VOLUME 10, Nº 03
2 0 0 8

Cubo

A Mathematical Journal

Universidad de La Frontera
Facultad de Ingeniería, Ciencias y Administración
Departamento de Matemática y Estadística
Temuco - Chile

Universidade Federal de Pernambuco
Centro de Ciências Exatas e da Natureza
Departamento de Matemática
Recife - Brazil

CUBO
A Mathematical Journal

The Department of Mathematics of the Universidad de La Frontera Temuco-Chile and the Department of Mathematics of the Universidade Federal de Pernambuco, Recife-Brazil, edit **CUBO,** A Mathematical Journal. **CUBO** appears in three issues per year and is indexed in ZentralBlatt Math., Mathematical Reviews, MathSciNet and Latin Index. **CUBO** publishes original research papers, preferably not longer than 20 printed pages, that contain substantial mathematical results in all branches of pure and applied mathematics.

CUBO (I.S.S.N. 0716.7776) Editorial Office : Departamento de Matemática, Universidade Federal de Pernambuco, CEP 50740-540, Recife-PE., Brazil, cch@dmat.ufpe.br.

CUBO
A MATHEMATICA JOURNAL
Universidad de La Frontera
Universidade Federal de Pernambuco
Volume 10/N<u>o</u>03 – OCTOBER 2008

CONTENTS

CUBO

A Mathematical Journal

EDITORIAL BOARD

CUBO
A MATHEMATICA JOURNAL
Universidad de La Frontera
Universidade Federal de Pernambuco
Volume 10/N$^{\underline{o}}$03 – OCTOBER 2008

SUMMARY

CUBO A Mathematical Journal
Vol.10, $N^{\underline{o}}$03, (1–12). October 2008

Proximal-Resolvent Methods for Mixed Variational Inequalities

Muhammad Aslam Noor and Khalida Inayat Noor
Mathematics Department,
COMSATS Institute of Information Technology, Islamabad, Pakistan
email: noormaslam@hotmail.com
email: khalidanoor@hotmail.com

ABSTRACT

It is well-known that the mixed variational inequalities are equivalent to the fixed point problem. We use this alternative equivalent formulation to suggest and analyze some new proximal resolvent methods for solving mixed variational inequalities. We also study the convergence of these new methods under some mild conditions. These new iterative methods include the projection, extragradient and proximal methods as special cases. Results obtained in this paper represent a refinement and improvement of the previously known results.

RESUMEN

Es bien conocido que las desigualdades variacionales mescladas son equivalentes a problemas de punto fijo. Nosotros usamos esta formulación alternativa equivalente para sugerir y analizar nuevos métodos resolventes proximales para resolver desigualdes variacionales mesclasdas. También estudiamos la convergencia de estos nuevos métodos bajo algunas condiciones blandas. Estos nuevos métodos iterativos incluyen como casos especiales la prejección, métodos extragradiente y proximales. Los resultados en este artículo representam un refinamiento y perfeccionamiento de resultados previamente conocidos.

Key words and phrases: *Variational inequalities, resolvent method, fixed point, proximal methods, convergence.*

