mathbf{x}\|}\right) \tag{7}$$

being constant in a small time stepsize.

From Equations (6) and (7) we thus need to solve a constant linear system:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{bmatrix}\mathbf{x}\\ \|\mathbf{x}\|\end{bmatrix}=\begin{bmatrix}\mathbf{a}\otimes\mathbf{b}-\mathbf{b}\otimes\mathbf{a} & \mathbf{a}\cdot\mathbf{b}\mathbf{b}\\ \mathbf{a}\cdot\mathbf{b}\mathbf{b}^{\mathrm{T}} & 0\end{bmatrix}\begin{bmatrix}\mathbf{x}\\ \|\mathbf{x}\|\end{bmatrix}. \tag{8}$$

Let

$$z=\mathbf{a}\cdot\mathbf{x},\quad w=\mathbf{b}\cdot\mathbf{x},\quad y=\|\mathbf{x}\|, \tag{9}$$

$$c_0=\mathbf{a}\cdot\mathbf{b}, \tag{10}$$

and Equation (8) becomes

$$\dot{\mathbf{x}}=w\mathbf{a}-z\mathbf{b}+c_0y\mathbf{b}, \tag{11}$$

$$\dot{y}=c_0w. \tag{12}$$

At the same time, from the above equations we can derive the following ODEs for z, w and y:

$$\frac{\mathrm{d}}{\mathrm{d}t}\begin{pmatrix}z\\ w\\ y\end{pmatrix}=\begin{pmatrix}-c_0 & a_0^2 & c_0^2\\ -1 & c_0 & c_0\\ 0 & c_0 & 0\end{pmatrix}\begin{bmatrix}z\\ w\\ y\end{bmatrix}, \tag{13}$$

where $a_0=\|\mathbf{a}\|$. Fortunately, the original $(n+1)$-dimensional problem in Equation (8) can be reduced to a three-dimensional problem in Equation (13). Depending on the signum function of

$$\mathrm{Sign}:=\mathrm{sign}\left(a_0^2-2c_0^2\right)=\mathrm{sign}\left(\|\mathbf{x}\|^2\|\mathbf{f}\|^2-2(\mathbf{f}\cdot\mathbf{x})^2\right), \tag{14}$$

there exist two different types solutions of (z, w, y). Here, we do not give a detailed derivation of the solutions for (z, w, y) and the corresponding second generation group preserving scheme (GPS2), but the reader can refer [9]-[11].

4. The Demonstration of Barcodes

It is significant that in Equation (14) we have derived a signum function to demand the algorithm into two branches. Without having the factor 2 before $\mathbf{f}\cdot\mathbf{x}$ in Equation (14) one has

$$\|\mathbf{f}\|^2\|\mathbf{x}\|^2-(\mathbf{f}\cdot\mathbf{x})^2=\|\mathbf{f}\|^2\|\mathbf{x}\|^2\left(1-\cos^2\theta\right)\geq 0,$$

where θ is the intersection angle between $\mathbf{x}$ and $\mathbf{f}$; hence, it makes no sense to define its signum function. On the contrast, by Equation (14) we have

$$\text{Sign} = \text{sign}\left(\|\mathbf{f}\|^2 \|\mathbf{x}\|^2 \left(1-2\cos^2\theta\right)\right) = -\text{sign}\left(\cos 2\theta\right), \tag{15}$$

which may be +1 or −1, depending on **x** and **f** by their intersection angle θ. When θ is in the range of $-\pi/2 < \theta < \pi/2$ or $3\pi/4 < \theta < 5\pi/4$, the value of Sign is Sign = −1; otherwise, Sign = +1. Thus we can observe the time-varying values of Sign and plot them as a barcode. In order to demonstrate the use of barcode, let us investigate the following examples. In all computations given below the initial conditions are set to be $\left(x(0), \dot{x}(0)\right) = (0,0)$, unless specifying otherwise.

4.1. Example 1

First we test an example with $\omega_n = \sqrt{2}$ and $\zeta = 0.15/\sqrt{2}$, and a harmonic force $p(t) = \cos t$ is applied to Equation (1) in a time interval of $t \in [0,100]$. The input, the following signum function:

$$\text{Sign}(t) = \text{sign}\left\{ f_1^2(t) + f_2^2(t) + x^2(t) + y^2(t) - 2\left[x(t) f_1(t) + y(t) f_2(t)\right]^2 \right\}, \tag{16}$$

where $f_1(t) = y(t)$ and $f_2(t) = p(t) - \omega_n^2 x(t) - 2\zeta\omega_n y(t)$, and the responses of $x(t)$ and $y(t)$ are plotted in **Figure 1**. It is interesting that after 0.78 sec, Sign = +1 for the harmonic input. As compared with the closed-form solution, in **Figure 1(d)** we plot the numerical error, which is quite accurate, and is smaller than the time stepsize h = 0.001 used in this computation.

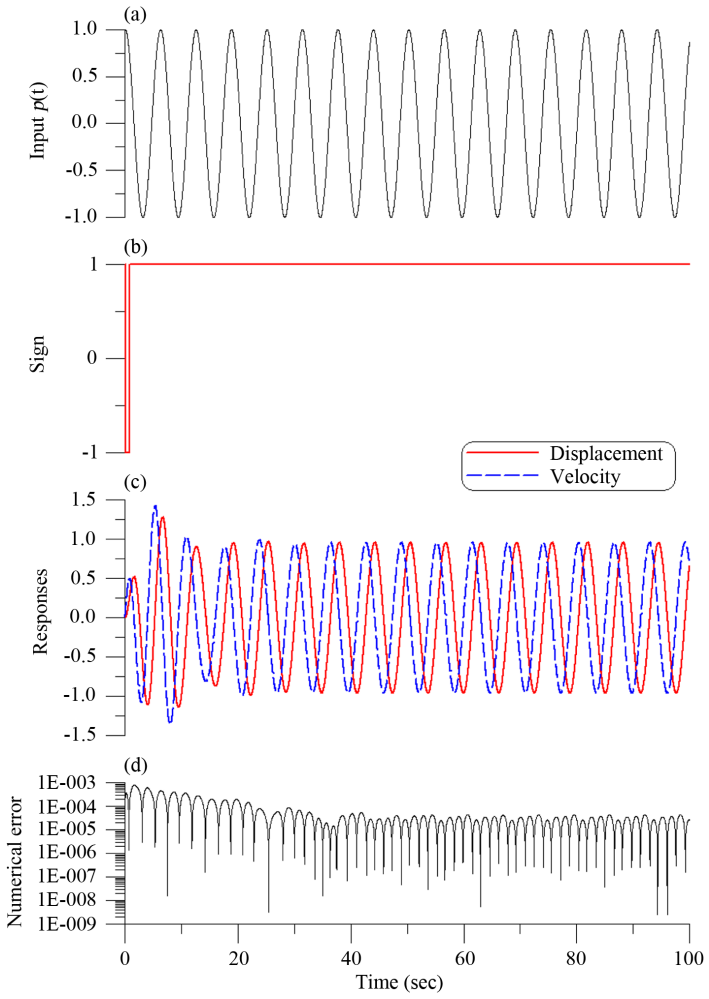


Figure 1. Under (a) a harmonic force, showing (b) the signum function, (c) the responses, and (d) numerical error of displacement.

4.2. Example 2

Then with $\omega_n = \sqrt{3}$ and $\zeta = 0.1/\sqrt{3}$, and under a random force

$$p(t_i) = 5R_i, \tag{17}$$

where R_i are random numbers between [−1, 1], the input, the Sign and the responses are plotted in **Figure 2**. It is interesting that the values of Sign are randomly jumping between +1 and −1 in some intervals of time.

4.3. Analyzing the Barcode under Earthquakes

With $\omega_n = 2\pi/1.5$ and $\zeta = 0.02$, and under the El Centro earthquake $\ddot{u}_g(t)$, which is adjusted to $p(t) = -0.98\ddot{u}_g(t)\ \mathrm{m/sec}^2$ as the input to Equation (1), the barcode and the responses are plotted in **Figure 3**. As shown in **Figure 3(b)**, the time history of Sign looks like a barcode, which is interesting that the values of Sign are frequently jumping between +1 and −1 with a certain structure.

Now we analyze the Sign and the length as shown in **Figure 4** within a short time interval of $t \in [0, 2]$, which is starting from the initial conditions $(x(0), \dot{x}(0)) = (0.1, 0.01)$, where we take $\omega_n = \sqrt{2}$ and $\zeta = 0.01/\sqrt{2}$ and $p(t) = -50\ddot{u}_g(t)$ being the input. First we note that the length as governed by Equation (5) has the following property:

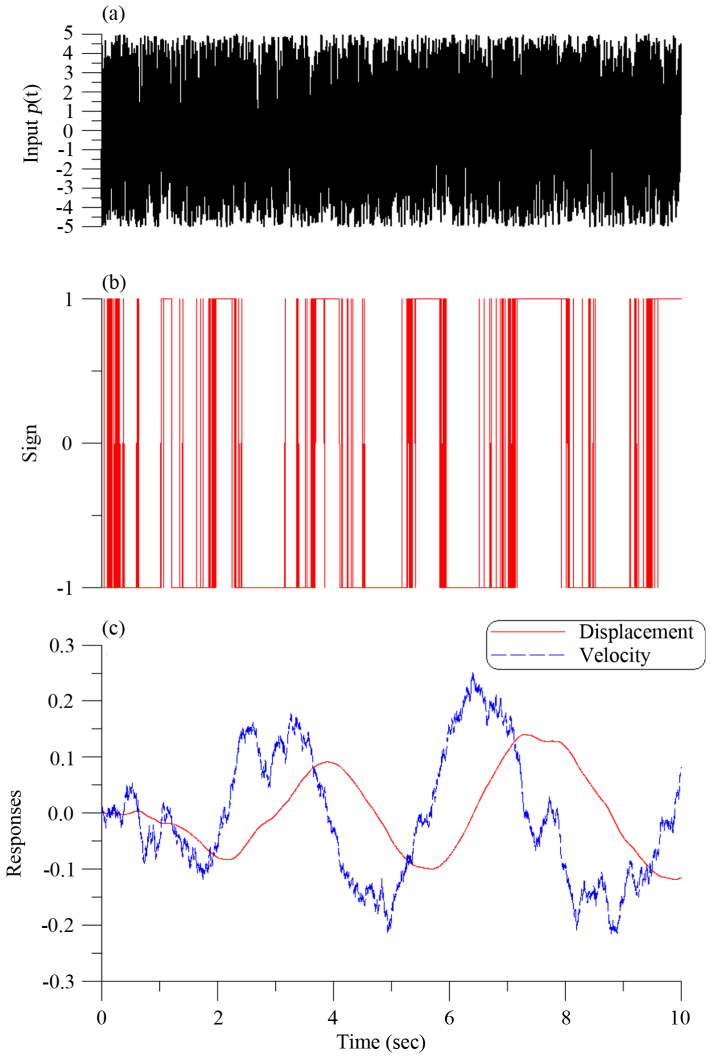


Figure 2. Under (a) a random force, showing (b) the signum function, and (c) the responses.

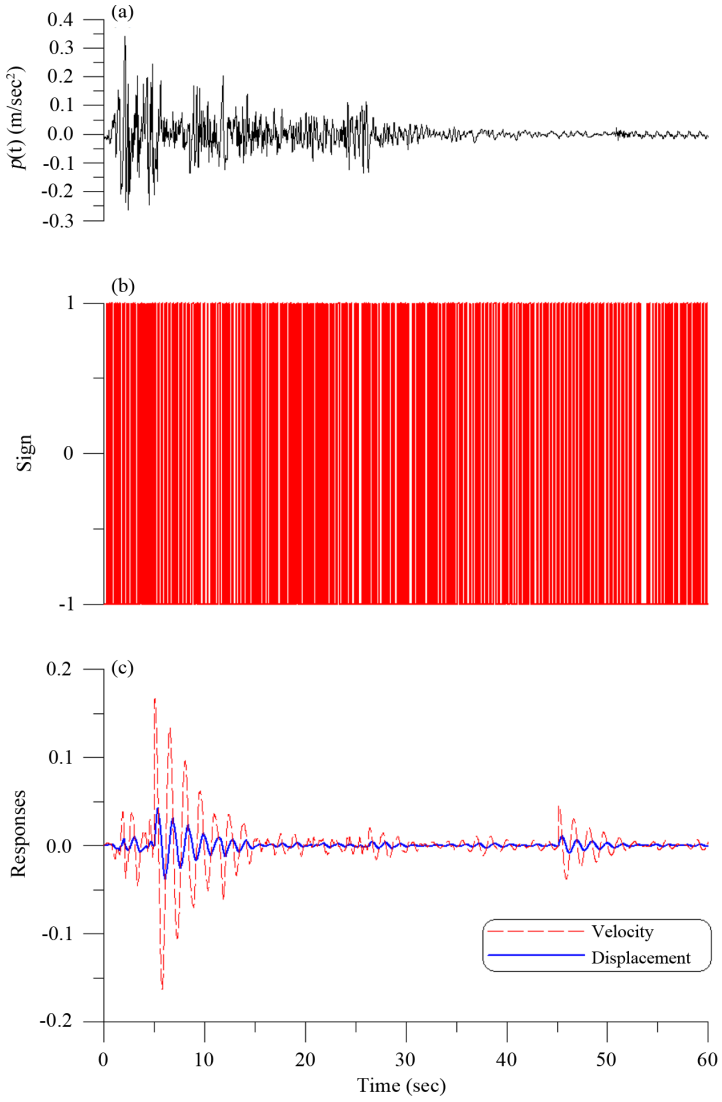


Figure 3. Under (a) the El Centro earthquake, showing (b) the barcode, and (c) the responses.

$$\operatorname{sign}(c_0) \underset{>}{<} 0 \Leftrightarrow \frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{x}\| \underset{>}{<} 0, \tag{18}$$

where $\operatorname{sign}(c_0) = \operatorname{sign}(\mathbf{f}\cdot\mathbf{x})$. Then, as a schematic plot shown in **Figure 5**, there are two dis-connected sets of $\|\mathbf{f}\|^2\|\mathbf{x}\|^2 - 2(\mathbf{f}\cdot\mathbf{x})^2 < 0$ in the hyper-plane (**x**, **f**):

$$\mathbf{f}\cdot\mathbf{x} > \frac{1}{\sqrt{2}}\|\mathbf{f}\|\|\mathbf{x}\|, \tag{19}$$

$$\mathbf{f}\cdot\mathbf{x} < -\frac{1}{\sqrt{2}}\|\mathbf{f}\|\|\mathbf{x}\|. \tag{20}$$

We may call them *the first set of dis-connectivity* and the second set of dis-connectivity, respectively. Clearly, the first set of dis-connectivity is a subset of sign(c_0) = +1, and the second set of dis-connectivity is a subset of sign(c_0) = −1.

If Equation (2) satisfies

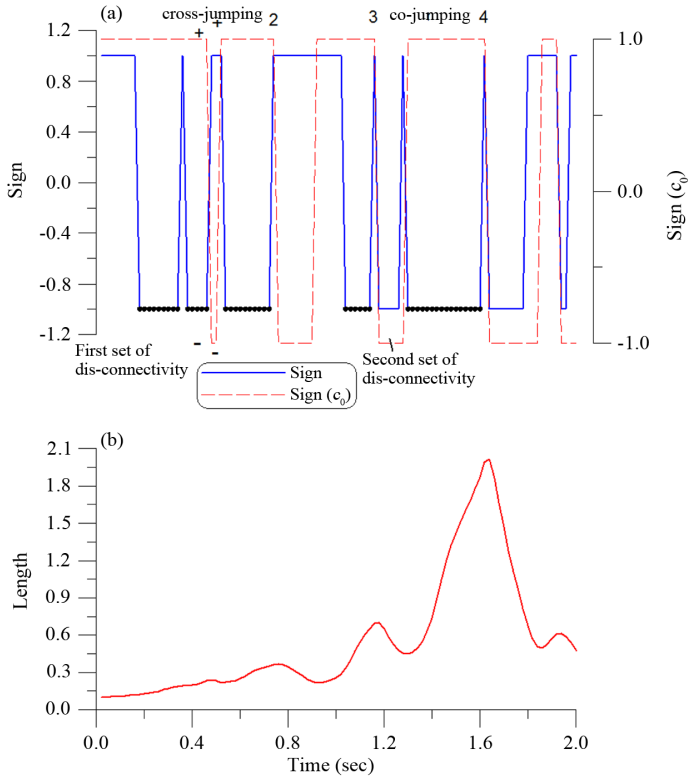


Figure 4. Under the El Centro earthquake, (a) showing two signum functions, and (b) the length with a small time interval.

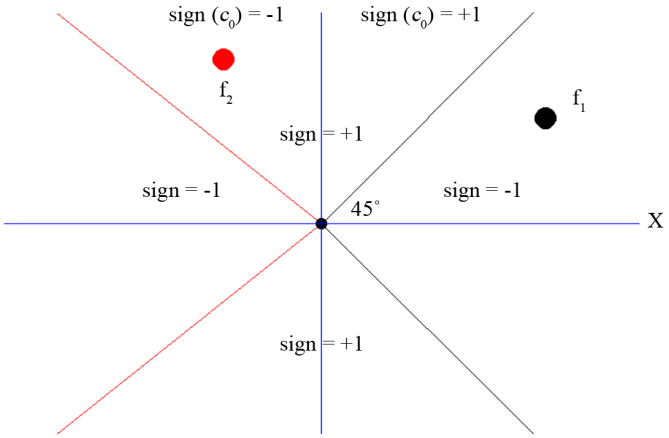


Figure 5. A schematic plot of the area of Sign = +1, Sign = −1, sign(c_0) = +1, and sign(c_0) = −1 in the plane.

$$\mathbf{f}\left(\mathbf{x}(t_0),t_0\right)\cdot\mathbf{x}(t_0)>0,\ \text{ and }\ \operatorname{sign}\left(\|\mathbf{f}\|^2\|\mathbf{x}\|^2-2(\mathbf{f}\cdot\mathbf{x})^2\right)=-1,\ \ \forall t\geq t_0,\ \text{ then }\mathbf{x}(t)\to\infty,\ \ t\to\infty. \tag{21}$$

Under the first assumption the case in Equation (20) is impossible because it contradicts to $\mathbf{f}\left(\mathbf{x}(t_0),t_0\right)\cdot\mathbf{x}(t_0)>0$. Then under the condition of $\operatorname{sign}\left(\|\mathbf{f}\|^2\|\mathbf{x}\|^2-2(\mathbf{f}\cdot\mathbf{x})^2\right)=-1$, it is always

$$\mathbf{f}\cdot\mathbf{x}>\frac{1}{\sqrt{2}}\|\mathbf{f}\|\|\mathbf{x}\|,\ \ \forall t\geq t_0, \tag{22}$$

because the two sets of dis-connectivity are dis-connected, and from the first set to the second set it must go to $\text{sign}\left(\|\mathbf{f}\|^2\|\mathbf{x}\|^2-2(\mathbf{f}\cdot\mathbf{x})^2\right)=+1$ in some time interval. Then using Equations (5) and (22) we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{x}\|>\frac{1}{\sqrt{2}}\|\mathbf{f}\|>0, \tag{23}$$

which means that the length grows with time. Thus, Equation (21) is proven.

For a system being excited by earthquake the value of Sign is not always +1 as that shown in **Figure 1(b)** for the harmonic input. As shown in **Figure 3(b)** the values of Sign are varying between −1 to +1 and +1 to −1 very fast, and then the Sign will return to +1 again as shown in **Figure 4(a)**; otherwise, by Equation (21) the system will respond unstably and thus the displacement will tend to infinity. In order to compare the values of sign(c_0) with the values of Sign we plot them in **Figure 4(a)** with solid line and dashed line, respectively, where we make a slight shift of the dashed line downward for clear. On the other hand, for a distinction the first set of dis-connectivity is filled by solid black points.

As shown in **Figure 5**, in the plane the areas of sign(c_0) = +1 and sign(c_0) = −1 are connected while that the sets of Sign = +1 and Sign = −1 are dis-connected as mentioned in the above. When sign(c_0) changes from +1 to −1, for example from point $\mathbf{f}_1$ to point $\mathbf{f}_2$ in the figure, the Sign must jump from −1 to +1. This jumping behavior is termed a *cross jumping*. On the other hand, we also find the *co-jumping* behavior which is happened when both the Sign and sign(c_0) are jumping from −1 to +1 simultaneously. The above two jumping behaviors are collected as a *coupled-jumping*, which renders a quite complex structure of the barcode as shown in **Figure 3(b)**.

From **Figure 4(a)** we can observe the following interesting phenomena: 1) While the first set of dis-connectivity is a subset of sign(c_0) = +1, the second set of dis-connectivity is a subset of sign(c_0) = −1. 2) In a time interval of the state with Sign = −1, the state is either in the first set of dis-connectivity or in the second set of dis-connectivity. 3) From one first set of dis-connectivity to another first set of dis-connectivity there must accompany a jump from Sign = −1 to Sign = +1. This also holds for the second set of dis-connectivity. 4) From one first set of dis-connectivity to one second set of dis-connectivity, or *vice-versa*, there must accompany a jump from Sign = −1 to Sign = +1. For some cases this jumping only happens at one time point. 5) When sign(c_0) jumps from +1 to −1 as remarked in **Figure 4(a)** by the symbols + and − (*i.e.*, the length is decreased), and if the Sign is in the state of −1, then the Sign will jump from −1 to +1 as remarked in **Figure 4(a)** by the symbols − and +. Indeed in **Figure 4(a)** we can observe that there are four times to happen the cross jumping behavior in the interval of $t\in[0,2]$. Besides the first one, the other three times are marked by the Arabic numbers 2, 3 and 4 in **Figure 4(a)**. Besides, there is also a co-jumping as remarked in **Figure 4(a)**. The proof of 5) is obvious by viewing **Figure 5** that the first set of Sign = −1 cannot directly jump to the second set of Sign = −1. When $\mathbf{f}_1$ goes to $\mathbf{f}_2$, the Sign changes from −1 to +1.

In **Figure 6** we compare the barcodes with 1) $\omega_n=2\pi/1.5$, $\zeta=0.02$ and $p(t)=-0.98\ddot{u}_g(t)\,\mathrm{m/sec^2}$, 2) $\omega_n=2\pi/1.5$, $\zeta=0.02$ and $p(t)=-4.9\ddot{u}_g(t)\,\mathrm{m/sec^2}$, 3) $\omega_n=2\pi/1.5$, $\zeta=0.05$ and $p(t)=-0.98\ddot{u}_g(t)\,\mathrm{m/sec^2}$, and 4) $\omega_n=2\pi/2$, $\zeta=0.02$ and $p(t)=-0.98\ddot{u}_g(t)\,\mathrm{m/sec^2}$.

4.4. The First Set of Dis-Connectivity under Earthquakes

In this example we let the Chi-Chi earthquake be the input to Equation (1). With $\omega_n=2\pi/1.5$ and $\zeta=0.02$, and under the Chi-Chi earthquake $\ddot{u}_g(t)$ recorded at the station of CHY028, we adjust it to be $p(t)=-0.98\ddot{u}_g(t)\,\mathrm{m/sec^2}$ as shown in **Figure 7(a)**. The barcode and the responses are plotted in **Figure 7(b)** and **Figure 7(c)**.

In **Figure 8** we compare the barcodes with 1) $\omega_n=2\pi/1.5$, $\zeta=0.02$ and $p(t)=-0.98\ddot{u}_g(t)\,\mathrm{m/sec^2}$, 2) $\omega_n=2\pi/1.5$, $\zeta=0.02$ and $p(t)=-4.9\ddot{u}_g(t)\,\mathrm{m/sec^2}$, 3) $\omega_n=2\pi/1.5$, $\zeta=0.05$ and $p(t)=-0.98\ddot{u}_g(t)\,\mathrm{m/sec^2}$, and 4) $\omega_n=2\pi/0.5$, $\zeta=0.02$ and $p(t)=-0.98\ddot{u}_g(t)\,\mathrm{m/sec^2}$.

In **Figure 9(a)** for the Chi-Chi earthquake with different scaling factor f_r in $p(t)=-9.8f_r\ddot{u}_g(t)$ as being the input to Equation (1), and with $\omega_n=2\pi/1.5$ and $\zeta=0.02$, we plot the percentage in the first set of dis-connectivity in **Figure 9(b)**, which is near to forty eight percentage, and at the same time we also display the numbers

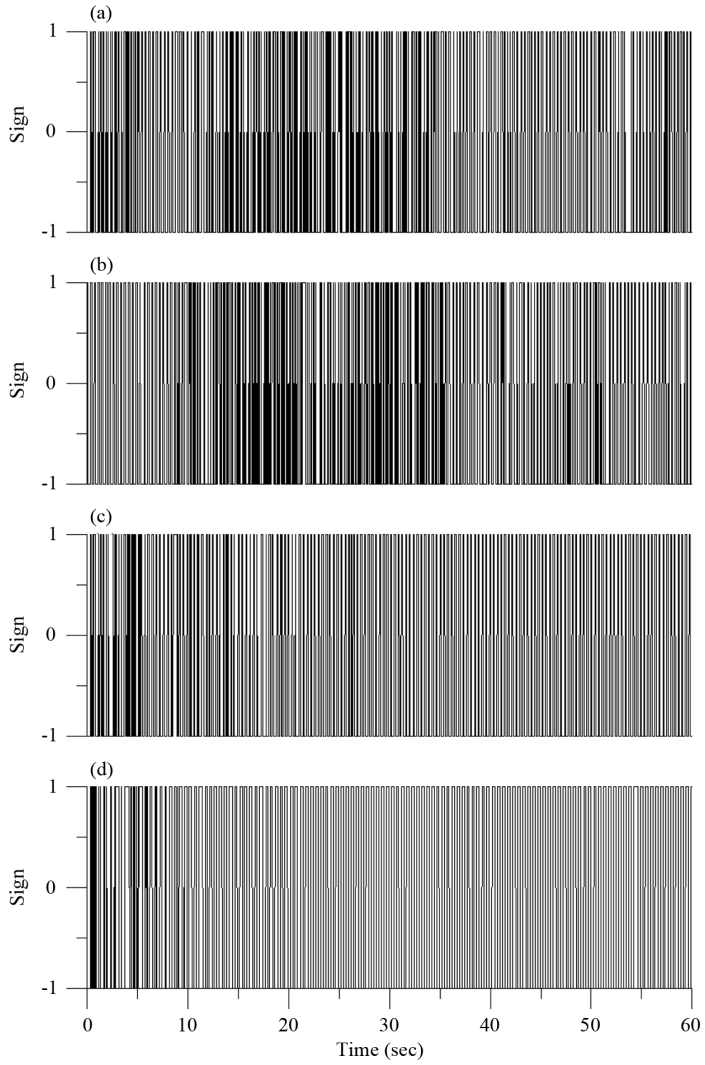


Figure 6. Under the El Centro earthquake, comparing the barcodes under different scaling factor, damping ratio and natural period.

of coupled-jumping as discussed above with respect to the scaling factor f_r in **Figure 9(c)**. It can be seen that the number is either 24 or 25. In the plane as shown in **Figure 5**, the first set of dis-connectivity with Sign = −1 is only in one quarter. But under a strong earthquake, the high occupation of the first set of dis-connectivity is near to fifty percentage as shown in **Figure 9(b)**.

In **Figure 10**, under the Chi-Chi earthquake (CHY080) as shown in **(a)**, we plot the first set of dis-connectivity which is near to forty nine percentage as shown in **Figure 10(b)**, and the number of coupled-jumping is plotted in **Figure 10(c)**, which is either 6 or 7.

In **Figure 11(a)** for the El Centro earthquake (ELC180) with different scaling factor f_r in $p(t) = -9.8 f_r \ddot{u}_g(t)$, and with $\omega_n = \pi$ and $\zeta = 0.02$, we plot the first set of dis-connectivity which is near to forty six percentage as shown in **Figure 11(b)**, and the number of coupled-jumping is plotted in **Figure 11(c)**, which is 0 or 1 for most scaling factors. In **Figure 12(a)** for the Loma Prieta earthquake (47379 Gilroy Array # 1) with different scaling factor f_r in $p(t) = -9.8 f_r \ddot{u}_g(t)$, and with $\omega_n = \pi$ and $\zeta = 0.02$, we plot the first set of dis-connectivity which is near to forty six percentage as shown in **Figure 12(b)**, and the number of coupled-jumping is plotted in **Figure 12(c)**, which is between 1 to 7.

It is surprising that the percentage of the first set of disconnectivity and the number of coupled-jumping are almost constant for each earthquake, which seems to be scale-invariant. In the plane as shown in **Figure 5**, the

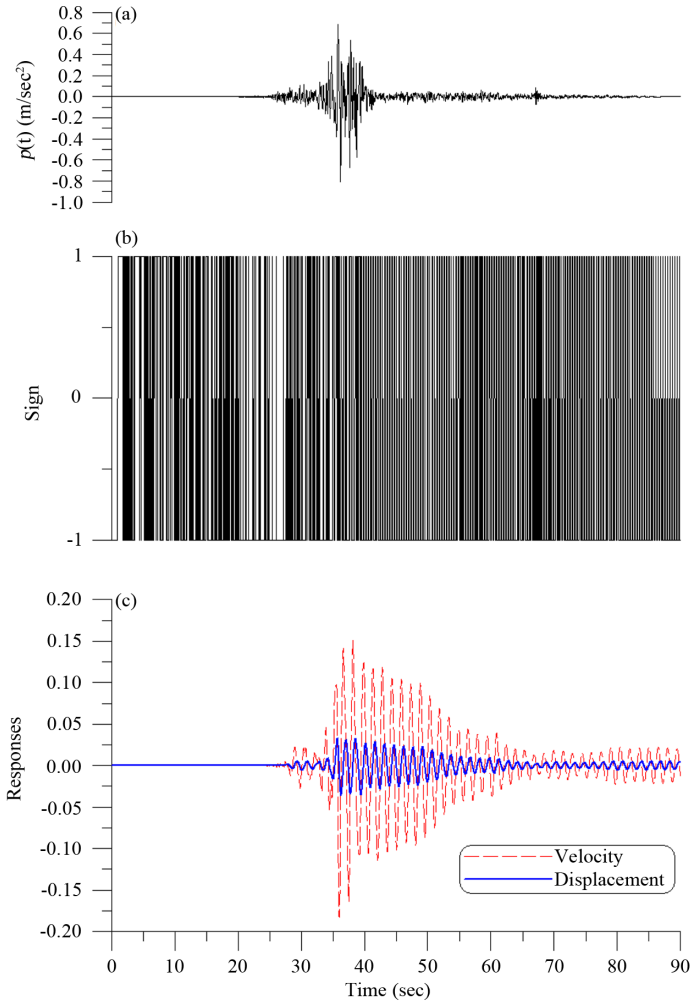


Figure 7. Under the Chi-Chi earthquake (CHY028) in (a), (b) showing the barcode, and (c) the responses.

first set of dis-connectivity with Sign = −1 is only in one quarter. But under strong earthquakes, the high occupation of the first set of dis-connectivity is near to fifty percentage. The high percentage of occupation demonstrates the importance of the first set of dis-connectivity in the strong earthquake motion. As comparing the above four earthquakes we can also observe that stronger earthquake leads to larger percentage of the first set of dis-connectivity. Finally, under the same scaling factor f_r = 0.5 in $p(t) = -9.8 f_r \ddot{u}_g(t)$ and the same $\omega_n = \pi$ and $\zeta = 0.02$, in **Figure 13** we compare the relative percentages of first set of dis-connectivity for the El Centro earthquake (ELC180), the Loma Prieta earthquake (47379 Gilroy Array #1), and two records (CHY028, CHY080) of Chi-Chi earthquake. The relative percentage is a time history of the time-duration of the first set of dis-connectivity, whose time is dividing by the total time of earthquake. This figure reflects that the Chi-Chi earthquake has larger values than that of the El Centro and the Loma Prieta earthquakes, and CHY080 is larger than CHY028. This can also be explained by calculating the power of earthquake with the ℓ_2-norm of earthquake, of which the power value of CHY080 is 0.7765, of CHY028 is 0.6193 and of the El Centro earthquake is 0.3444. The Loma Prieta earthquake (47379 Gilroy Array #1) leads to a close curve with that produced by the El Centro earthquake, whose power is 0.33.

5. Conclusions

For a single-degree-of-freedom system subjected to an earthquake excitation, we have investigated the response

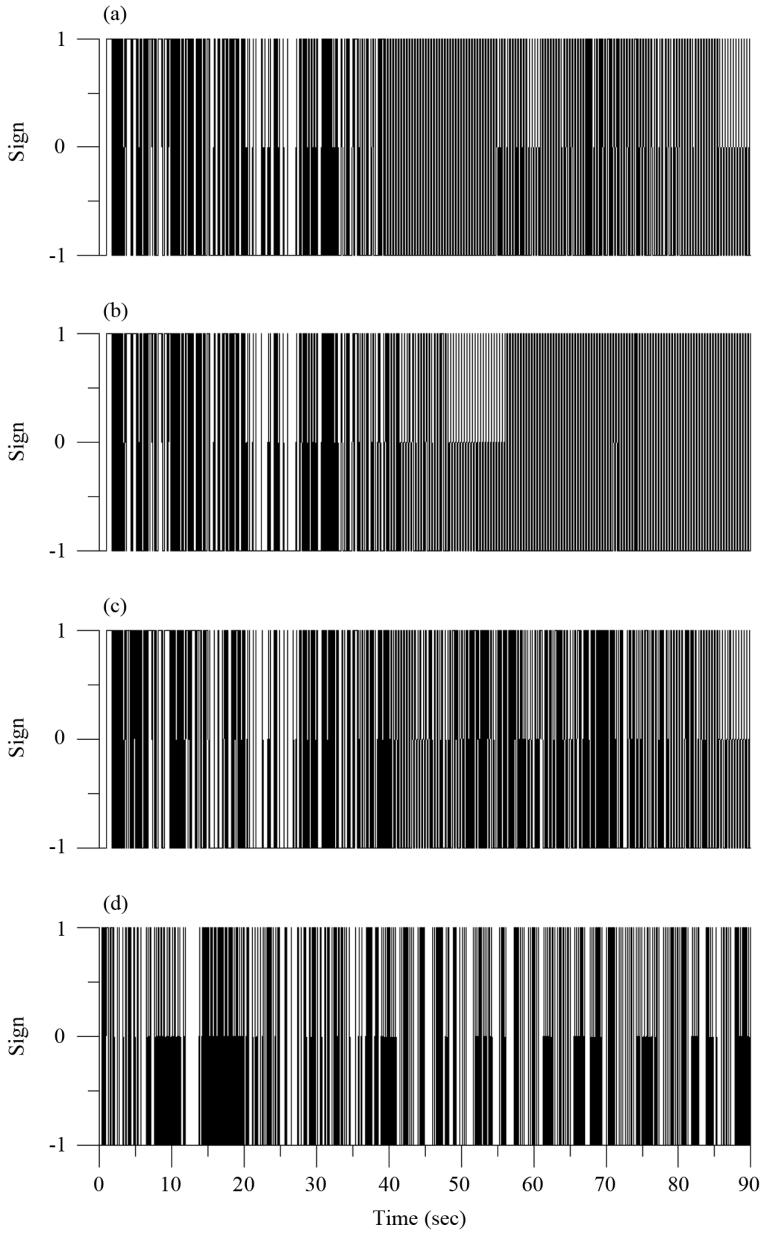


Figure 8. Under the Chi-Chi earthquake (CHY028), comparing the barcodes under different scaling factor, damping ratio and natural period.

behavior from a newly developed theory about the dynamical system of $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t)$ by viewing the signum function of $\|\mathbf{f}\|^2 \|\mathbf{x}\|^2 - 2(\mathbf{f} \cdot \mathbf{x})^2$. The main idea in this paper is that the earthquake will make the function **f** highly complex in time, even the linear system is simple in **x**. Under this situation, we expect that the signum function being a time varying function will disclose the complexity of earthquake, such that we can study its structure. The barcode was formed by plotting the signum function Sign with respect to the time history of an earthquake. We have analyzed the jumping behavior from two famous earthquake inputs of 1940 El Centro earthquake and 1999 Chi-Chi earthquake. Several barcodes were used to reveal a quite complex structure of earthquake, and we have demonstrated the importance of the first set of dis-connectivity, whose high percentage occupation is very interesting when one considers the linear system under a strong earthquake. Through the comparisons of four earthquake inputs, we have observed that stronger earthquake leads to larger percentage of the first set of dis-connectivity. The coupled-jumping behavior may happen due to the random property of

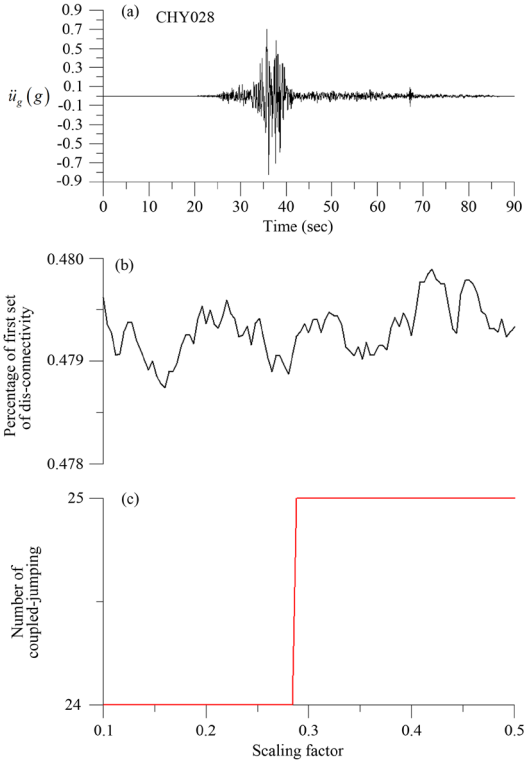


Figure 9. Under (a) the Chi-Chi earthquake (CHY028) with different scaling factor, showing (b) the percentage in the first set of dis-connectivity, and (c) the number of coupled-jumping.

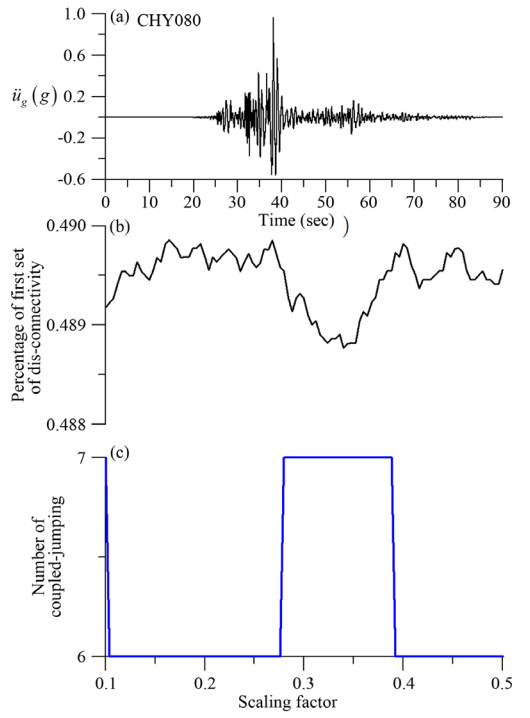


Figure 10. Under (a) the Chi-Chi earthquake (CHY080) with different scaling factor, showing (b) the percentage in the first set of dis-connectivity, and (c) the number of coupled-jumping.

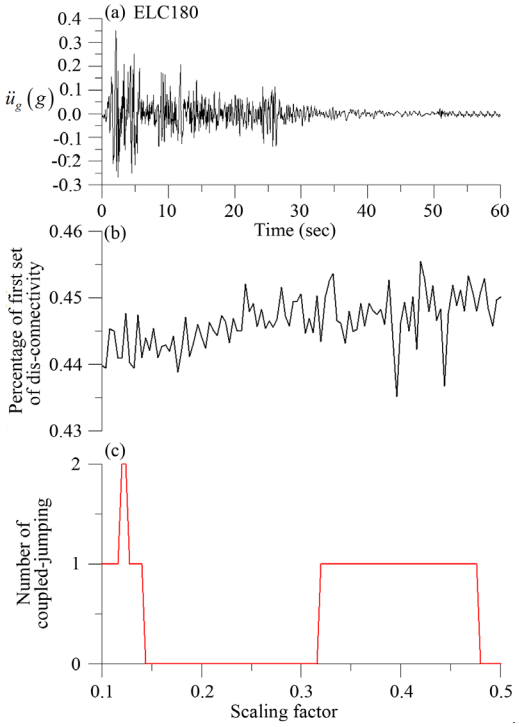


Figure 11. Under (a) the El Centro earthquake (ELC180) with different scaling factor, showing (b) the percentage in the first set of dis-connectivity, and (c) the number of coupled-jumping.

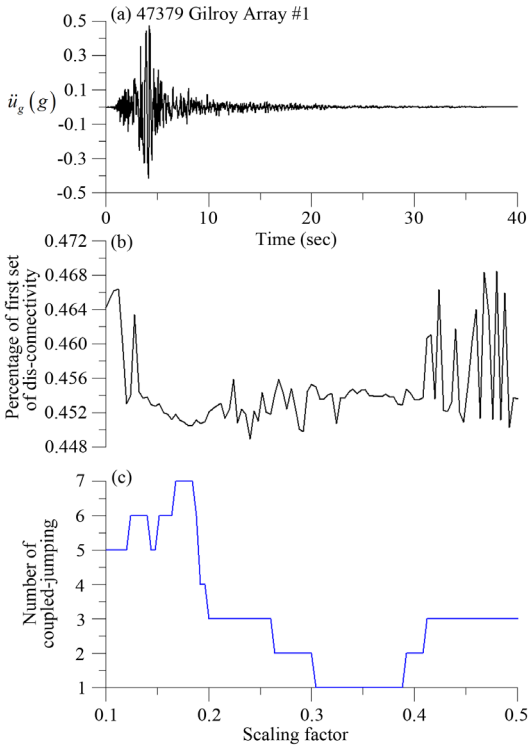


Figure 12. Under (a) the 1989 Loma Prieta earthquake (47379 Gilroy Array #1) with different scaling factor, showing (b) the percentage in the first set of dis-connectivity, and (c) the number of coupled-jumping.

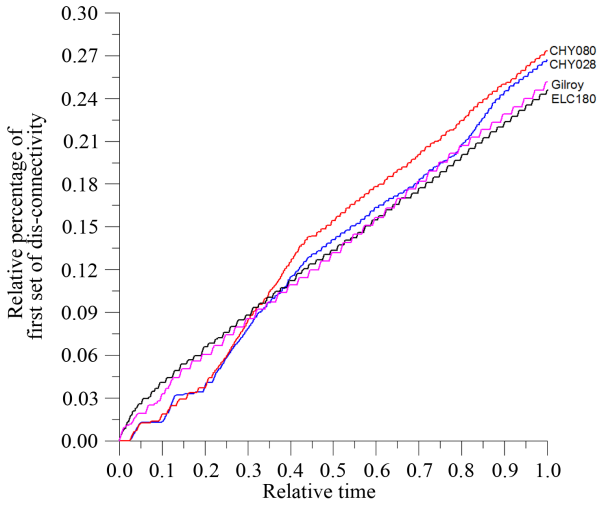


Figure 13. Comparing the relative percentages of first set of dis-connectivity for the El Centro earthquake (ELC180), two records (CHY028, CHY080) of Chi-Chi earthquake, and the Loma Prieta earthquake (47379 Gilroy Array #1).

earthquakes, and its number seems scale-independence. This is just an initial study by correlating the percentage of the first set of dis-connectivity to the earthquake, and there are still many works to study the barcode and the use of it to forecast the intensity of earthquake before the rapid coming of main shock impulse. In the future, we will analyze the signal of barcode for many earthquake records, and thus we can correlate the intensity and frequency content of an earthquake to the first set of dis-connectivity.

References

[1] Loh, C.H., Lee, Z.K., Wu, T.C. and Peng, S.Y. (2000) Ground Motion Characteristics of the Chi-Chi Earthquake of 21 September 1999. *Earthquake Engineering and Structural Dynamics*, **29**, 867-897. http://dx.doi.org/10.1002/(SICI)1096-9845(200006)29:6<867::AID-EQE943>3.0.CO;2-E

[2] Sokolov, V.Y., Loh, C.H. and Wen, K.L. (2002) Comparison of the Taiwan Chi-Chi Earthquake Strong-Motion Data and Ground-Motion Assessment Based on Spectral Model from Smaller Earthquakes in Taiwan. *Bulletin of the Seismological Society of America*, **92**, 1855-1877. http://dx.doi.org/10.1785/0120010178

[3] Sokolov, V.Y., Loh, C.H. and Wen, K.L. (2003) Evaluation of Hard Rock Spectral Models for the Taiwan Region on the Basis of the 1999 Chi-Chi Earthquake Data. *Soil Dynamics and Earthquake Engineering*, **23**, 715-735. http://dx.doi.org/10.1016/S0267-7261(03)00075-7

[4] Chopra, A.K. (1995) Dynamics of Structures: Theory and Applications to Earthquake Engineering. Prentice-Hall, New Jersey.

[5] Kramer, S.L. (1996) Geotechnical Earthquake Engineering. Prentice-Hall, New Jersey.

[6] Kanai, K. (1957) Semi-Empirical Formula for the Seismic Characteristics of the Ground. *Bulletin of the Earthquake Research Institute, University of Tokyo*, **35**, 309-325.

[7] Tajimi, H. (1960) A Statistical Method of Determining the Maximum Responses of a Building Structure during an Earthquake. *Proceeding of the 2nd World Conference on Earthquake Engineering*, Japan, 781-797.

[8] Liu, C.-S. (2001) Cone of Non-Linear Dynamical System and Group Preserving Schemes. *International Journal of Non-Linear Mechanics*, **36**, 1047-1068. http://dx.doi.org/10.1016/S0020-7462(00)00069-X

[9] Liu, C.-S. and Jhao, W.S. (2014) The Second Lie-Group $SO_o(n,1)$ Used to Solve Ordinary Differential Equations. *Journal of Mathematic Research*, **6**, 18-37.

[10] Liu, C.-S. (2014) Disclosing the Complexity of Nonlinear Ship Rolling and Duffing Oscillators by a Signum Function. *CMES*: *Computer Modeling in Engineering & Sciences*, **98**, 375-407.

[11] Liu, C.-S. (2015) A Novel Lie-Group Theory and Complexity of Nonlinear Dynamical Systems. *Communications in Nonlinear Science and Numerical Simulation*, **20**, 39-58. http://dx.doi.org/10.1016/j.cnsns.2014.05.004

Natural Science, 2015, 7, 32-41
Published Online January 2015 in SciRes. http://www.scirp.org/journal/ns
http://dx.doi.org/10.4236/ns.2015.71004

Adaptation of Resilience against Disaster—Case Study of 2000 Tokai Flood and 2011 Flood in Shonai River, Japan

Marie Thomas, Makiko Obana, Tetsuro Tsujimoto

Department of Civil Engineering, Nagoya University, Nagoya, Japan
Email: ttsujimoto@genv.nagoya-u.ac.jp

Received 11 December 2014; accepted 21 December 2014; published 14 January 2015

Abstract

In this study, the application of the resilience concept of the flood event depending on progress of the time is analyzed as the hazard occurrence, the disaster risk, the damage risk, and the evolution of the damages. Flood disaster is defined as the occurrence of an inundation in an exposed area. The human exposure (loss of life, injury, ···), structural (buildings, roads, ···) and functional (economic, political, ··· functions of an area) economic exposure cause high risk of damage if the area in which the hazard occurs is at low resilience. Furthermore the damage will increase without adequate response against disaster. The flood disaster risk is decreased by flood control measures, reducing structural and functional exposure. Non-structural measures, such as appropriate prior-evacuation, decrease the human exposure to flood disaster. This study reviews the events of 2000 and 2011 floods in the Shonai River basin in Japan to help assess resilience to flood disaster. These two events had the same type of hazards in intensity and location, allowing the study in terms of adaptation to flood disaster in the river basin to focus on the structural and nonstructural effort to increase resilience of the disaster depending on progress of the time.

Keywords

Heavy Rainfall, Flood, Flood Mitigation, Evacuation, Hazard, Resilience, Risk Management

1. Introduction

Flood disaster is one of the major natural disasters around the world in terms of frequency, victims and economic losses [1]. Flood management and mitigation consisting in the establishment of structural measures in flood-prone areas such as levees setup to protect directly exposed areas, dams and retarding pond upstream to reduce

How to cite this paper: Thomas, M., Obana, M. and Tsujimoto, T. (2015) Adaptation of Resilience against Disaster—Case Study of 2000 Tokai Flood and 2011 Flood in Shonai River, Japan. *Natural Science*, **7**, 32-41.
http://dx.doi.org/10.4236/ns.2015.71004

the peak discharge of floods and so on, even if sometimes criticized [2], not only are still up to date the only way of reducing the geographical impact of a hazard, but also are part of holistic the concept called "the modes of coping with natural hazard" of Burton *et al*. [3]. It is however true that structural measures have unpredicted impacts as the effect of a disaster in case of levee breach [4] or reduction of risk awareness; therefore, they should be complemented by non-structural measures to face potential levee breach and maintain risk knowledge.

The purpose of this research is to link conceptual research done during the past 30 years in social sciences on disaster management and applied previous hydrology with river hydraulic engineering. Because in social sciences there is no consensus on concepts meaning [5] [6], it is necessary to define clearly what word will be used with what meaning.

Japan is known for its exposition to numerous hazards, and the high exposure of its population in megacities in deltaic floodplains. Continuous flood levees and flood control dams are characteristic and reasonably significant infrastructures for flood protection in Japan since the eighteenth century. But the rapid urban changes during the high economic growth era (1950-1980) had an influence on the lifestyle and the risk culture of urbanites changing the acceptation of risk and increasing people's vulnerability to hazard.

From the above recognition of the topics, firstly in this paper we review the characteristics of floods in Japan simply; then we will present the definitions of several concepts used in this study to assess the adaptation to flood risk by increasing structural and non-structural measure to fight flood risk.

Then two floods in the Shonai River basin near Nagoya metropolis which affected unban activity are taken as examples to be investigated, based on the above scenario described by the fundamental concepts. Two floods occurred in the Shonai River basin in 2000 and 2011, and we analyze how the adaptation of resilience against flood disaster was achieved.

2. Flood Disasters in Japan

Japan stretches from the south west to the north east, and mountains higher than 2000 m form a backbone which brings short and steep rivers with rapid runoff. It is located at the west of the Pacific Ocean (see **Figure 1**), and is exposed to typhoons and characteristic fronts which bring heavy torrential downpours in summer and autumn. Such conditions cause serious floods. Population and human activities are concentrated in limited alluvial plains, where floods with large discharge threat them. Alluvial plains are protected against river floods by continuous flood levees additionally with flood control by dams reducing peak discharge. Such infrastructures against floods have been improved with gradual river improvement projects based on the master plan and concentrated efforts especially after flood disasters. Actually we have experienced a lot of flood disasters in alluvial plains where inundations due to levee breaches have caused serious damage in human life and activities. We have also experienced disasters in mountain area where debris flows have caused serious losses of life and properties, but this category of flood is not dealt with in this paper.

The safety level of the river improvement is measured by the return period of flood discharge which is related to the accumulated rainfall of the river basin for one event. In Japan, the duration of a flood (rainfall) is 1 - 3 days. The discharge is related to the return period, and the safety level is measured whether corresponding discharge can be safely conveyed within the flood levees considering the free board. In other words, if the flood stage exceeds the "high water level" (HWL, the elevation of the top of the levee minus the free board), the chances of a flood disaster (inundation due to levee breach) is increasing (hypothesis in flood mitigation planning). The return period of the maximum discharge which can be safely conveyed within the levees indicates the

Figure 1. Map of Japan.

safety level of the river or the floodplain of that river. At present in Japan, each river has the master plan which must be completed within 30 years, though we have higher safety level as project.

Japan has a disadvantage against floods because the population and human activities are concentrated in low-land floodplain, therefore sometimes the floodplain protected by levees is lower than the river bed. Furthermore under poor river improvement conditions, the flood risk is extremely higher against the possible hazard which must be measured by the flood discharge or the accumulated rainfall in a river basin for an event. Poorly prepared river means not only one with insufficient levees but also with insufficient cross section (higher bed level and river with higher hydraulic resistance). Previously, inundations were frequently occurring due to poor achievement of levee construction and higher bed level (poorly dredged), but inhabitants were ready to protect their daily lives by several strategies recognized as non-structural measures: living in higher lands, small-scale ring levees surrounding homes or villages and evacuation on flooding.

Recently, river improvement has been advanced though it has not reached the master plan level. Because of accumulation of flood protection infrastructures, the decreasing trend of the number of casualties has been achieved from a few thousand to a few dozen during this half century. However, the recent problems exist on the casualties due to insufficient strategy of evacuation once inundation is occurred. The river improvement is still undergoing (under the master plan level), non-structural measures are needed to reduce the risk of disaster once the hazard exceeds the mitigation measures (inundation). In order to reduce the risk of damage by flood, human activities would avoid those in the lower area. However, on the urban development process, population and human activities have expanded to the extremely risky areas because of the lower price of land. Furthermore, non-structural measures have not very effective (low evacuation rates) because inhabitants in flood plains had less experience of inundation and their life style has been extremely changed. We have to devise new system to operate such non-structural measures even in modern society.

3. Concepts Employed in Discussion of Flood Disaster

Though some concepts necessary in discussion of flood mitigation were already employed in the previous chapter, they are carefully defined in the following (see **Figure 2**) in order to make the further discussions clear. The concepts to be discussed here are “hazard”, “disaster risk”, “exposure”, “damage risk”, and “resilience” and/or “vulnerability” for the discussion on flood disaster in a floodplain with particular reference to examples in Japan.

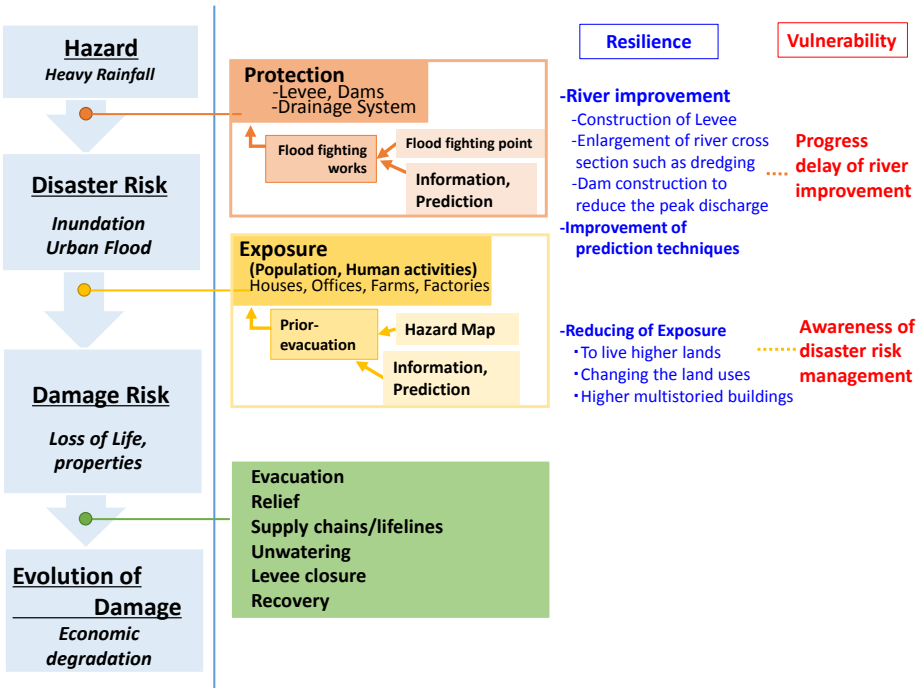


Figure 2. Some concepts employed in discussion of flood disaster.

"*Hazard*" is defined as "the manifestation of a natural or human originated phenomenon with uncertain but supposed extreme intensity and/or very low occurrence potentiality", in our flood case it is the discharge in a river which may cause a "disaster risk". Discharge is brought by accumulated rainfall of a river basin because duration of one-event heavy rainfall and its runoff into a river are in the same order in Japan and within a few days. If the water stage due to the discharge hydrograph exceeds the safety level (HWL = "high water level") of the flood levee, the risk of flood "disaster" becomes extremely high. "*Disaster*" is defined as "the occurrence of a hazard exceeding the mitigation-structural measures—and preparedness-nonstructural measures setup up to this time to prevent disaster, therefore causing structural, functional and human damages to the exposed areas". In our case, the flood disaster can be caused by levee overtopping or levee breach. Simply saying, the risk of flood is determined by the comparison between forthcoming flood discharge and the "protection level" (height of levee and height of river bed level). In order to reduce the risk of flood disaster, river improvement (flood protection) should be done (construction of levee, enlargement of river cross section such as dredging, and dam construction to reduce the peak discharge). Due to the megacity environment of our study field, it has to be remembered that one hazard causing a disaster (levee overtopping) can also occasion events increasing the original disaster (shortage in electricity, inability to reach persons in danger⋯) therefore preparedness (nonstructural measures setup) like prior evacuation, risk culture maintenance, real-time information dispatch are needed.

When inundation (flood disaster) occurs, "*damage*": human, functional and structural losses and disturbances is related to the "*exposure*": human structural and functional activities located in areas where a hazard can be expected. Reducing the exposure must be done by structural (reducing the hazard-prone area) and non-structural measures (maintain activities, functions and preserve human life) to reduce the risk of damage. Prior-evacuation is a strategy to reduce temporarily the human exposure in addition to land use change and building surrounding dikes. In modern urbanized area where life style has been changed, we have to construct a system to assist a strategic prior-evacuation with proper information and alert and to support the shelters.

Furthermore, we have to consider how to reduce the evolution of damage, and it requires rapid recover and restoration. For the time being, the areas of human activities in floodplains particularly in urbanized area are *vulnerable* (inability to cope with a disaster) against flood disaster and the subsequent damages. In order to overcome such vulnerability or to increase *resilience* (ability to prepare for, cope with, and recover from a disaster) against flood disaster and the subsequent damages, we have to prepare firstly structural measures (flood management infrastructures) to reduce the structural and functional exposure (inundation), then the nonstructural measures to reduce the human exposure against damages by inundation, and furthermore the disaster management system has to support rapid recover and restoration to reduce the expansion of damages.

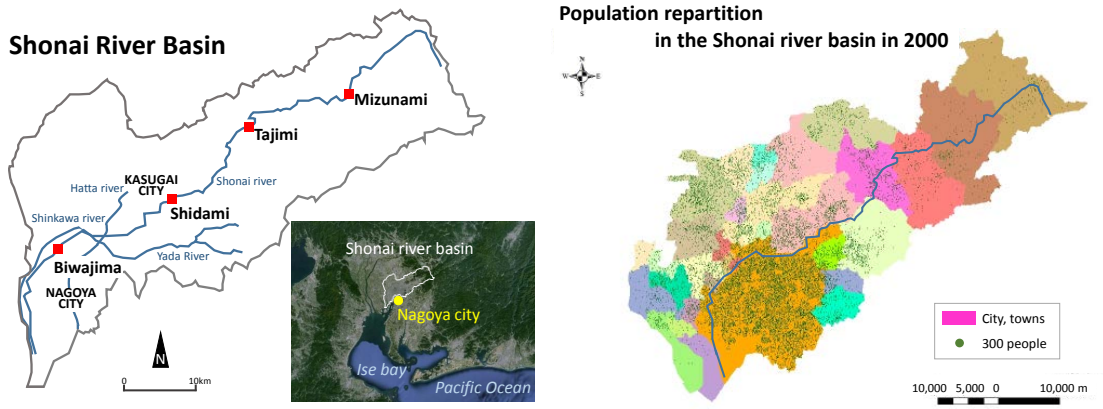
4. Tokai Heavy Rainfall Disaster in 2000 and Shonai River Flood in 2011

4.1. Shonai River Basin and Nagoya City

The Shonai River basin is located in the center of Japan (see **Figure 1** and **Figure 3(a)**). The total length of Shonai River is 96 km and the area of its river basin is 1010 km^2. There are 17 municipalities in this river basin. The lower reach (0 - 35 km) flows in an alluvial plain, and Nagoya metropolitan is located in this area where the main central function such as industries, economies and transportation system and the population are concentrated. The population in this river basin is about 4.3 million, and more than 3.9 million is concentrated in the lower reach including in Nagoya and Kasugai cities (see **Figure 3(b)**). The site of Shidami (34 km from the river mouth) and Biwajima (lower reach, 14.5 km from the river mouth) is selected as the reference site of river master plan.

4.2. Heavy Rainfall Events in 2000 (Tokai Heavy Rainfall Disaster)

In September 2000, we experienced a severe flood disaster in Nagoya metropolitan area. The heavy torrential downpour recorded 100 mm within an hour at Nagoya meteorological observatory and the accumulated rainfall of the Shonai River basin exceeded 500 mm, then the peak discharge of the Shonai River exceeded the capacity at that time. The safety level of the river reached at most the level with return period of 30 years in those days. Hence, the water stage along the river exceeded HWL, several levee breaches occurred and the most serious one was a levee breach in "Shinkawa River," a tributary (diversion) of the Shonai River in northern Nagoya. In addition, because of insufficient urban drainage, a wide area was seriously inundated. Insufficient urban drainage



(a) (b)

Figure 3. (a) Shonai River basin; (b) Population distribution map in the Shonai River basin in 2000.

was caused not only by limited capacity of pumping but also by stopping pumps mechanically by inundation and institutionally because of higher river stage at the outlet of drainage.

The casualties were fortunately few (less than 10 casualties in the Shonai River basin), but the economic loss due to this disaster reached 70 billion yen (US $7 billion) and around 90% was from general economic activities (not related to public works loss). There were a lot of inhabitants and small factories in lowland area near the river, which were seriously affected.

In addition, various damages were focused on related to daily urban activities not only in seriously inundated area but also even in not-inundated area. In particular, inundated water invaded into the subway facilities (stations and railways) and metropolitan functions were paralyzed. Needless to say, the transportation failure is serious in daily life and economic activities particularly in urban area and the flood caused the Japanese bullet train to stop, causing communication problems between the two major cities of Japan, increasing the scale of the disaster's consequence at national level. Not only transportation, sewage system were damaged. In general, serious inundation brings paralysis of life lines such as water supply, energy (electricity and city gas) and telecom. Because of inundation and subsequent damage of lifelines, many people required refuges and could not return to their houses and offices for several days. Furthermore the garbage treatment problem after inundation was serious. Nagoya metropolis area had a serious issue to look for garbage dump area even before the disaster, and in addition they had to treat large amount of garbage due to inundation disaster. In particular, electric products such as refrigerators and computers, cars, and furniture such as tatami mats and so on could not be reused after inundation and they were bulky and heavy to deal with, slowing the recovery period after the disaster.

Some areas were inundated because of insufficient urban drainage system, but some of them were caused by insufficient river improvement because the pumping from the sewage system to the river was not allowed when the water stage exceeds the limit. During Tokai flood, however, some levee breaches occurred and some inhabitants doubted illegal release from the pumps after the water stage exceeded HWL.

4.3. Resilience Adaptation after Tokai Flood

The Tokai flood in 2000, helped to recognize the vulnerability of this, and river improvements have been undertook using the framework of "special recovering project against serious disaster" and the successive planned river improvement according to the master plan. The former is the framework to repair and improve infrastructure to avoid the disaster due to the hazard of the same pattern then the previous (Tokai flood) disaster which has been done during the 5 following years (2001-2006). This project was realized on the Shonai River and its tributary the Shinkawa River, and for the Tenpaku River (another river system in Nagoya metropolis). This project consisted of the expansion of the cross section of the river to keep the capacity of flood passing by river-bed dredging and the strengthening of the levee (seepage protection against levee breach). The expected return period for structural measures was increased and the Shonai River has now the protection ability against the level of Tokai flood. After Tokai flood after finishing the special project, the master plan project understanding successive river improvements to aim the level of 50 years return period was undertook and should be completed

within 30 years.

In addition to the structural measures (flood protection), the non-structural measures have been discussed. The effective way is reducing the exposure to the risk of inundation. The targeted project was to improve prior-evacuation, and firstly the proper information system has been improved at all levels of the disaster management system. First, the river manager made an "inundation map for design flood" (prediction for 200 years return period flood intensity, which are more severe than the master plan level) and provide it to the city disaster managers in charge of evacuation, and then the city made a "hazard map" distributed it to the inhabitants to show individuals inundation to expect and evacuation information at the community level. Then the inhabitants can acknowledge the hazardous area for flood risk, and can recognize the risk of inundation with successive damages for each forthcoming event and take necessary pre-evacuation measures (information provided on pre-evacuation measures by the city with the hazard map). Some of the inhabitants can evacuate based on the information on real time weather and river (available on the Shonai River Bureau website) and with the "hazard map". And the suggestions of the timing of prior-evacuation will help inhabitants who may not start actions without incentive.

The evacuation suggestion (advice) or order (obligation to evacuate) is issued from the community head (mayor), for whom it is generally difficult to judge the proper timing of evacuation with insufficient information on rivers and knowledge on inundation processes. They often hesitate to issue such orders, therefore, in order to avoid the situation of delayed evacuation suggestion, the process was systematized as follows: The river manager provides the information not only for the real time data on the river (water stage) but also their forecast. Then, the standardized process to judge the timing of evacuation with necessary procedures in the community (municipality) was discussed and the framework of the risk management system (information of river data from the river manager and suggested tasks of community) was established as shown in **Figure 4**. This figure shows timeline for the river manager which may help community to issue evacuation suggestion or commend.

Next, the information from the community to the inhabitants should be improved. In this area, a system was constructed through mobile-phone mail system to send automatically necessary information for emergency. Then, a necessary issue is the promotion for the inhabitants to carry out prior-evacuation (education and training). The system to support evacuees on the way of evacuation and in the refuges should be improved too. At the present time, the community cannot deal with all the evacuees properly.

4.4. Shonai River Flood in 2011 Compared with Tokai Flood in 2000

In September 2011, 11 years later after the Tokai flood, heavy rainfall stroke the Shonai River basin again and the peak discharge of flood along the Shonai River was almost same to each other (design flood level of the master plan). **Figure 5** shows the weather maps for those events for comparison, and it is recognized that the weather conditions are very similar to each other where the typhoon existed on the south east of Kyushu island and the front expanded from the vicinity of the typhoon to the north east. Rainfalls within an hour were near 100 mm and accumulated rainfalls reached around 500 mm, though highest intensity of the hazard was slightly different for the two events. Heavy rainfall concentrated near Nagoya city in 2000 while it concentrated near Shidami in the middle reach of the Shonai River as shown in **Figure 6**.

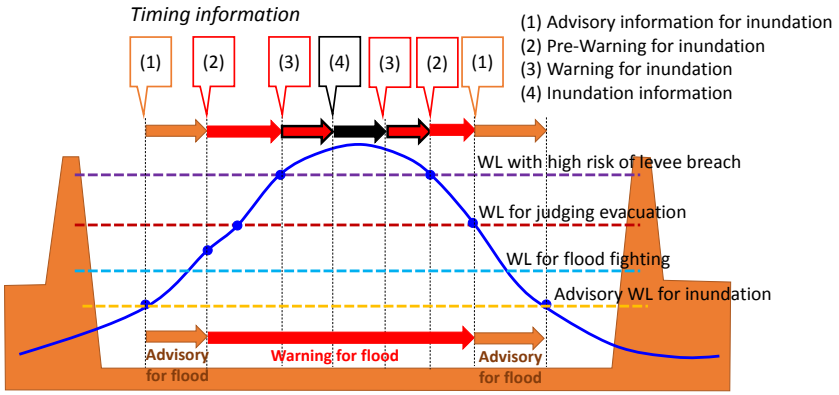


Figure 4. Timeline of timing information for river manager in stage hydrograph.

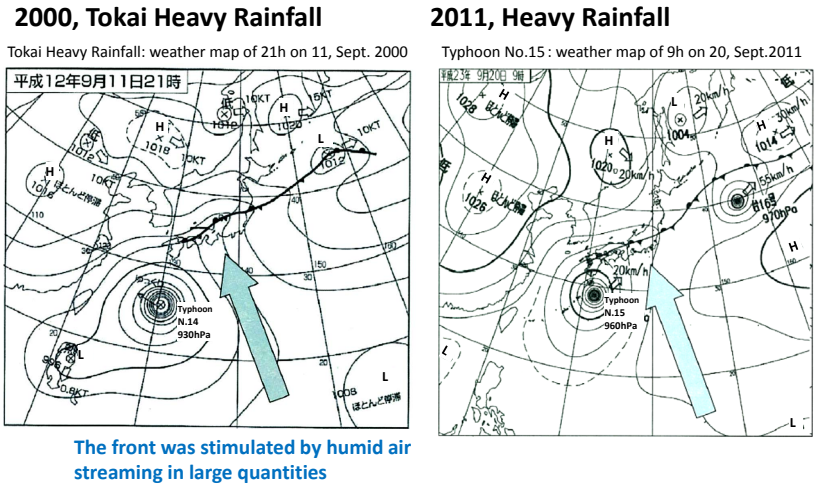


Figure 5. Weather maps for 2000 and 2011 of heavy rainfall.

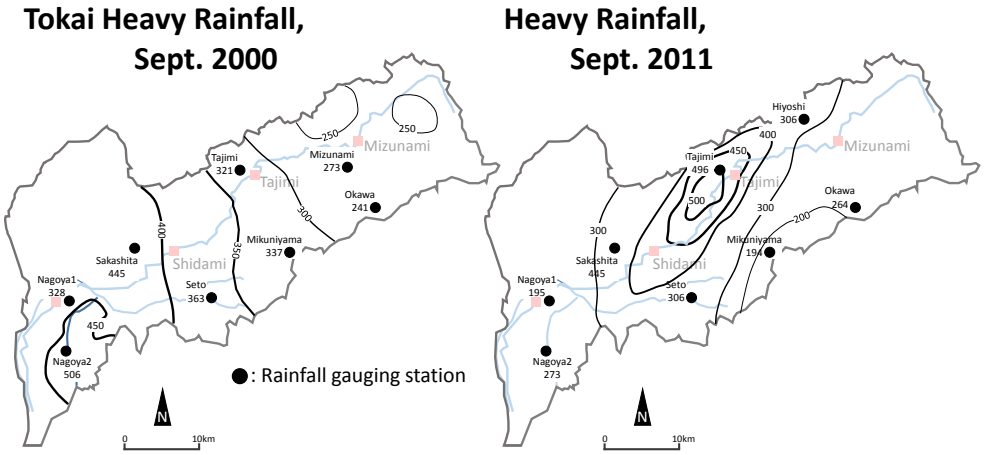


Figure 6. Spatial distribution of accumulated rainfalls for 2000 and 2011 rainfalls.

The observed peak discharges were almost the same on the both events and it was around the designed flood discharge (3500 m^3/s). As seen in **Figure 7** (comparison of the stage hydrographs between 2000 and 2011 floods), the peak stage at Shidami (middle reach) for 2011 flood was higher than the one for 2000 flood, but the peak stage at Biwajima was lower. In addition, the duration time of flood in 2011 was shorter than that in 2000.

We would say that the flood (heavy rainfall) in 2011 was the same level of "hazard" with that in 2000 though there were slight differences. In particular the spatial difference brought a great difference in exposure. The lower reach of the Shonai River where the 2000 flood occurred is a highly metropolitan exposed flood plain area with much more population and economic activities than the middle reach where the 2011 flood occurred.

4.5. Assessment of Resilience Adaptation after Tokai Flood

Because of the advancement of river improvement by the special project after 2000 flood and the successive river improvement according to the master plan, the impact of the hazard (inundation) were quite different from each other. The "disaster risk" is suggested by comparing the flood stage along the river with HWL. **Figure 8** shows the longitudinal profiles of HWL with flood marks (peak water stage) for 2000 and 2011 floods. When the floods in 2000 and 2011 are compared with each other in this figure, in the lower reach the water stage in 2011 was obviously lower and it exceeded HWL nowhere though being almost the same discharge flow there. This was caused mainly by the result of the special project after 2000 flood. The hazard being contained by structural measures, the lower reach didn't suffer any disaster, while in the middle reach near Shidami, the flood marks were higher than those in 2000 and exceeded HWL in some places. In Shidami district, the over flow

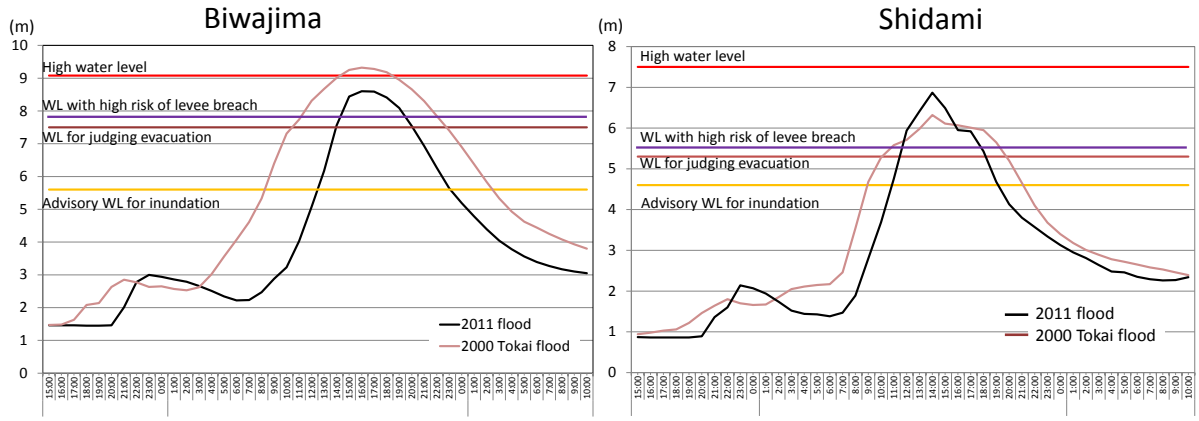


Figure 7. The stage hydrographs at Biwajima and Shidami for 2000 and 2011 floods (time axes set for the peaks to be the same).

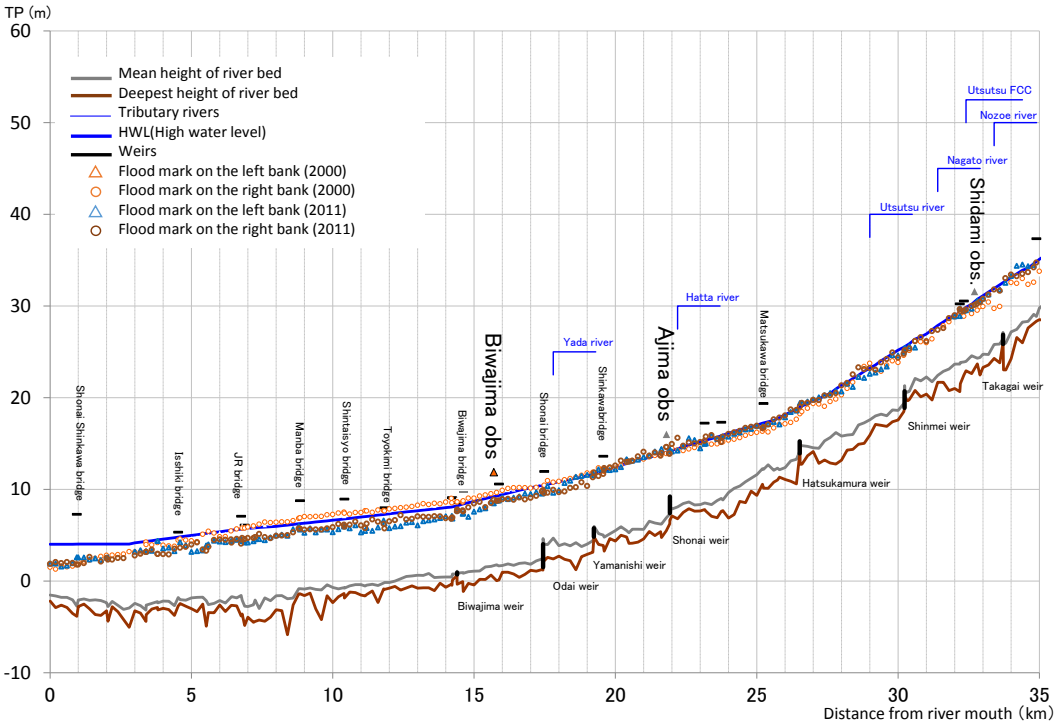


Figure 8. The longitudinal profiles of HWL with flood marks for 2000 and 2011 floods.

occurred to cause successive inundation (and high risk of levee breach to be accompanied with flood fighting activities), and higher stage of the river disturbed the drainage activity from the areas along the river (Tajimi and Kasugai cities).

The special project for 2000 flood was concentrated only in the lower reach, and even the successive river improvement according to the master plan has been progressed from the downstream. The project of the expansion of the cross section is an improvement that has to be realized progressing from the downstream from the viewpoint of hydraulic influence.

In addition, the concentration of rainfall in the middle reach in 2011 brought the rapid runoff with sharp hydrograph in the middle reach. Furthermore, it might be guessed that bed morphological change was promoted by the 2000 flood's propagation and the vegetation invasion increased the hydraulic resistance leading to higher water stage consequences. It is reasonable to expect river improvement to progress from the downstream but we have to take care of the maintenance of river bed and vegetation even in the upstream reach.

In 2011 flood, higher water stage in the middle reach brought the higher downstream condition for the flood in the tributaries. Particularly along the Hatta River, a tributary of the Shonai River, on which a levee overtopping occurred, caused serious inundation with poor drainage, and the inhabitants along this river were seriously threatened by extremely high risk of levee breach (see **Figure 9**).

Because fortunately no obvious levee breaches occurred, the disaster was not serious. However from the viewpoint of risk management, comparatively large area of the river basin area was threatened by flood risk because the water stage exceeded HWL at several points along the middle reach of the Shonai River and its tributaries. After 2000 flood, the citizen was presumed to know the hazard map. The information of the water stage of the river was supplied to the municipalities from the river manager with its forecast, and the municipalities behaved almost exactly as expected to the presumed scenario (the standard was already shown by **Figure 4**) as shown in **Figure 10**.

In 2011, Nagoya mayor issued "evacuation recommendation" smoothly according to the aforementioned framework, and the number of targeted inhabitants was more than a million (1,095,204). The newspaper on the next day commented comparatively favorably. However, actually only 4394 peoples (around 0.4%) followed the suggestions to evacuate. In 2000, the prior evacuation was not suggested though 32,155 peoples finally evacuated. While, on the heavy rainfall in 2008, based on the standard framework, Nagoya city issued the evacuation recommendation to 360,000 peoples. Then, already the standard format for issuing recommendation was functioning but the number of peoples who evacuated was only 375 (0.1%). Because of the development of area mailing service for alert through mobile phones have been developed and became popular, the ratio of the evacuation has increased gradually. On 2011 flood, prior-evacuation was recommended at appropriate timing based on smooth process for issuing evacuation recommendation, but the following remains to be discussed: 1) Are the inhabitants targeted for evacuation reasonable, and if not how can we explain lack of interest in evacuation in order to fix the problem? 2) Are the communities ready to accept and support such number of evacuee?

According to the present manual, the target of evacuation recommendation is all inhabitants in the administrative districts parts of which may be inundated, and thus all of them are not necessary to evacuate. The inhabitants

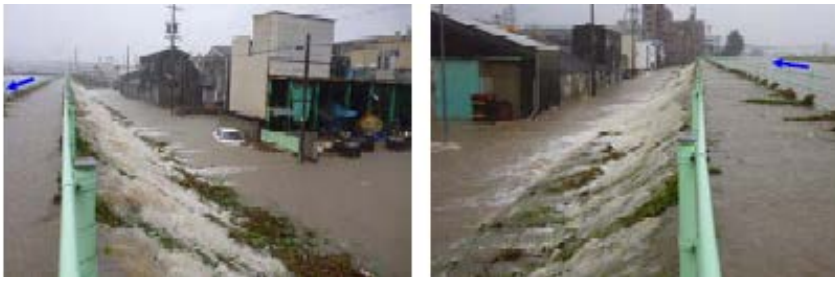


Figure 9. Overflowing from top of the levee in Hatta River.

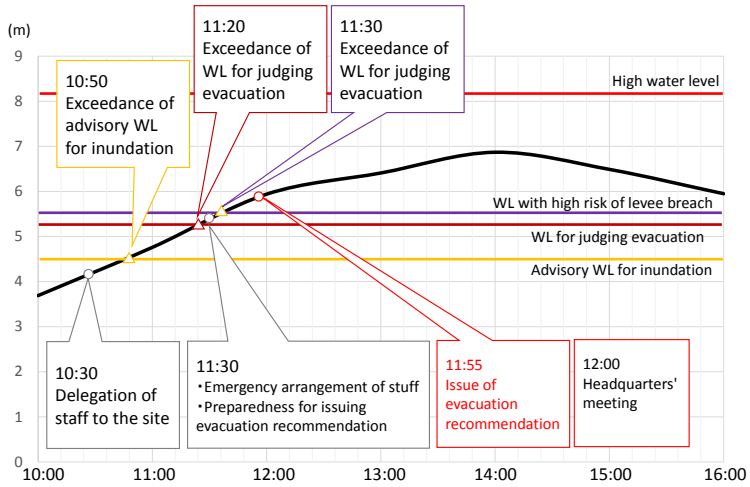


Figure 10. The presumed scenario of timing information and action for river manager.

can recognize whether it is necessary to evacuate or not if they consult the hazard map, but only a few people will evacuate if they know not all are required to evacuate. Moreover, hazard map is based on an inundation map where inundated areas for some scenarios with different breaching points are overlaid, and it may include the area to be inundated later after the beginning of inundation (breach of the levee). In other words, nobody recognize the exact numbers of inhabitants necessary to evacuate.

On the other hand, the support for evacuees is not well prepared. The refuges or the shelters have less capacity compared with numbers of evacuees, while nobody has checked the exact number of evacuees.

5. Conclusions

In this study, we focused on the floods due to heavy rainfalls in Shonai River basin in 2000 and 2011, which were caused by hazards that could almost match the actual master plan projected structural measures improvements. Nagoya metropolis, the 4th largest city, exists on the Shonai River basin, and the floods threatened the urban functions both at local and national scale.

The research was conducted in the framework constituted by the following concepts: “hazard”, “disaster risk”, “exposure”, “damage risk”, “mitigation” as structural measures against flood, “preparedness” as non-structural measures mitigate the damages by flood, and “resilience” and/or “vulnerability” from more integrated viewpoints including prevention, coping, repair and restoration abilities. We clearly defined those concepts for floods in Japan. Based on these concepts, the two floods were investigated with particular reference of resilience adaptation after 2000 flood, and finally the remaining problems were pointed out.

Acknowledgements

The authors would show sincere gratitude to the Regional Bureau of MLIT (ministry of land, infrastructure, transportation and tourism), Shonai River Office and Nagoya Municipal Office for their efforts to collect and publicize data for floods in 2000 and 2011 and valuable discussion with them.

References

[1] Jonkman, S.N. (2005) Global Perspectives on Loss of Human Life Caused by Floods. *Natural Hazards*, **34**, 151-175. http://dx.doi.org/10.1007/s11069-004-8891-3

[2] Dauphiné, A. and Provitolo, D. (2007) La résilience: Un concept pour la gestion des risques. *Annales de Géographie*, **654**, 115-125. http://dx.doi.org/10.3917/ag.654.0115

[3] Burton, I., Kates, R.W. and White, G.F. (1978) The Environment as Hazard. Oxford University Press, New York.

[4] White, G.F., Calef, W.C., Hudson, J.W., Mayer, H.M., Schaeffer, J.R. and Volk, D.J. (1958) Changes in Urban Occupance of Flood Plains in the United States. Department of Geography Research Papers, 57.

[5] Gallopin, C.G. (2006) Linkages between Vulnerability, Resilience, and Adaptive Capacity. *Global Environmental Change*, **16**, 293-303. http://dx.doi.org/10.1016/j.gloenvcha.2006.02.004

[6] Thomas, M. and Tsujimoto, T. (2012) Vulnerability to Flood Risk in Japanese Urban Areas: Crisis Management and Emergency Response in Nagoya, for an Efficient Evacuation Management. *WIT Transaction on Ecology and the Environment, Flood Recovery and Response IV*, **184**, 61-71.

Natural Science, 2015, 7, 42-54
Published Online January 2015 in SciRes. http://www.scirp.org/journal/ns
http://dx.doi.org/10.4236/ns.2015.71005

The Classical Binary and Triplet Distribution Functions for Dilute Relativistic Plasma

N. A. Hussein[1], D. A. Eisa[2], E. G. Sayed[1]

[1]Mathematics Department, Faculty of Science, Assiut University, Assiut, Egypt
[2]Mathematics Department, Faculty of Science, Assiut University, New Valley, Egypt
Email: aragamal@yahoo.com

Received 1 January 2015; accepted 19 January 2015; published 22 January 2015

Abstract

The aim of this paper is to calculate the binary and triplet distribution functions for dilute relativistic plasma in terms of the thermal parameter μ where $\mu = \frac{mc^2}{KT}$, m is the mass of charge; c is the speed of light; K is the Boltzmann's constant; and T is the absolute temperature. Our calculations are based on the relativistic Bogoliubov-Born-Green-Kirkwood-Yvon (BBGKY) hierarchy. We obtain classical binary and triplet distribution functions for one- and two-component plasmas. The excess free energy and pressure are represented in the forms of a convergent series expansions in terms of the thermal parameter μ.

Keywords

Relativistic Plasma, Binary and Triplet Distribution Functions, BBGKY Hierarchy

1. Introduction

Studying the properties of plasma has received a great interest in both astrophysical and laboratory plasma application. The relativistic binary distribution function is one of the most important functions of statistical mechanics. The importance of the distribution function in statistical mechanics is due to the fact that all the thermodynamic quantities, such as the pressure, the internal energy and the free energies, can be calculated from it.

Relativistic statistical mechanics has a long story, but we may notice that, whereas the theory of relativistic ideal gases has received deep and detailed developments, little has been achieved in order to account for mutual interactions between particles [1]. In statistical physics, the BBGKY hierarchy (Bogoliubov-Born-Green-Kirkwood-Yvon hierarchy, sometimes called Bogoliubov hierarchy) is a set of equations describing the dynam-

How to cite this paper: Hussein, N.A., Eisa, D.A. and Sayed, E.G. (2015) The Classical Binary and Triplet Distribution Functions for Dilute Relativistic Plasma. *Natural Science*, **7**, 42-54. http://dx.doi.org/10.4236/ns.2015.71005

ics of a system of a large number of interacting particles. The equation for an *s*-particle distribution function (probability density function) in the BBGKY hierarchy includes the $(s+1)$-particle distribution function thus forming a coupled chain of equations [2]. Their history of 70 years has brought enormous progress in the investigation of the transition from the microscopic to the macroscopic world, and they are still an attractive starting point for new developments. In particular, great advance has come by clever methods of truncation, approximation and scaling limits of the hierarchy, providing in various cases a justification of the kinetic equations describing particle systems on mesoscopic (intermediate) scales [3]. Many authors studied the BBGKY hierarchy [4] and [5]; Hussein and Eisa [6] [7] calculated the binary and triplet distribution function for one- and two-component plasmas for quantum and classical non-relativistic plasma. In this paper, we will calculate these distributions in the relativistic case in terms of plasma thermal parameter. Arendt and Eilek [8] showed that the pair-plasma distribution functions could be described by a thermal parameter $\mu = \frac{mc^2}{KT}$ that was moderately relativistic. The thermal parameter value plays a significant role for the stability of our system. Barcons and Lapiedra (1984) [9] gave explicit expressions for the thermodynamic functions of a high-temperature electron-positron plasma and gave expression for distribution functions for a classical dilute arbitrarily hot plasma in equilibrium which we compared our results with it.

Special relativity, however, does not permit velocities greater than the speed of light c and is thus incompatible with a Maxwellian distribution that predicts a non-zero probability for every velocity [10]. It is possible to obtain the relativistic Maxwellian distribution (Jüttner distribution) of a moving gas when the so-called Planck-Einstein case is analyzed. It should be reiterated that the Planck-Einstein theory suffered a modification, principally in the transformation law of energy (the Planck-Einstein case) [11]. Treumann *et al*. [12] have also shown that the isotropic thermal equilibrium distribution function in relativistic plasma analytically has the form of a modified Jüttner distribution.

A little is known about relativistic distribution functions involving more than two particles, and in particular about the three-particle (or triplet) distribution function. This is of course due to the greater mathematical complexity of higher order correlation functions, and to a lack of a direct link with experiment. Although one can consider experimental determination of triplet distribution function from triple elastic scattering similar to that of the binary distribution function, but to our knowledge, the measurement of three-body correlation function requires the knowledge of the positions of three particles at the same time which is technically very demanding to obtain in 3D samples [13]; this means that the experiment to measure the triplet distribution function directly requires high precision so that the multiple scattering can be differentiated from single scattering. Essentially, one follows the phase-space trajectory of the system as it evolves in time, and has thus the same amount of information as one obtains in a simulation. Similarly the particle positions have measured in dusty plasmas [14]. Lapiedra *et al*. [15] have undertaken an application of predictive methods to relativistic statistical mechanics. Coulomb forces between point charges are purely repulsive and charges approach very close to each other. Coulomb systems, such as plasma or electrolytes, are made of charged particles interacting through Coulomb's law. The simplest model of a Coulomb system is the one-component plasma (OCP), also called jellium: an assembly of identical point charges, embedded in a neutralizing uniform background of the opposite sign. Here we consider the classical (*i.e*. non-quantum) equilibrium statistical mechanics of the OCP. It is rather straightforward to calculate higher-order correlation functions from the measured configurations. Moreover we study the model of two-component plasma (TCP) *i.e*. neutral system of point like particles of positive and negative charges such as electrons and ions. For the numerical calculation we restrict ourselves to the case of (TCP) which anti-symmetric with respect to the charges $e_e = -e_i = -e$ and therefore symmetrical with respect to the densities $\rho_e = \rho_i = \rho$. To simplify the numeric investigations, we simulated so far only mass symmetrical electron-ion plasma with $m_e = m_i = m$.

Homogeneous plasma is characterized by two parameters: the density of particles ρ and the temperature T; if there are several kinds of particles, it is also necessary to state their concentrations. There are different energies associated with the plasma, namely:

1) Energy of the rest mass per particle mc^2;

2) Kinetic energy per particle of order kT;

3) Coulomb energy of order $e^2\rho^{1/3}$ per particle.

The ratios of these energies give us the two main dimensionless parameters of the plasma: the thermal para-

meter of plasma $\mu = \frac{mc^2}{KT}$ and the dilution parameter $\epsilon_d = \frac{e^2\rho^{1/3}}{KT}$. The value of μ is very important. There are four different regimes characterized by it: plasma with $\mu \le 1$ is relativistic, $\mu \ge 1$ weakly relativistic plasma, $\mu \gg 1$ low plasma temperature and $\mu \ll 1$ ultra-relativistic high temperatures plasma. In our study we note that the system is safely classical if where $\lambda = \frac{\hbar}{P} \ll \rho^{1/3}$, $\hbar = \frac{h}{2\pi}$, h is the Planck constant, and P is the typical linear momentum of particles. A relativistic plasma with a thermal distribution function has temperatures greater than around 260 keV, or 3.0 GK (5.5 billion degrees Fahrenheit), where approximately 10% of the electrons have $\gamma > 2$ where $\gamma = 1/\sqrt{1-(v/c)^2}$ is the Lorentz factor. It is occurs in many environments in astrophysics, including gamma-ray bursts, AGN jets, and pulsar winds. In physics, a particle is called ultrarelativistic when its speed is very close to the speed of light c. Or, similarly, in the limit where the Lorentz factor is very large $\gamma \gg 1$ and $\mu = 10^{-2}$. Our study is only valid for dilute plasmas $(\epsilon_d < 1)$. The plasma temperatures and velocities also plays a significant role for the stability of our system **Figure 1** show qualitative sketch of the four different regimes characterized by the thermal parameter μ and Lorentz factor γ. Our model lies in the region of high speeds, $\gamma > 2$ and $\mu < 1$.

In the astrophysical environment, the fraction of ionized particles varies widely from nearly no ionization in cold regions to fully ionized in regions of high temperature. This leads to a wide range of parameters where astrophysical plasmas can exist. While the astrophysical environment is frequently dominated by the presence of the plasma, this plasma is often strongly influenced by and coupled to the presence of embedded particulates (*i.e.*, dust). These dust grains which range in size from a few nanometers to micron-sized objects can become either positively or negatively charged due to interactions with the background plasma environment and ionizing radiation sources in the astrophysical environment [16]. Understanding the processes that govern these plasma particle interactions is critical to the study of astrophysics. The agglomeration and growth of larger particles from single atoms and dust grains leads to the eventual formation of objects so large that gravity becomes the dominant force controlling their subsequent evolution [17].

Another important point is that of the walls. We shall study plasma which is homogeneous and isotropic but the plasma must be confined or it will expand. We may assume that the plasma is confined by some kind of walls which prevent the escape of particles, but that the container is so large that the effects of the walls are negligible. Our calculations are based on the phase-space distribution function; this is defined as the number of particles per unit volume of space per unit volume of velocity space: At time *t*, number of particles in elementary volume of space, with velocities in range $u \to u + \mathrm{d}u = f(x,u,t)\mathrm{d}^3x\mathrm{d}^3u$.

2. The Basic Equations and Hierarchy

The statistical state of a macroscopic system of *N* particles is in a complete—though in an intractably complex—

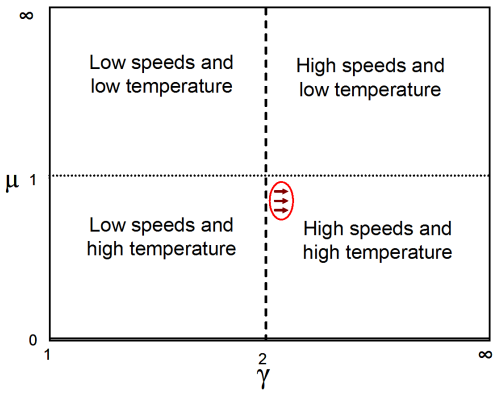


Figure 1. Qualitative sketch of the four different regimes characterized by the thermal parameter μ and Lorentz factor γ: our model lies in the region of high speeds, $\gamma > 2$ and $\mu \lesssim 1$.

way described by the distribution function F^N in $6N$ dimensional phase space, which is spanned by the coordinates and velocities of all individual particles $F^N(t, x_A, u_A)$, $A = 1, 2, \cdots, N$, x_A being the three position of particle A, and $u_A = v_A \Big/ \sqrt{1-(v_A/c)^2}$ the spatial components of its four velocity in a given frame.

The *s*-particle reduced distribution is giving by:

$$F^{(s)}(t, x_A, u_A) = \int F^N \prod_{R=s+1}^{N} \mathrm{d}^3 x_R \mathrm{d}^3 u_R. \tag{1}$$

The relativistic BBGKY hierarchy [15] is given by

$$u_A^\alpha \frac{\partial F^{(s)}}{\partial x_A^\alpha} + \sum_B \zeta_{AB}^\alpha \frac{\partial F^{(s)}}{\partial u_A^\alpha} + \sum_R \int \zeta_{AB}^\alpha \frac{\partial F^{(s+1)}}{\partial u_A^\alpha} \prod_{R=s+1}^{N} \mathrm{d}^3 x_R \mathrm{d}^3 u_R = 0, \quad A \neq B \tag{2}$$

where ζ_{AB} is the acceleration of the charge A in the presence of the charge B.

The Einstein summation convention is only valid for Greek labels. Therefore, for each value of *s* we have *s* equations, since $A = 1, 2, \cdots, s$. As in the non-relativistic case, the determination of the reduced generalized distribution function $F^{(s)}$ can only be made when the hierarchy is cut off somewhere, that is, when for some value s we give $F^{(s+1)}$ as a function of the other $F^{(r)}$ functions with $r < s+1$.

According to what we done in non-relativistic case [7] we set, whatever particles 1, 2, 3 are,

$$F^{(2)}(1,2) = F^{(1)}(1) F^{(1)}(2) \left[1 + G(1,2) + \frac{1}{2} G^2(1,2) + \frac{1}{3!} G^3(1,2) \right] \tag{3}$$

$$F^{(3)}(1,2,3) = F^{(1)}(1) F^{(1)}(2) F^{(1)}(3) \times \left[1 + G(1,2) + G(1,3) + G(2,3) + \frac{1}{2}\left(G(1,2) + G(1,3) + G(2,3)\right)^2 + \frac{1}{3!}\left(G(1,2) + G(1,3) + G(2,3)\right)^3 \right]. \tag{4}$$

Let us consider the case of homogeneous plasma in equilibrium. Then $F^{(1)}$ is the one-particle distribution function of an ideal gas see **Figure 2** which shows one particle relativistic distribution function in the particle velocity interval (0, 0.9*c*) for different values of the thermal parameter μ. That is, in a frame relative to which the system is macroscopically at rest, we must set for $F^{(1)}$ the relativistic Maxwellian distribution [18]

$$F^{(1)}(v_1) = \frac{\mu_1}{4\pi (mc^3) K_2(\mu_1)} \exp\left(\frac{-\mu_1}{\sqrt{1-(v_1/c)^2}} \right) \tag{5}$$

where $K_2(\mu)$ denotes the modified Bessel function.

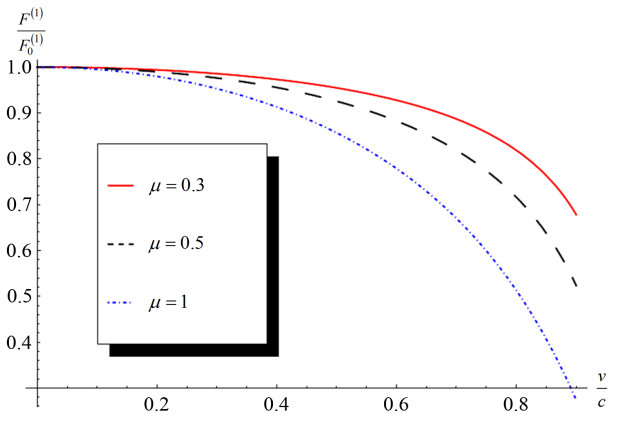


Figure 2. The one particle relativistic distribution function in the particle velocity interval (0, 0.9*c*) for different values of the thermal parameter μ.

3. The Binary Relativistic Distribution Function

At first sight, the calculation of the relativistic interaction between two charged particles seems rather involved, because the force on particle 1 at time t would depend on the position and velocity of particle 2 at *a* retarded time. Also, the position and velocity of 1 at that time depends on the position and velocity of 1 at an earlier time, and so on. If one chooses a frame of reference, the acceleration on particle 1 due to the presence of 2 can be calculated from the position and velocity of 2 at that time. This does not mean that actions propagate instantaneously, but rather that there is a precise scheme to take retardation into account automatically, through the equations of the theory. All we shall need in that chapter is the time component of the acceleration of particle 2 due to particle 1 [19].

To calculate the binary relativistic distribution function substituting from Equations (3) and (5) into (2) for $s=1,2$ we obtain

$$\sum_R \int F^{(1)}(R)\zeta_{1R}^{\alpha}\frac{\partial}{\partial u_1^{\alpha}}F^{(1)}(1)G(1,R)\mathrm{d}^3x_R\mathrm{d}^3u_R=0 \tag{6}$$

$$u_A^{\alpha}\frac{\partial G(1,2)}{\partial x_A^{\alpha}}-\mu_A\zeta_{AB}^0+\sum_R\int F^{(1)}(R)\zeta_{AR}^0G(B,R)\mathrm{d}^3x_R\mathrm{d}^3u_R=0. \tag{7}$$

Then, by solving the integro-differential Equation (7) and by using the Fourier transform of $G(1,2)=G(\boldsymbol{x},\boldsymbol{v}_1,\boldsymbol{v}_2)$ *i.e.*, $G(\boldsymbol{k},\boldsymbol{v}_1,\boldsymbol{v}_2)=(2\pi)^{-3/2}\int G(\boldsymbol{x},\boldsymbol{v}_1,\boldsymbol{v}_2)\exp(i\boldsymbol{k}\cdot\boldsymbol{x})\mathrm{d}^3x$. We get the relativistic binary generalized distribution function $F^{(2)}(1,2)$ at $t_1=t_2$ in the following form

$$\begin{aligned}F^{(2)}(1,2)=F^{(1)}(1)F^{(1)}(2)\sum_{i,j=1,i\neq j}^{2}\Bigg\{&1-\frac{e_ie_j}{m_ir_{ij}}e^{-\kappa r_{ij}}\mu_i+\frac{e_ie_j}{4m_j\pi^2}l_{ij}\,\mu_j+\frac{e_i^2e_j^2}{2m_i^2r_{ij}^2}e^{-2\kappa r_{ij}}\mu_i^2+\frac{e_i^2e_j^2\mu_j\mu_i}{2m_im_j\pi^2r_{ij}}e^{-\kappa r_{ij}}l_{ij}\\&+\frac{e_i^2e_j^2}{16m_j^2\pi^4}l_{ij}^2\mu_j^2-\frac{e_i^3e_j^3}{6m_i^3r_{ij}^3}e^{-3\kappa r_{ij}}\mu_i^3-\frac{e_i^3e_j^3}{64m_j^3\pi^6}l_{ij}^3\mu_j^3-\frac{3e_i^3e_j^3}{8m_i^2m_j\pi^2r_{ij}^2}e^{-2\kappa r_{ij}}\mu_i^2\mu_jl_{ij}\\&-\frac{3e_i^3e_j^3}{16m_im_j^2\pi^4r_{ij}}e^{-\kappa r_{ij}}\mu_i\mu_j^2l_{ij}^2+\cdots\Bigg\},\end{aligned} \tag{8}$$

where $\boldsymbol{n}=\dfrac{\boldsymbol{k}}{k}$, $\kappa^2=\dfrac{4\pi}{V}\sum_R\dfrac{\mu_R}{m_R}e_R^2$, $\alpha^2=\dfrac{4\pi}{V}\sum_R e_R^2\dfrac{\mu_R}{m_R}B(\mu_R)$,

$$l_{ij}=\int\frac{(\boldsymbol{n}\cdot\boldsymbol{v}_i)(\boldsymbol{n}\cdot\boldsymbol{v}_j)-\boldsymbol{v}_i\cdot\boldsymbol{v}_j}{(i+\boldsymbol{n}\cdot\boldsymbol{v}_i)(i+\boldsymbol{n}\cdot\boldsymbol{v}_j)}\frac{e^{-i\boldsymbol{k}\cdot\boldsymbol{x}}}{k^2+\alpha^2}\mathrm{d}^3k,$$

$$B(\mu_R)=\pi\int F^{(1)}(R)\left[\frac{3-v_R^2}{2v_R}\ln\frac{1+v_R}{1-v_R}-3\right]u_R^2\mathrm{d}u_R.$$

Let us now study the two-particle relativistic distribution for the model of two-component plasma (TCP) *i.e.* neutral system of point like particles of positive and negative charges such as electrons and ions. For the numerical calculation we restrict ourselves to the case of two-component plasma which anti-symmetric with respect to the charges $e_e=-e_i=-e$ and therefore symmetrical with respect to the densities $n_e=n_i=n$.

For two-component plasma we can use the two-particle correlation function $G(1,2)$ which is given by

$$G(1,2)=g_1(r_{12})+g_2(r_{12})\frac{\boldsymbol{v}_1\cdot\boldsymbol{v}_2}{c^2}+g_3(r_{12})\frac{(\boldsymbol{x}_{12}\cdot\boldsymbol{v}_1)(\boldsymbol{x}_{12}\cdot\boldsymbol{v}_2)}{2c^2r_{12}^2} \tag{9}$$

$$g_1(r_{12})=-\frac{e_1e_2}{KTr_{12}}e^{-\kappa r_{12}} \tag{10}$$

where $g_1(r_{12})$ is the Debye-Hiickel solution [20] for two-component plasma, by substituting Equation (9) into Equation (3) we get the binary distribution function for two-component plasma in the following form:

$$F^{(2)}(1,2)=\frac{\mu_1\mu_2}{8\pi^2\left(mc^3\right)^2K_2(\mu_1)K_2(\mu_2)}\sum_{i,j=1,i\neq j}^{2}\left\{1-\frac{e_ie_j}{m_ir_{ij}}\left[e^{-\kappa r_{ij}}+\frac{\boldsymbol{v}_i\cdot\boldsymbol{v}_j}{2c^2}+\frac{\left(\boldsymbol{x}_{ij}\cdot\boldsymbol{v}_i\right)\left(\boldsymbol{x}_{ij}\cdot\boldsymbol{v}_j\right)}{2c^2r_{ij}^2}\right]\mu_i\right.$$
$$+\frac{e_i^2e_j^2}{2m_i^2r_{ij}^2}+\left[e^{-\kappa r_{ij}}+\frac{\boldsymbol{v}_i\cdot\boldsymbol{v}_j}{2c^2}+\frac{\left(\boldsymbol{x}_{ij}\cdot\boldsymbol{v}_i\right)\left(\boldsymbol{x}_{ij}\cdot\boldsymbol{v}_j\right)}{2c^2r_{ij}^2}\right]^2\mu_i^2 \tag{11}$$
$$\left.-\frac{e_i^3e_j^3}{6m_i^3r_{ij}^3}\left[e^{-\kappa r_{ij}}+\frac{\boldsymbol{v}_i\cdot\boldsymbol{v}_j}{2c^2}+\frac{\left(\boldsymbol{x}_{ij}\cdot\boldsymbol{v}_i\right)\left(\boldsymbol{x}_{ij}\cdot\boldsymbol{v}_j\right)}{2c^2r_{ij}^2}\right]^3\mu_i^3+\cdots\right\}\exp\left(\frac{-\mu_1}{\sqrt{1-\left(v_1/c\right)^2}}+\frac{-\mu_2}{\sqrt{1-\left(v_2/c\right)^2}}\right)$$

with $F^{(1)}$ given by Equation (5) (in fact only terms of $F^{(1)}$ up to $1/c^2$ must be retained). It can be verified that $G(1,2)$ given by Equation (9), with ζ_{1R}^0 satisfies identically the condition (6) for $t_{12}=0$, to first order in $1/c^2$.

The standard two-body distribution function for dilute slightly relativistic plasma has been calculated previously by Kosachev and Trubnikov [21] starting from the Darwin Lagrangian. Lapiedra and Santos [19] result agrees with theirs to order $1/c^2$, but not to higher-order terms. We think that these terms are meaningless unless one goes beyond the Darwin Lagrangian, which is only correct to order $1/c^2$. The Lagrangian approach introduces some modification of the expression for the statistical sum and what is more important the use of the restricted Breit-Darwin Hamiltonian leads to the wrong behavior of the pair correlation function and therefore to the incorrect expressions for various thermodynamical quantities [22].

According to our knowledge few scientists studied the binary distribution function for relativistic dilute plasma [15]; the new in our article is using effects of thermal parameter on the values of binary relativistic distribution function for one and two-component plasma.

4. The Triplet Relativistic Distribution Function

The triplet distribution function $F^{(3)}(1,2,3)$ is defined in such a way that $F^{(3)}\left(t,x_i,u_i\right)\mathrm{d}x_i\mathrm{d}u_i$, $i=1,2,3$ the calculation of the relativistic interaction between three charged particles seems rather involved, because the force on particle 1 at time *t* would depend on the position and velocity of particles 2 and 3 at a retarded time. Also, the position and velocity of 1 at that time depends on the position and velocity of 1 at an earlier time.

Substituting Equation (4) and (5) into (2) for $s=1,2,3$ we obtain

$$F^{(3)}(1,2,3)=F^{(1)}(1)F^{(1)}(2)F^{(1)}(3)\sum_{i,j=1,i\neq j}^{3}\left\{1-\frac{e_ie_j}{m_ir_{ij}}e^{-\kappa r_{ij}}\mu_i+\frac{e_ie_j}{4m_j\pi^2}l_{ij}\,\mu_j+\frac{e_i^2e_j^2}{2m_i^2r_{ij}^2}e^{-2\kappa r_{ij}}\mu_i^2\right.$$
$$+\frac{e_i^2e_j^2\mu_j\mu_i}{2m_im_j\pi^2r_{ij}}e^{-\kappa r_{ij}}l_{ij}-\frac{3e_i^3e_j^3}{8m_i^2m_j\pi^2r_{ij}^2}e^{-2\kappa r_{ij}}\mu_i^2\mu_jl_{ij}-\frac{3e_i^3e_j^3}{16m_im_j^2\pi^4r_{ij}}e^{-\kappa r_{ij}}\mu_i\mu_j^2l_{ij}^2 \tag{12}$$
$$\left.+\frac{e_i^2e_j^2}{16m_j^2\pi^4}l_{ij}^2\mu_j^2-\frac{e_i^3e_j^3}{6m_i^3r_{ij}^3}e^{-3\kappa r_{ij}}\mu_i^3-\frac{e_i^3e_j^3}{64m_j^3\pi^6}l_{ij}^3\mu_j^3+\cdots\right\}.$$

If we used the Kirkwood superposition approximation (KSA) [23]; which is consisting of the assumption that the potential in a set of three particles is the sum of the three pair potentials, this is equivalent to assuming that the triplet distribution function is the product of the three radial distribution functions

$$F^{(3)}(1,2,3)=F^{(2)}(1,2)F^{(2)}(2,3)F^{(2)}(1,3). \tag{13}$$

It has been mentioned that before 2003 there is no direct measurement of three-body correlation function and this is the first study for the triplet distribution function in the case of dilute relativistic plasma. Such measurement requires the knowledge of the positions of three particles at the same time which is technically very demanding to obtain in 3D samples [13]; this means that the experiment to measure the triplet distribution function directly requires high precision so that the multiple scattering can be differentiated from single scattering. Now because of the existence of video-microscopy, a modern experimental technique applied to colloidal systems to directly measure all particles' positions at all times. Essentially, one follows the phase-space trajectory of the

system as it evolves in time, and has thus the same amount of information as one obtains in a simulation. Similarly the particle positions have measured in dusty plasmas [24].

From Equation (12) we can obtain the classical TDF for two-component plasma in the following form:

$$F^{(3)}(1,2,3)=F^{(1)}(1)F^{(1)}(2)F^{(1)}(3)\sum_{i,j=1,i\neq j}^{3}\left\{1-\frac{e_ie_j}{m_ir_{ij}}\left[e^{-\kappa r_{ij}}+\frac{\boldsymbol{v}_i\cdot\boldsymbol{v}_j}{2c^2}+\frac{(\boldsymbol{x}_{ij}\cdot\boldsymbol{v}_i)(\boldsymbol{x}_{ij}\cdot\boldsymbol{v}_j)}{2c^2r_{ij}^2}\right]\mu_i+\frac{e_i^2e_j^2}{2m_i^2r_{ij}^2}\right.$$
$$\left.+\left[e^{-\kappa r_{ij}}+\frac{\boldsymbol{v}_i\cdot\boldsymbol{v}_j}{2c^2}+\frac{(\boldsymbol{x}_{ij}\cdot\boldsymbol{v}_i)(\boldsymbol{x}_{ij}\cdot\boldsymbol{v}_j)}{2c^2r_{ij}^2}\right]^2\mu_i^2-\frac{e_i^3e_j^3}{6m_i^3r_{ij}^3}\left[e^{-\kappa r_{ij}}+\frac{\boldsymbol{v}_i\cdot\boldsymbol{v}_j}{2c^2}+\frac{(\boldsymbol{x}_{ij}\cdot\boldsymbol{v}_i)(\boldsymbol{x}_{ij}\cdot\boldsymbol{v}_j)}{2c^2r_{ij}^2}\right]^3\mu_i^3+\cdots\right\}. \tag{14}$$

And we also can used (KSA) which is given in Equation (13) to get it in the form

$$F^{(3)}(1,3,2)=\left(F^{(1)}(1)F^{(1)}(2)F^{(1)}(3)\right)^2\left\{1-\frac{\mu_1e_1e_2}{m_1r_{12}}\left[e^{-\kappa r_{12}}+\frac{\boldsymbol{v}_1\cdot\boldsymbol{v}_2}{2c^2}+\frac{(\boldsymbol{x}_{12}\cdot\boldsymbol{v}_1)(\boldsymbol{x}_{12}\cdot\boldsymbol{v}_2)}{2c^2r_{12}^2}\right]\right.$$
$$\left.-\frac{\mu_2e_3e_2}{m_2r_{32}}\left[e^{-\kappa r_{32}}+\frac{\boldsymbol{v}_3\cdot\boldsymbol{v}_2}{2c^2}+\frac{(\boldsymbol{x}_{32}\cdot\boldsymbol{v}_3)(\boldsymbol{x}_{32}\cdot\boldsymbol{v}_2)}{2c^2r_{ij}^2}\right]-\frac{\mu_3e_1e_3}{m_3r_{13}}\left[e^{-\kappa r_{13}}+\frac{\boldsymbol{v}_1\cdot\boldsymbol{v}_3}{2c^2}+\frac{(\boldsymbol{x}_{13}\cdot\boldsymbol{v}_1)(\boldsymbol{x}_{13}\cdot\boldsymbol{v}_3)}{2c^2r_{ij}^2}\right]+\cdots\right\}. \tag{15}$$

The triplet and quadruple distribution functions as well as binary distribution function must be incorporated for a more accurate and complete discussion of macroscopic equilibrium properties. A little is known about distribution functions involving more than two particles, and in particular about the three-particle (or triplet) distribution functions. This is of course due to the greater mathematical complexity of higher order correlation functions, and to a lack of a direct link with experiment [25]. And this is considered the first study for the triplet distribution functions for dilute relativistic plasma.

5. Conclusions

In many physical systems, the description of a plasma as a Coulomb system is sufficient to reproduce most of the properties of interest. If the system is cold enough, the mean velocities of the particles are much smaller than the speed of light, and the charges may be assumed to interact via the instantaneous Coulomb potential. However, at sufficiently high temperatures, this approximation is no longer valid, and the contributions of the relativistic effects (which include, apart from the trivial kinetic corrections and of course all the retardation effects) must be incorporated when studying the equilibrium properties of the system [26].

In the classical (non-quantum) case, the systematic approach adopted here follows the traditional route of the relativistic Bogoliubov-Born-Green-Kirkwood-Yvon (BBGKY) hierarchy for the reduced distribution functions by formal density expansion. Interestingly, this study is the first to display the effect of thermal parameter of plasma in the classical binary and triplet distribution function. Also the triplet relativistic distribution function for dilute plasma was calculated from the relativistic BBGKY hierarchy. We used the results to obtain the analytical forms of the classical triplet distribution functions for one- and two-component plasmas.

In **Figure 3** and **Figure 4** we noticed that the value of both binary and triplet relativistic distribution function increased when the μ value decreased at very high temperature. Physically these results seem acceptable because from the definition of the phase-space distribution function which means the number of particles per unit volume of space per unit volume of velocity space, and at high temperature, the velocity of the particles increases and as a result the number of particles increases per unit volume.

Our calculations are grounded in the classical relativistic statistical mechanics. Plasma is non-degenerate. The system is not dense, so one may neglect the contributions of higher order particle interactions. **Figure 5** and **Figure 6** show that the two- and three-particle distribution functions have become concentrated to ever-smaller region of speed v, dramatically increasing the thermal parameter μ. When we made a comparison between the two-particle relativistic distribution function from Equation (11) and the result of Barcons and Lapiedra as shown in **Figure 7**, the results were nearly similar at very high velocities.

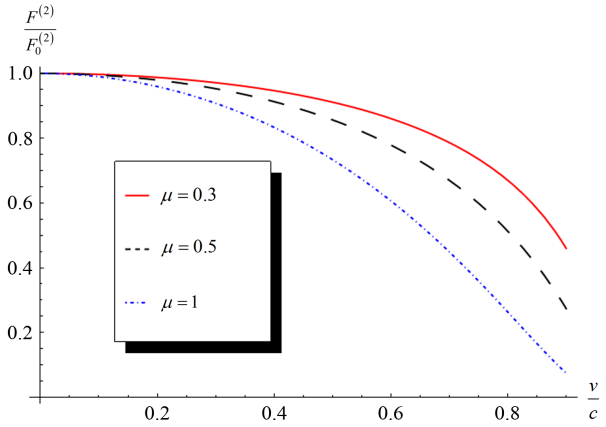


Figure 3. The two-particle relativistic distribution function in the particle velocity interval (0, 0.9*c*) for one-component plasma.

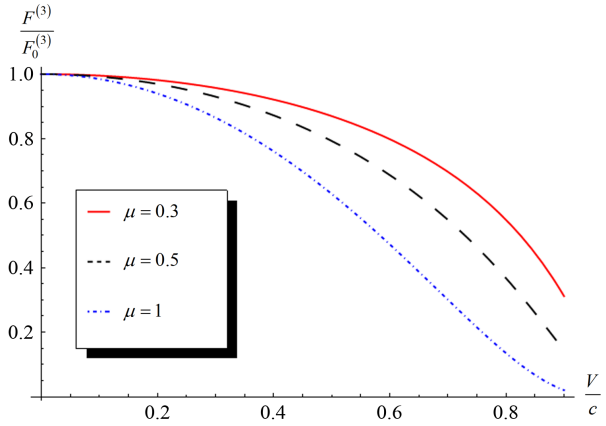


Figure 4. The three-particle relativistic distribution function in the particle velocity interval (0, 0.9*c*) for one-component plasma for different values of μ, $v_1 = v_2 = v_3 = v$.

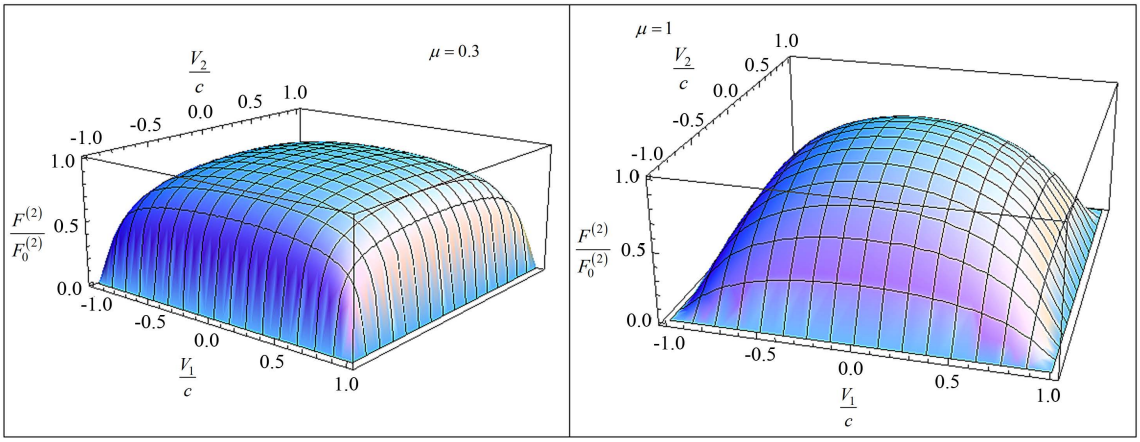


Figure 5. The two-particle relativistic distribution function for two-component plasma in the particle velocity interval $(-c, c)$ at $\mu = 1$; $\mu = 0.3$ and $v_1 \neq v_2$.

Figure 8 and **Figure 9** show that the two- and three-particle distribution functions in the thermal parameter of plasma μ interval $(0,1)$ for one-component plasma for different values of speed $v = 0.8c$, $v = 0.86c$, and $v = 0.9c$. We note that the curves are very close to each other, dramatically increasing the thermal parameter

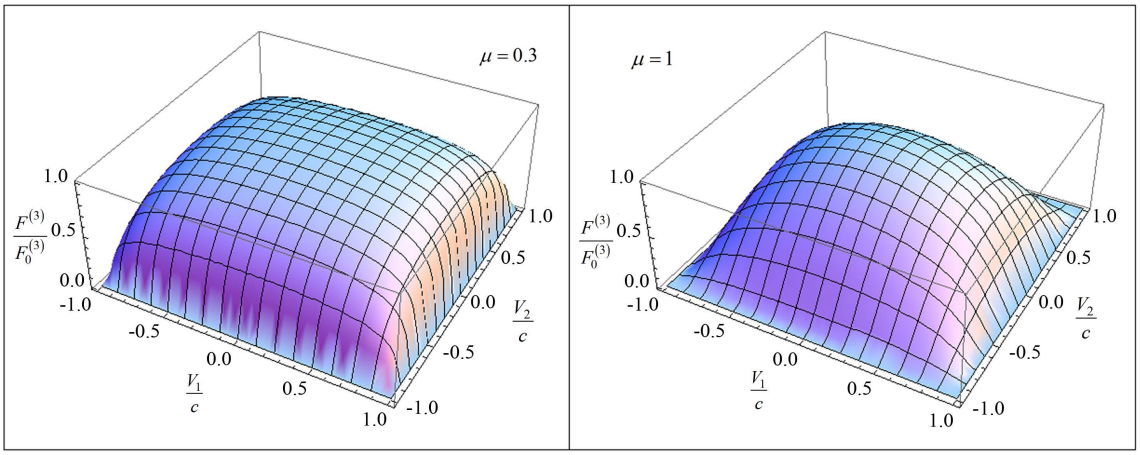


Figure 6. The three-particle relativistic distribution function in the particle velocity interval (0, 0.9*c*) for one-component plasma for different values of μ; $v_1 = v_2 = v_3 = v$.

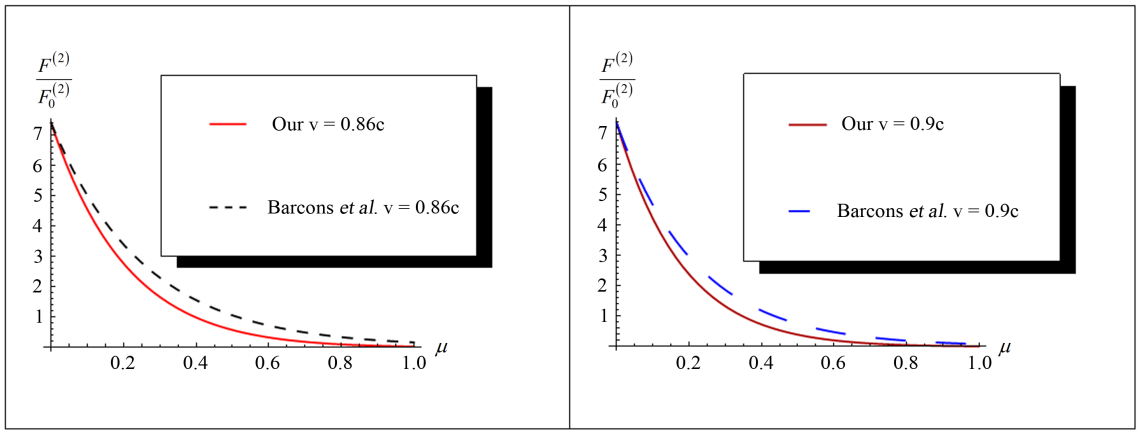


Figure 7. The comparison between $F^{(2)}$ from our result and from Barcons and Lapiedra [9] for $v = 0.86c$ and $v = 0.9c$.

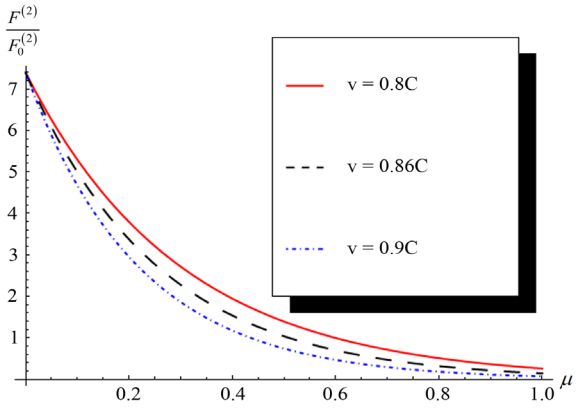


Figure 8. The two-particle relativistic distribution function in the thermal parameter of plasma μ interval (0, 1) for one-component plasma for different values of speed, $v = 0.8c$, $v = 0.86c$, $v = 0.9c$.

μ and for three-particle distribution function more than two-particle distribution function.

Figure 10 shows the comparison between the three-particle relativistic distribution function from Equations (14) and (15). One of them is based on the Kirkwood superposition approximation (KSA) which is consisting of the assumption that the potential in a set of three particles is the sum of the three pair potentials. This is equivalent

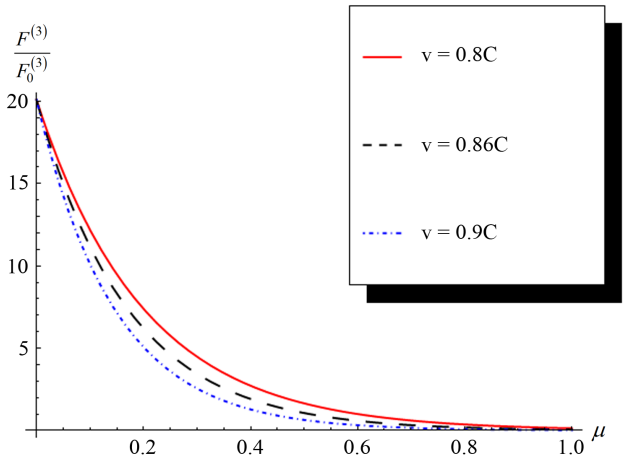


Figure 9. The three-particle relativistic distribution function in the thermal parameter of plasma μ interval (0, 1) for one-component plasma for different values of speed, $v = 0.8c$, $v = 0.86c$, $v = 0.9c$.

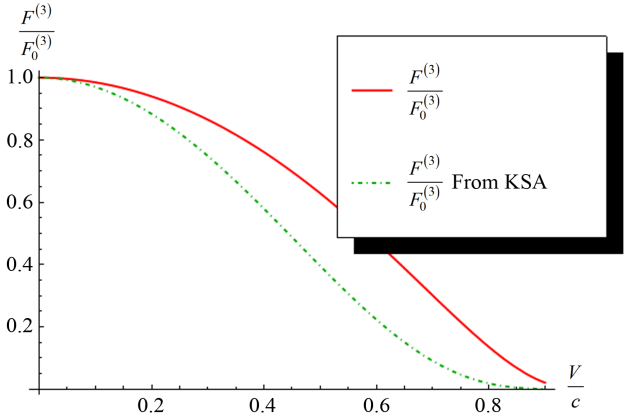


Figure 10. The comparison between $F^{(3)}$ from Equation (14) and $F^{(3)}$ from (KSA) for two-component plasma for $\mu = 1$; at $v_1 = v_2 = v_3 = v$.

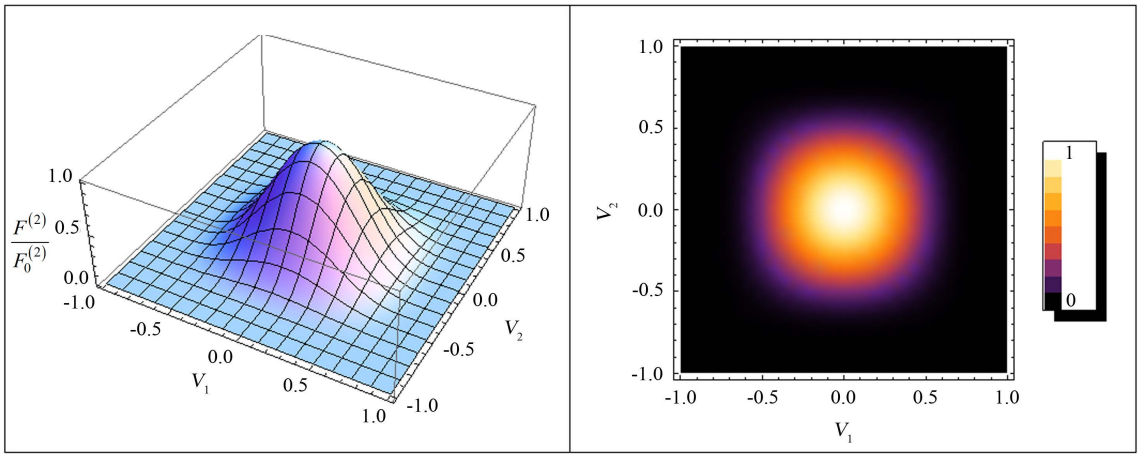


Figure 11. The two-particle relativistic distribution function in the particle velocity interval $(-c, c)$ for $\mu > 1$ (weakly relativistic case).

to assume that the triplet distribution function is the product of the three radial distribution functions, and the other form is calculated by using the relativistic BBGKY hierarchy.

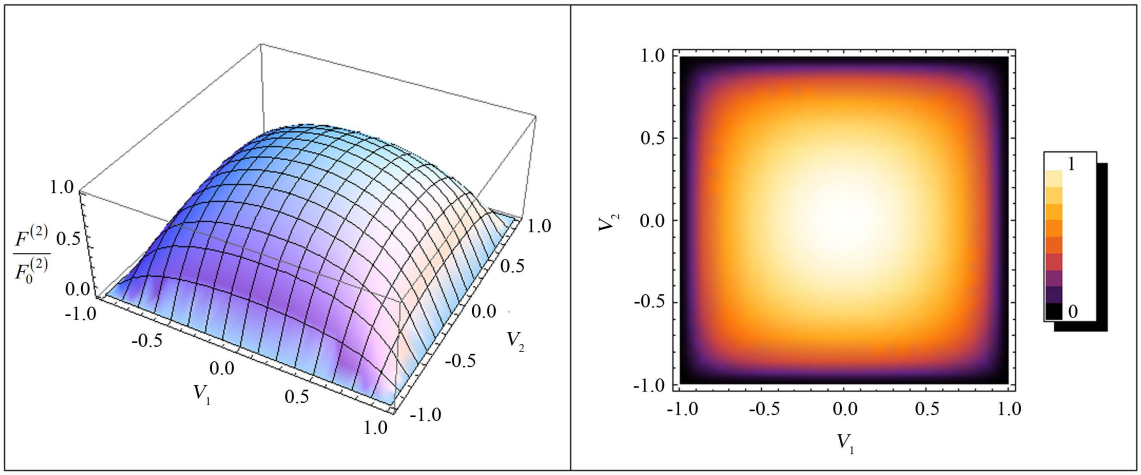


Figure 12. The two-particle relativistic distribution function in the particle velocity interval $(-c, c)$ for $\mu = 1$ (relativistic case).

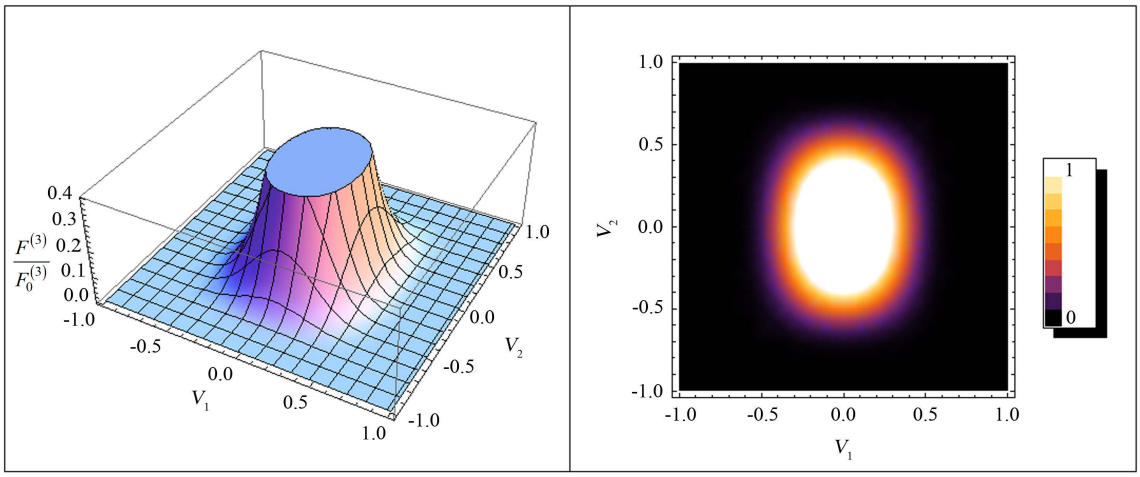


Figure 13. The two-particle relativistic distribution function in the particle velocity interval $(-c, c)$ for $\mu > 1$ (weakly relativistic case).

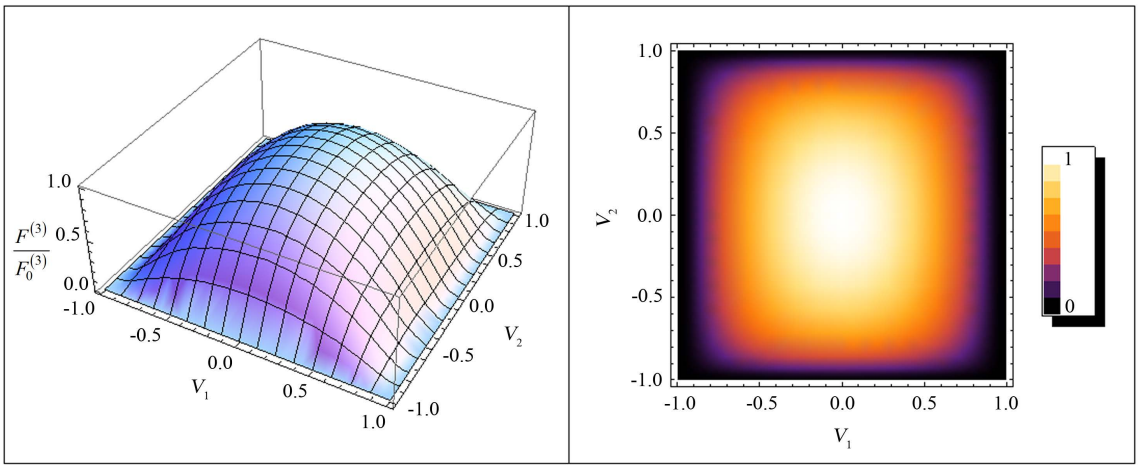


Figure 14. The three-particle relativistic distribution function in the particle velocity interval $(-c, c)$ for $\mu = 1$ (relativistic case).

Finally, we can note that from **Figures 11-14**, the distribution function has become more concentrated to ever-smaller region of speed v, dramatically increasing the thermal parameter μ and for three-particle distribution function more than two-particle distribution function.

References

[1] Droz-Vincent, P. (1997) Direct Interactions in Relativistic Statistical Mechanics. *Foundations of Physics*, **27**, 363-387. http://dx.doi.org/10.1007/BF02550162

[2] Bogoliubov, N.N. (1946) Probleme der Dynamischen Theorie in der Statistischen Physik. Moskau.

[3] Simonella, S. (2011) BBGKY Hierarchy for Hard Sphere Systems. Ph.D. Dissertation, Sapienza University of Rome, Rome.

[4] Kraeft, W.D., Kremp, D. and Ebeling, W. (1986) Quantum Statistics of Charged Particle Systems. Akademie Verlag, Berlin. http://dx.doi.org/10.1007/978-1-4613-2159-0

[5] Kahlbaum, T. (1999) Density Expansions of the Reduced Distribution Functions and the Excess Free Energy for Plasma with Coulomb and Short-Range Interactions. *Contributions to Plasma Physics*, **39**, 181-184. http://dx.doi.org/10.1002/ctpp.2150390144

[6] Hussein, N.A. and Eisa, D.A. (2011) The Quantum Equation of State of Fully Ionized Plasmas. *Contributions to Plasma Physics*, **51**, 44-50. http://dx.doi.org/10.1002/ctpp.201110003

[7] Eisa, D.A. (2012) The Classical Binary and Triplet Distribution Functions for Two Component Plasma. *Contributions to Plasma Physics*, **52**, 261-275. http://dx.doi.org/10.1002/ctpp.201100030

[8] Arendt Jr., P.N. and Eilek, J.A. (2000) The Pair Cascade in Strong and Weak Field Pulsars. *Pulsar Astronomy*—2000 *and beyond*, **202**, 445-448.

[9] Barcons, X. and Lapiedra, R. (1985) Statistical Mechanics of Classical Dilute Relativistic Plasmas in Equilibrium. *Journal of Physics A*: *Mathematical and General*, **18**, 271-285. http://dx.doi.org/10.1088/0305-4470/18/2/017

[10] Swisdak, M. (2013) The Generation of Random Variates from a Relativistic Maxwellian Distribution. *Physics of Plasmas*, **20**, Article ID: 062110. http://dx.doi.org/10.1063/1.4812459

[11] Ares de Parga, G. and Lopez-Carrera, B. (2011) Relativistic Statistical Mechanics vs. Relativistic Thermodynamics. *Entropy*, **13**, 1664-1693. http://dx.doi.org/10.3390/e13091664

[12] Treumann, R.A., Nakamura, R. and Baumjohann, W. (2011) A Model of So-Called "Zebra" Emissions in Solar Flare Radio Burst Continua. *Annales Geophysicae*, **29**, 1673-1682. http://dx.doi.org/10.5194/angeo-29-1673-2011

[13] Sciortino, F. and Kob, W. (2001) Debye-Waller Factor of Liquid Silica: Theory and Simulation. *Physical Review Letters*, **86**, 648-651. http://dx.doi.org/10.1103/PhysRevLett.86.648

[14] Rosenfeld, Y., Levesque, D. and Weis, J.J. (1990) Free-Energy Model for the Inhomogeneous Hard-Sphere Fluid Mixture: Triplet and Higher-Order Direct Correlation Functions in Dense Fluids. *The Journal of Chemical Physics*, **92**, 6818-6832. http://dx.doi.org/10.1063/1.458268

[15] Lapiedra, R. and Santos, E. (1981) Classical Relativistic Statistical Mechanics: The Case of a Hot Dilute Plasma. *Physical Review D*, **1**, 2181-2188. http://dx.doi.org/10.1103/PhysRevD.23.2181

[16] Alam, M.S., Masud, M.M. and Mamun, A.A. (2014) Effects of Two-Temperature Superthermal Electrons on Dust-Ion-Acoustic Solitary Waves and Double Layers in Dusty Plasmas. *Astrophysics and Space Science*, **349**, 245-253. http://dx.doi.org/10.1007/s10509-013-1639-3

[17] Masud, M.M., Kundu, N.R. and Mamun, A.A. (2013) Obliquely Propagating Dust-Ion Acoustic Solitary Waves and Their Multidimensional Instabilities in Magnetized Dusty Plasmas with Bi-Maxwellian Electrons. *Canadian Journal of Physics*, **91**, 530-536. http://dx.doi.org/10.1139/cjp-2012-0390

[18] Lazar, M., Stockem, A. and Schlickeiser, R. (2010) Towards a Relativistically Correct Characterization of Counter-streaming Plasmas. I. Distribution Functions. *Open Plasma Physics Journal*, **3**, 138-147. http://dx.doi.org/10.2174/1876534301003010138

[19] Lapiedra, R. and Santos, E. (1983) Classical Dilute Relativistic Plasma in Equilibrium. Two-Particle Distribution Function. *Physical Review A*, **27**, 422-430. http://dx.doi.org/10.1103/PhysRevA.27.422

[20] Turski, L.A. (1974) Pair Correlation Function for a System with Velocity-Dependent Interactions. *Journal of Statistical Physics*, **11**, 1-16. http://dx.doi.org/10.1007/BF01019474

[21] Kosachev, V.V. and Trubnikov, B.A. (1969) Relativistic Corrections to the Distribution Functions of Particles in a High-Temperature Plasma. *Nuclear Fusion*, **9**, 53-56. http://dx.doi.org/10.1088/0029-5515/9/1/006

[22] Bel, L., Salas, A. and Sanchez, J.M. (1973) Approximate Solutions of Predictive Relativistic Mechanics for the Electromagnetic Interaction. *Physical Review D*, **7**, 1099-1106. http://dx.doi.org/10.1103/PhysRevD.7.1099

[23] Kirkwood, J.G. (1935) Statistical Mechanics of Fluid Mixtures. *The Journal of Chemical Physics*, **3**, 300-313. http://dx.doi.org/10.1063/1.1749657

[24] Bonitz, M., Henning, C. and Block, D. (2010) Complex Plasmas: A Laboratory for Strong Correlations. *Reports on Progress in Physics*, **73**, Article ID: 066501. http://dx.doi.org/10.1088/0034-4885/73/6/066501

[25] Barrat, J.L., Hansen, J.P. and Pastore, G. (1988) On the Equilibrium Structure of Dense Fluids: Triplet Correlations, Integral Equations and Freezing. *Molecular Physics*, **63**, 747-767. http://dx.doi.org/10.1080/00268978800100541

[26] Kalman, G.J., Rommel, J.M., Blagoev, K. and Blagoev, K. (1998) Strongly Coupled Coulomb Systems. Springer, Berlin. http://www.springer.com/physics/particle+and+nuclear+physics/book/978-0-306-46031-9

Natural Science, 2015, 7, 55-63
Published Online January 2015 in SciRes. http://www.scirp.org/journal/ns
http://dx.doi.org/10.4236/ns.2015.71006

Scientific Research Publishing

An Econophysics Model of Financial Bubbles

Bodo Herzog

Department of Economics, ESB Business School, Reutlingen, Germany
Email: Bodo.Herzog@Reutlingen-Universtiy.de

Received 7 January 2015; accepted 25 January 2015; published 29 January 2015

Abstract

Usually financial crises go along with bubbles in asset prices, such as the housing bubble in the US in 2007. This paper attempts to build a mathematical model of financial bubbles from an econophysics, and thus a new perspective. I find that agents identify bubbles only with a time delay. Furthermore, I demonstrate that the detection of bubbles is different on either the individual or collective point of view. Second, I utilize the findings for a new definition of asset bubbles in finance. Finally, I extend the model to the study of asset price dynamics with news. In conclusion, the model provides unique insights into the properties and developments of financial bubbles.

Keywords

Financial Bubbles, Econophysics, Wave Equation, Financial Crises

1. Introduction

The phenomenon of a financial bubble is debated in economics for centuries. There are several important questions: What is a bubble? How can we identify a bubble? What is the root cause of a bubble and what triggers a bubble burst? So far, different economic and mathematical models study different issues of bubbles. However to my knowledge, there is no consistent model that addresses all issues at once. In addition, the complexity in financial markets due to social interaction and behavioural elements makes the study of those bubbles difficult.

This paper attempts to develop a new model of financial bubbles. My model utilizes the idea of particle dynamics in physics. I translate this idea into the dynamics of financial assets without relying on stochastic and martingale theory [1]. Hence, I intend to throw light on the bubble definition, detection, and dynamics without using stochastic theory. I find that the model is tractable and, at the same time, provides valuable evidence. First, I demonstrate that traders always recognize bubbles with a time delay. Second, I study the properties of bubble detection in detail. It turns out that the modeling of herd behaviour and the impact of news are major forces of financial bubbles. Third, I analyze news of asset prices and the relationship to financial bubbles. At the end, I develop some regulatory policy recommendations for financial markets.

How to cite this paper: Herzog, B. (2015) An Econophysics Model of Financial Bubbles. *Natural Science*, **7**, 55-63.
http://dx.doi.org/10.4236/ns.2015.71006

The remainder of the paper is structured as follows. Section 2 discusses the different forms of financial bubbles and presents a literature review. In Section 3, I derive the mathematical model and study the main implications. Section 4 extends the model and studies the impact of news. Finally, Section 5 provides concluding remarks.

2. Motivation and Literature Review

Let me start with a quote by Charles MacKay about the forces of financial bubbles in general: "*Men, it has been well said, think in herds*; *it will be seen that they go mad in herds, while they only recover their senses slowly, and one by one*." No doubt, herd behaviour turns out to be important during almost all financial bubbles. Thus it was studied in the economic literature for decades. In general, there are different types of herd behaviour and they are characterized by imperfect information, a systematic bias in expectations, and regulatory arbitrage due to perverse incentives.

The need for better models in this field has to do with the tremendous social costs of a bubble burst [2] [3]. Indeed, a bubble exacerbates volatility, destabilize markets and increase the fragility of the overall financial system [4]. In the economics literature there are three different lines of modeling but the reference point is always the seminal work by [5].

First, economists have analyzed "information-based" or "rational" herd behaviour. That research is pioneered by DeLong *et al.* [6], Froot *et al.* [7], and Lux and Marchesi [8]. They argue that individuals observe the action of each other but not the private information that they receive. The arrival of too little information and idiosyncratic news leads to either rational but always erratic decision-making. Indeed, agents do not realize that they are in a financial bubble due to flawed expectations and imperfect information. This leads to fragile financial markets in the end. There are several modeling variants of information-based herd behaviour in economics [9]-[13]. For instance, Maug and Naik [14] find that risk avers investors' that have imperfect information about stock returns tend to follow a herd rather soon.

A second line of literature in economics argues for "reputation-based herding". This was first explored by Scharfstein and Stein [15]. They show that uncertainty about the own ability leads to path-dependency; *i.e.* you follow a naive trend or a heuristic. Commonly, investors use heuristics in finance, such as a rule-of-thumb. Therefore, skilled agents with reputation in finance are less prone to herd behaviour and financial bubbles. A third approach is "compensation-based herding". That form of herd behaviour appears if an investors' compensation depends on the performance to others.

This paper develops a unique link between the different types of herd behaviour and a general "Model of Financial Bubbles". To analyze the complex financial dynamics, I build a mathematical model based on particle physics. In fact, there exist an obvious parallel of both particle and financial dynamics. The dynamics of both systems is based on the interaction of single elements/agents that lead to a collective outcome. From an aggregate perspective, the frequent events of traffic jams are often the result of uncoordinated interaction of individual's at overcrowded highways and not always a car accident. Interestingly, this is similar to the root cause of herd behaviour in financial markets. Consequently, a financial bubble, *i.e.* a massive price increase (decrease), can be caused by a large number of buying (selling) individuals, especially if they build a herd. Thus, bubbles can come into existence without any change of the fundamental value. The ups and downs of asset prices attract similar agents and thus form a herd. Soon later, I obtain a bubble or crash. Unfortunately, the collective behaviour creates even more volatility, higher risks, and massive welfare losses. A second parallel between physics and financial bubbles is the impact of news as a root cause or trigger of a bubble. Ad-hoc news either positive or negative can be compared to an unforeseeable red light or traffic flows via an exit or entrance to a highway. Hence, the main objective of my model is twofold: 1) build a coherent model and 2) find new implications for empirical finance in future.

There is also a large descriptive literature about collective behaviour in speculative bubbles [16]-[19]. The model by Galbraith [20] identifies accelerating bubbles as a precursory pattern for essentially all crashes. A similar result is found by Kindleberger and Fransman [21]. Although these papers provide interesting insights, they have limitations. They cannot sufficiently explain the causes and triggers of both bubbles and crashes. One reason for this limitation is due to the methodological disadvantage of economic models. The economic models utilize the individualistic and equilibrium methodology. These two elements have a serious downside in the study of complex dynamics and social interactions [22]. Therefore, my econophysic model provides a new per-

spective and is ready to analyse complex dynamics and social interactions.

3. Econophysics Model

The financial asset price dynamic is denoted by $p_i(f,t)$. This is a function of asset *i* dependent on two variables: the fundamental price vector *f* and time *t*. The benefit-loss-rate (BLR) of each asset *i* is defined by $v_i = \partial p_i(f,t)/\partial t$. The BLR can be interpreted as the price change or the speed of asset *i*. The acceleration of the price change, *i.e.* the rate of change of the BLR, is given by the second derivative, such as $a_i = \partial^2 p_i(f,t)/\partial t^2$. There are different ways to define bubbles or herd behaviour. Obviously, a financial bubble is closely connected to a high BLR, v_i, and a high rate of change of the BLR, a_i. Hence, I define a unique benefit-loss-field (BLF)

$$u_i^k\left(p_i(f,t),t\right) \tag{1}$$

where *k* captures the agents' expectation or reputation level. I assume $k^H > k^L$ and the *H* denotes a high and *L* a low reputation level respectively. The existence of a BLF implies that for each price process, $p_i(f,t)$, exists only one unique benefit or loss. Thus, this model does not simultaneously allow different BLF for the same asset price dynamic; *i.e.* $v_i(f,t) = v_j(f,t) \Rightarrow u_i = u_j$ for $i \neq j$. However, I assume different dynamics for the rate of change, $a_i \neq a_j$. This assumption is consistent with financial market dynamics [23]. At the beginning, I study the market dynamics along with continuous functions without jumps. This is known as the continuum hypothesis. In addition, let me define the following variables:

Definition 1. Trading volume (flow) of asset *i* is defined as the amount of buy and sell orders (number of trades) times the buy and sell price per day. It is denoted by the function $q\left(p(f,t),t\right)$.

Definition 2. Trading density of asset *i* is defined as the number of trades within a certain price range $p_a < p_i < p_b$. It is denoted by $\rho\left(p(f,t),t\right)$.

Next, based on both definitions, I define a fundamental relationships:

Definition 3. There exists a relationship between trading volume, trading density and the benefit-loss-field, such as

$$q(p,t) = \rho(p,t) * u(p,t). \tag{2}$$

The last equation is a fundamental law in applied physics and in my econophysics model on financial assets. In the following subsections, I derive the model and study the identification of a financial bubble.

3.1. Derivation of Model Equation

Next, I consider the two fundamental variables $\rho(p,t)$ and $u(p,t)$. The function $\rho(p,t)$ can be measured in reality, while the BLF, $u(p,t)$, is unobservable. As defined above, the motion of each asset price satisfies a first-order differential equation, such as

$$\frac{\mathrm{d}p}{\mathrm{d}t} = u(p,t) \quad \text{with} \quad p(0) = p_0. \tag{3}$$

Solving this equation would determine the asset price for which an agent is willing to buy or sell at a later time under a uniform stock market. However, finding the function $u(p,t)$ is not an easy task in reality. Therefore, I choose a different approach. Suppose, the aggregate number of trades is equal to *N*. Then, you can define *N* as the integral of the trading density $\rho(p,t)$ in the given price range $p_a = a$ and $p_b = b$, as

$$N = \int_a^b \rho(p,t)\,\mathrm{d}p. \tag{4}$$

In general, the number of trades change over time. It increases with the inflow of buying agents at the lower bound $q(p_a,t)$ and it decreases with the outflow of sellers at the upper price-limit $q(p_b,t)$. Hence, I obtain

$$\frac{\mathrm{d}N}{\mathrm{d}t} = q(p_a,t) - q(p_b,t). \tag{5}$$

Combining Equations (4) and (5), per definition yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_a^b \rho(p,t)\,\mathrm{d}p = q(p_a,t) - q(p_b,t). \tag{6}$$

The last equation can be labelled a "conversation law". Hence, the equation denotes that the number of trades in a certain price range is equal to the difference of the trading volume in this price range. Starting with that equation, I derive the model. Consider the integral conversation law over a small interval from $p = a$ to $p = a + \Delta a$. Equation (6), yields

$$\frac{\partial}{\partial t}\int_a^{a+\Delta a} \rho(p,t)\,\mathrm{d}p = q(a,t) - q(a+\Delta a,t). \tag{7}$$

Now, divide by $-\Delta a$ and take the limit as $\Delta a \to 0$:

$$\lim_{\Delta a\to 0} \frac{\partial}{\partial t}\frac{1}{-\Delta a}\int_a^{a+\Delta a} \rho(p,t)\,\mathrm{d}p = \lim_{\Delta a\to 0} \frac{q(a,t) - q(a+\Delta a,t)}{-\Delta a}. \tag{8}$$

The right-hand side of that equation is the first derivative of $q(a,t)$ with respect to *a*, *i.e.* $(\partial/\partial a)q(a,t)$. Next, let me redefine the left-hand side with a new function. I introduce the function $N(\bar{p},t)$, the number of trades between any price p_0 and $\bar{p}$. I obtain

$$N(\bar{p},t) = \int_{p_0}^{\bar{p}} \rho(p,t)\,\mathrm{d}p. \tag{9}$$

The average number of trades in the defined price interval $(a, a+\Delta a)$ is given by

$$-\frac{1}{\Delta a}\int_a^{a+\Delta a} \rho(p,t)\,\mathrm{d}p = \frac{N(a+\Delta a,t) - N(a,t)}{-\Delta a}. \tag{10}$$

In limit for $\Delta a \to 0$, the right-hand side obtains $\partial N(a,t)/\partial a$. Thus, I can rewrite Equation (10), as

$$\rho(a,t) = \frac{\partial N(a,t)}{\partial a}. \tag{11}$$

The left-hand side of Equation (8) equals now $-(\partial/\partial t)(\rho(a,t))$. The following equation holds for all *a*. Therefore, it is more appropriate to replace *a* by *p*. I finally obtain

$$\frac{\partial \rho(p,t)}{\partial t} + \frac{\partial}{\partial p}\left[q(p,t)\right] = 0 \tag{12}$$

which is a wave equation in form of a partial differential equation (PDE). This equation can be rewritten with the help of the fundamental law from Equation (2), where *q* is defined as $q = \rho * u$. Hence, trading volume is a function of the density, $q(\rho)$, and thus I obtain a non-linear PDE of the form

$$\frac{\partial \rho}{\partial t} + \frac{\mathrm{d}q}{\mathrm{d}\rho}\frac{\partial \rho}{\partial p} = 0. \tag{13}$$

Solving this PDE requires further assumptions. Suppose that the BLF is defined as a function of the density, such as $u = u(\rho)$. Moreover, this function has the following properties: 1) $u^j(0) = u^j_{k,\max}$, where $u^j_{H,\max} > u^j_{L,\max}$. That is intuitively clear because the higher the reputation (ability) the greater the likelihood of gains and hence the level of the BLF. 2) The BLF and thus the profits are zero in a perfect competitive market according to standard economics; $u^j(\rho_{\max}) = 0$. Hence, agents are price-taker in this market environment. 3) Logically the function value, $u(\rho)$, is declining in ρ. Consequently, the first derivative is negative: $\frac{\mathrm{d}u}{\mathrm{d}\rho} = u'(\rho) \le 0$.

Figure 1 illustrates these properties in a linear model graphically. If the trading density is low, then the BLF is maximal, *i.e.* assets with a limited amount of trades have a bigger potential for benefits (or losses) by a few active traders. This assumes that a low amount of orders significantly change the price in spare markets. Simultaneously, this assumption indicates either a high degree of asymmetric information or uncertainty. Both problems are revealed by a low trading density. In the other extreme of maximal trading density, I obtain a market environment with "full information" and that would imply a benefit (loss) of almost zero. Hence, in this market environment agents are price-taker.

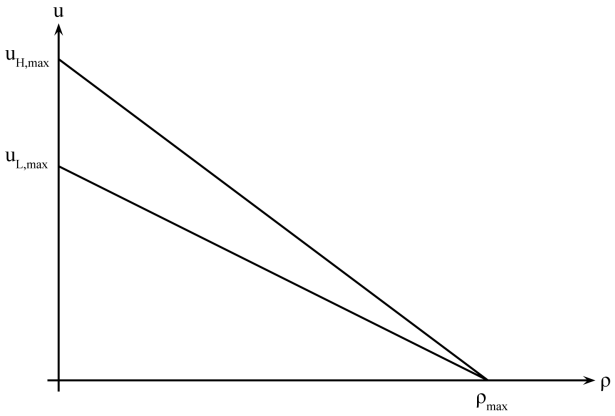


Figure 1. Benefit-loss-ratio (BLR).

3.2. Solution

The solution of the non-linear PDE Equation (13) determines the trading density and thus the price at all future times. Hence, I solve the following initial value problem

$$\frac{\partial \rho}{\partial t} + \frac{\mathrm{d}q}{\mathrm{d}\rho}\frac{\partial \rho}{\partial p} = 0, \quad s.t. \quad \rho(p,t) = 0. \tag{14}$$

This problem requires a linear approximation. Suppose the density is uniform, then it can be approximated as

$$\rho(p,t) = \rho_0 + \epsilon\rho_1(p,t) \tag{15}$$

where $|\epsilon\rho_1| \ll \rho_0$, and $\epsilon\rho_1(p,t)$ is called the "perturbed" trading density. For $p = 0$, I obtain

$$\rho(0,t) = \rho_0 + \epsilon f(0) \tag{16}$$

where $f(0) = \rho_1(0,t)$. Differentiation of Equation (14) in respect to t, result in

$$\epsilon\frac{\partial \rho_1}{\partial t} + \frac{\mathrm{d}q}{\mathrm{d}\rho} * \epsilon\frac{\partial \rho_1}{\partial p} = 0. \tag{17}$$

Taylor-series approximation of the second term, yields

$$\frac{\mathrm{d}q}{\mathrm{d}\rho}\left[\rho_0 + \epsilon\rho_1[p,t]\right] \cong \frac{\mathrm{d}q(\rho_0)}{\mathrm{d}\rho} + \epsilon\rho_1\frac{\mathrm{d}^2 q(\rho_0)}{\mathrm{d}\rho^2} + |O|^2 \tag{18}$$

where $\frac{\mathrm{d}q}{\mathrm{d}\rho} =: c$. Now, I have transformed the non-linear PDE into a linear-homogenous partial differential equation of the form:

$$\frac{\partial \rho}{\partial t} + c * \frac{\partial \rho}{\partial p} = 0. \tag{19}$$

I solve this PDE via variable transformation. Use both $p' := p - ct$ and $t' := t$, and rewrite it in the following terms:

$$\begin{aligned} \frac{\partial}{\partial p} &= \frac{\partial p'}{\partial p} * \frac{\partial}{\partial p'} + \frac{\partial t'}{\partial p} * \frac{\partial}{\partial t'} = \frac{\partial}{\partial p'}, \\ \frac{\partial}{\partial t} &= \frac{\partial p'}{\partial t} * \frac{\partial}{\partial p'} + \frac{\partial t'}{\partial t} * \frac{\partial}{\partial t'} = \frac{\partial}{\partial t'} - c * \frac{\partial}{\partial p'}. \end{aligned} \tag{20}$$

Substitute both findings in Equation (19), yields finally

$$\frac{\partial \rho}{\partial t'} = 0. \tag{21}$$

Proposition 1. The general solution of the PDE in Equation (19) and (21), is given as $\rho = g(p,t) = g(p-ct)$.

Proof. First, compute

$$\frac{\partial \rho}{\partial p} = \frac{\mathrm{d}g}{\mathrm{d}(p-ct)} * \frac{\partial (p-ct)}{\partial p} = \frac{\mathrm{d}g}{\mathrm{d}(p-ct)}$$

$$\frac{\partial \rho}{\partial t} = \frac{\mathrm{d}g}{\mathrm{d}(p-ct)} * \frac{\partial (p-ct)}{\partial t} = -c * \frac{\mathrm{d}g}{\mathrm{d}(p-ct)}$$

and then substitute both expressions in Equation (19): $-c * \frac{\mathrm{d}g}{\mathrm{d}(p-ct)} + c\frac{\mathrm{d}g}{\mathrm{d}(p-ct)} = 0$. □

3.3. Definition of a Financial Bubble

The solution of the model enables me to define a financial bubble in general.

Definition 4. A bubble is defined as a trading density greater than $\rho_{\mathrm{mac}}/2$ or accordingly by $c := \frac{\mathrm{d}q}{\mathrm{d}\rho} < 0$.

Figure 2 illustrates the definition graphically. Let me use this definition and discuss the implications.

Proposition 2. The existence of a bubble implies that the benefit-loss-field (BLF) is different to the specific BLR of an asset.

Proof. Due to $c < 0$. □

A negative slope of the BLF is a necessary and sufficient condition for a bubble in my model. In this constellation all traders buy (sell) the respective asset and that lowers the risk of the individual's BLF. However, an increasing trading density creates herd behaviour and finally a financial bubble in the overall market (**Figure 2**).

Proposition 3. In general, a bubble is defined by an upper limit for $\frac{\mathrm{d}q}{\mathrm{d}\rho} \le u(\rho)$. In other words, the left-hand side is less than the BLF $u(\rho)$.

Proof. If $q = \rho * u(p)$ then $\frac{\mathrm{d}q}{\mathrm{d}\rho} = \rho * \frac{\mathrm{d}u(\rho)}{\mathrm{d}\rho} + u(\rho)$ and therefore $\frac{\mathrm{d}q}{\mathrm{d}\rho} \le u$. □

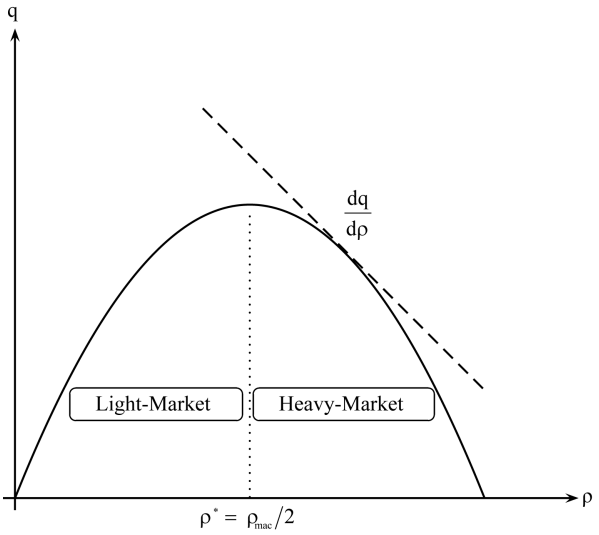


Figure 2. Shock wave of a "financial bubble".

This proposition demonstrates that a bubble can be characterized as a social interaction problem. Hence, a financial bubble is not just determined by individual rationality or irrationality as assumed in economic models. On the contrary, more buyers automatically indicate a higher trading density and higher risk for an exuberance. But from the individual point of view it is the other way round. A high density imply a small price movement and thus a lower risk and higher willingness to buy (sell) assets. This interaction and finally the imbalance of both mechanisms trigger a financial bubble. This new insight is in line with findings in sociology and psychology, for instance in the social system theory [22], and thus underlines the interdisciplinary character of the model.

4. Extended Model with Jumps

Finally, I model the impact of "news" in financial markets with bubbles. I utilize the idea of a "jump-discontinuity". In this case, I obtain

$$N(t) = \int_{p_a}^{p_b} \rho(p,t)\,\mathrm{d}p. \tag{22}$$

This integral is well defined if even $\rho(p,t)$ has a jump-discontinuity. A jump can be interpreted as a trading stop or news. In connection with Equation (2), $q = \rho * u$, the derivation of Equation (22) to dt yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{p_{(a)}}^{p_{(b)}} \rho(p,t)\,\mathrm{d}p = \left[q(p_a,t) - \rho(p_b,t)\frac{\mathrm{d}p_a}{\mathrm{d}t}\right] - \left[q(p_a,t) - \rho(p_b,t)\frac{\mathrm{d}p_b}{\mathrm{d}t}\right]. \tag{23}$$

Let's assume that the trading density changes on both ends of the interval at the same amount, $\frac{\mathrm{d}p_a}{\mathrm{d}t} = \frac{\mathrm{d}p_b}{\mathrm{d}t} = \frac{\mathrm{d}p_S}{\mathrm{d}t}$. Set Equation (23) equal to zero, results in $\frac{\mathrm{d}p_s}{\mathrm{d}t}\left[\rho(p_a,t) - \rho(p_b,t)\right] = q(p_a,t) - q(p_b,t)$ and after a transformation, I obtain

$$\frac{\mathrm{d}p_s}{\mathrm{d}t} = \frac{\left[q(p_a,t) - q(p_b,t)\right]}{\left[\rho(p_a,t) - \rho(p_b,t)\right]} =: \frac{[q]}{[\rho]} \tag{24}$$

where $[q] := q(p_S^-,t) - q(p_S^+,t)$ and $[\rho] := \rho(p_S^-,t) - \rho(p_S^+,t)$. Substituting this back in Equation (23), together with Equation (2), yields

$$\frac{\mathrm{d}p_s}{\mathrm{d}t} = \frac{\left[q(p_a,t) - q(p_b,t)\right]}{\rho(p_a,t) - \rho(p_b,t)} = \frac{\rho_{\max} * u_{\max} - \rho_0 * u(\rho_0)}{\rho_{\max} - \rho_0}. \tag{25}$$

Proposition 4. A financial bubble is characterized by a negative BLF in respect of time: $\frac{\mathrm{d}p_s(t)}{\mathrm{d}t} < 0$.

Proof. Use the properties of the function $u_{\max} = 0$ together with Equation (25). Compute the first derivative, I obtain $\frac{\mathrm{d}p_s(t)}{\mathrm{d}t} = -\frac{\rho_0 u(\rho_0)}{\rho_{\max} - \rho_0} < 0$. □

Consequently, the proposition confirms that good news imply a jump with a decline in the change of asset prices (BLR) and thus a lower probability of a bubble burst. Even if the result seems to be counterintuitive at first, it has a rational. News reduce asymmetric information and thus lowers the BLF. This reduces the market reaction time. In other words, the bubble may burst in case of small price jumps. Consequently, efficient markets with many news may trigger only small financial bubbles, but inefficient markets (=high asymmetric information) with little news may trigger huge financial bubbles. The solution of the ODE (Equation (25)), demonstrates the evolution of the asset price with news and it yields

$$p_S = -\frac{\rho_0 * u(\rho_0)}{\rho_{\max} - \rho_0} * t. \tag{26}$$

This equation explains the fact that with increasing time, t, the asset price declines automatically due to the lag of news. Therefore, more agents enter the market and buy/sell assets. In the end, this leads to herd behaviour and a financial bubble.

5. Conclusion

This paper attempts to explain the major properties and developments of financial bubbles from an econophysics point of view. I utilize a new theoretical model to detect ex ante financial bubbles. This model makes the complexity tractable without stochastic and martingale theory. In general, this approach has several advantages. First, it is a simple model and uniquely defines a financial bubble. Second, the model is general and thus explains the origins and developments of bubbles, too. Third, the model contains novel implications for empirical studies on financial bubbles. Finally, and most importantly, this model enables you to study the effect of news on financial bubbles for the first time.

Acknowledgements

I thank the editor and two anonymous referees for their helpful comments and suggestions. Research is funded by the RRI—Reutlingen Research Institute. This support is greatly appreciated.

References

[1] Protter, P. (2013) A Mathematical Theory of Financial Bubbles. *Paris-Princeton Lectures on Mathematical Finance*, **2081**, 1-108.

[2] Pindyck, R.S. and Wang, N. (2009) The Economic and Policy Consequences of Catastrophes. MIT Sloan School Working Paper No. 4751-09, 1-26.

[3] Barro, R.J. (2009) Rare Disasters and Asset Markets in the Twentieth Century. *American Economic Review*, **99**, 243-264. http://dx.doi.org/10.1257/aer.99.1.243

[4] Bikhchandani, S. and Sharma, S. (2000) Herd Behavior in Financial Markets: A Review. IMF Working Paper No. 00/48.

[5] Diamond, D. and Dybvig, P. (1983) Bank Runs, Deposit Insurance, and Liquidity. *Journal of Political Economy*, **91**, 401-409. http://dx.doi.org/10.1086/261155

[6] DeLong, J.B., Schleifer, A., Summers, L. and Waldman, R. (1990) Positive Feedback Investment Strategies and Destabilizing Rational Speculation. *Journal of Finance*, **45**, 379-395. http://dx.doi.org/10.2307/2328662

[7] Froot, K., Scharfstein, D. and Stein, J. (1992) Herd on the Street: Informational Efficiencies in a Market with Short-Term Speculation. *Journal of Finance*, **47**, 1461-1484. http://dx.doi.org/10.1111/j.1540-6261.1992.tb04665.x

[8] Lux, T. and Marchesi, M. (1999) Scaling and Criticality in a Stochastic Multiagent Model of a Financial Market. *Nature*, **397**, 498-500. http://dx.doi.org/10.1038/17290

[9] Banerjee, A. (1992) A Simple Model of Herd Behavior. *Quarterly Journal of Economics*, **107**, 797-818. http://dx.doi.org/10.2307/2118364

[10] Bikhchandani, S., Hirshleifer, D. and Welch, I. (1992) A Theory of Fads, Fashion, Custom and Cultural Change as Informational Cascades. *Journal of Political Economy*, **100**, 992-1026. http://dx.doi.org/10.1086/261849

[11] Bikhchandani, S., Hirshleifer, D. and Welch, I. (1998) Learning from the Behavior of Others: Conformity, Fads and Informational Cascades. *Journal of Economic Perspectives*, **12**, 151-170. http://dx.doi.org/10.1257/jep.12.3.151

[12] Devenov, A. and Welch, I. (1996) Rational Herding in Financial Economics. *European Economic Review*, **40**, 603-615. http://dx.doi.org/10.1016/0014-2921(95)00073-9

[13] Calvo, G. and Mendoza, E. (1998) Rational Herd Behavior and Globalization of Securities Markets. Mimeo, University of Maryland, College Park.

[14] Maug, E. and Naik, N. (1996) Herding and Delegation Portfolio Management. Mimeo, London Business School, London.

[15] Scharfstein, D. and Stein, J. (1990) Herd Behavior and Investment. *American Economic Review*, **80**, 465-479.

[16] Wemers, R. (1999) Mutual Fund Herding and the Impact on Stock Prices. *Journal of Finance*, **54**, 581-622. http://dx.doi.org/10.1111/0022-1082.00118

[17] Jain, A.K. and Gupta, S. (1987) Some Evidence on "Herding" Behavior of US Banks. *Journal of Money, Credit and Banking*, **19**, 78-89. http://dx.doi.org/10.2307/1992247

[18] Teh, L.L. and de Bondt, W.F.M. (1997) Herding Behavior and Stock Returns: An Exploratory Investigation. *Swiss Journal of Economics and Statistics*, **133**, 293-323.

[19] Graham, J.R. (1999) Herding among Investment Newsletters: Theory and Evidence. *Journal of Finance*, **54**, 237-269. http://dx.doi.org/10.1111/0022-1082.00103

[20] Galbraith, J.K. (1997) Savings, Investment, and Functional Efficiency: A Comparative Examination of National Financial Complexes: Comment. Chapter Comment, University of Michigan Press, Ann Arbor, 305-308.

[21] Fransman, M. (2004) The Telecoms Boom and Bust 1996-2003 and the Role of Financial Markets. *Journal of Evolutionary Economics*, **14**, 369-406.

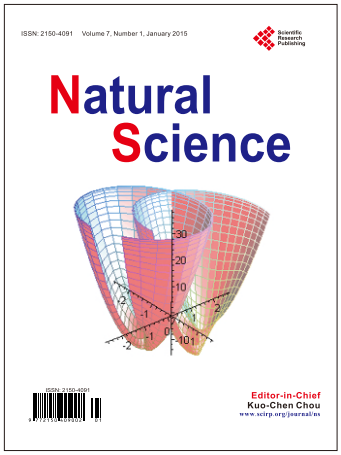
[22] Luhmann, N. (1994) Die Wirtschaft der Gesellschaft (suhrkamp taschenbuch wissenschaft). Suhrkamp Verlag, Berlin.

[23] McCauley, J.L. (2004) Dynamics of Markets: Econophysics and Finance. Cambridge University Press, Cambridge. http://dx.doi.org/10.1017/CBO9780511606588

Call for Papers

Natural Science

A Journal Published by Scientific Research Publishing, USA
www.scirp.org/journal/ns



Editor-in-Chief

Editorial Advisory Board

Editorial Board

Natural Science is an international journal dedicated to the latest advancement of natural sciences. The goal of this journal is to provide a platform for scientists and academicians all over the world to promote, share, and discuss various new issues and developments in different areas of natural sciences. All manuscripts must be prepared in English, and are subject to a rigorous and fair peer-review process. Accepted papers will immediately appear online followed by printed hard copy. The journal publishes original papers including but not limited to the following fields:

- **Astronomy & Space Sciences**
  - Astronomy
  - Astrophysics
  - Atmospheric Science
  - Space Physics
- **Earth Science**
  - Geography
  - Geology
  - Geophysics/Geochemistry
  - Oceanography
- **Chemistry**
  - Analytical Chemistry
  - Biochemistry
  - Computational Chemistry
  - Inorganic Chemistry
  - Organic Chemistry
  - Physical Chemistry
- **Life Science**
  - Cell Biology
  - Computational Biology
  - Genetics
  - Immunology
  - Medicine/Diseases
  - Microbiology
  - Molecular Biology
  - Neuroscience
  - Pharmacology/Toxicology
  - Physiology
  - Psychology
  - Virology
- **Physics**
  - Applied Physics
  - Atomic, Molecular, and Optical Physics
  - Biophysics
  - High Energy/Particle Physics
  - Material Science
  - Plasma Physics
- **Others**
  - Education
  - History of Science
  - Science and Innovations

We are also interested in: 1) Short Reports—2-5 page papers where an author can either present an idea with theoretical background but has not yet completed the research needed for a complete paper or preliminary data; 2) Book Reviews—Comments and critiques.

Notes for Intending Authors

Submitted papers should not be previously published nor be currently under consideration for publication elsewhere. Paper submission will be handled electronically through the website. For more details, please access the website.

Website and E-Mail

http://www.scirp.org/journal/ns E-mail: ns@scirp.org

What is SCIRP?

Scientific Research Publishing (SCIRP) is one of the largest Open Access journal publishers. It is currently publishing more than 200 open access, online, peer-reviewed journals covering a wide range of academic disciplines. SCIRP serves the worldwide academic communities and contributes to the progress and application of science with its publication.

What is Open Access?

All original research papers published by SCIRP are made freely and permanently accessible online immediately upon publication. To be able to provide open access journals, SCIRP defrays operation costs from authors and subscription charges only for its printed version. Open access publishing allows an immediate, worldwide, barrier-free, open access to the full text of research papers, which is in the best interests of the scientific community.

- High visibility for maximum global exposure with open access publishing model
- Rigorous peer review of research papers
- Prompt faster publication with less cost
- Guaranteed targeted, multidisciplinary audience

Scientific Research Publishing

Website: http://www.scirp.org

Subscription: sub@scirp.org

Advertisement: service@scirp.org