Math. Subj. Class.: *49J40, 90C30.*

1. Introduction

Variational inequalities, which were introduced and considered by Stampacchia [26] in 1964, have had a great impact and influence in the development of almost all branches of pure and applied sciences. It has been shown that the variational inequalities provide a simple, unified, natural, novel and general framework to study a wide class of problems arising in various branches of pure and applied sciences. The ideas and techniques of variational inequalities are being used in a variety of diverse fields and proved to innovative and productive, see [1-26] and the references therein. In recent years, variational inequalities have been extended and generalized in several directions. A useful and important generalization of variational inequalities is called the mixed variational inequality or variational inequality of the second kind containing the nonlinear term. Due to the presence of the nonlinear term, the projection method and its variant forms including the Wiener-Hopf equations can not be extended for solving the mixed variational inequality. To overcome these drawbacks, some iterative methods have been developed and investigated for solving mixed variational inequalities using the technique of auxiliary principle technique, the origin of which can be traced back to Lions and Stampacchia [10] and Glowinski, Lions and Tremolieres [7]. This technique has been used by several researchers to develop implicit and explicit methods for solving the mixed variational inequalities and the equilibrium problems, see [14-23] and the references therein. We would like to mention that, if the nonlinear term in the mixed variational inequalities is a proper, convex and lower-semicontinuous, then it has been shown [14] that the mixed variational inequalities are equivalent to the fixed point problem. This alternative equivalent formulation has been used to suggest and analyze several iterative methods for solving the mixed variational inequalities. The convergence of these resolvent iterative methods requires that the underlying operator is strong monotone and Lipschitz continuous. Secondly it is very difficult to evaluate the resolvent of the operator. These facts have motived to modify the resolvent iterative method. Noor [16-20] used the technique of updating the solution to suggest and analyze some modified extraresolvent type method. The extraresolvent method overcomes this difficulty by using the technique of updating the solution, which modified the resolvent method by performing additional step and resolvent at each step according to double resolvent formula. It is worth mentioning that the convergence of the extraresolvent method requires that the solution exists and the operator to be monotone and Lipschitz continuous. When the operator is not Lipschitz continuous or when the Lipschitz continuous is not known, the extraresolvent method and its variant forms require an Armijo-like line search procedure to compute the step size with a new resolvent needed for each trial, which leads to expensive computation. To overcome these draw backs, many authors have suggested and proposed some modified methods for solving mixed variational inequalities. We also note that if the nonlinear term involving the mixed variational inequalities is an indicator function of a convex set in the Hilbert space, then the mixed variational inequalities are equivalent to the classical variational inequalities. He at el. [9] and Noor [19] have considered a class of modified proximal-extragradient methods for solving the classical variational inequalities, which uses a better step-size rule (inexactness criteria) and includes the proximal and the extragradient methods as a special cases. They have shown the convergence of this approximate proximal method requires either monotonicity or pseudomonotonicity. It has been shown [9] that these proximal-extragradient methods are numerically efficient and robust. It is worth mentioning that there are no such methods for solving the mixed variational inequalities. Inspired and motivated by the research going in this dynamic field, we suggest some new proximal-resolvent methods for solving the mixed variational inequalities. We show that the convergence of our methods requires only the pseudomonotonicity, which is a weaker condition than monotonicity. Results obtained in this paper include the results of He et al [9] and Noor [19] as special cases and improve the convergence criteria of methods of He et al [9]. Our results can also be viewed as a significant extension and generalization of the previously known methods for solving the mixed variational inequalities and related optimization problems.

2. Formulation

Let K be a nonempty closed and convex set in a real Hilbert space H, whose inner product and norm are denoted by $\langle\cdot,\cdot\rangle$ and $\|.\|$ respectively. Let $T : H \longrightarrow H$ be a nonlinear operator and S be a nonexpansive operator. Let P_K be the projection of H onto the convex set K. Let $\varphi : H \longrightarrow R \cup \{\infty\}$ be a continuous function. It is well known that the subdifferential $\partial\varphi(.)$ of a proper, convex and lower-semicontinuous function φ is a maximal monotone operator.

We consider the problem of finding $u \in H$ such that

$$\langle Tu, v-u\rangle + \varphi(v) - \varphi(u) \geq 0, \quad \forall v \in H, \tag{1}$$

which is known as the mixed variational inequality introduced or variational inequality of the second type, see Glowinski, Lions and Tremolieres [7] and Lions and Stampacchia [10]..

We note that, if the function φ in the mixed variational inequality is a proper, convex and lower-semicontinuous, then problem (1) is equivalent to finding $u \in H$ such that

$$0 \in Tu + \partial\varphi(u), \tag{2}$$

which is known as the problem of finding a zero of sum of two (or more) monotone operators. Here $\partial\varphi$ is the subdifferential of the function φ. It is well known that a large class of problems arising in industry, ecology, finance, economics, transportation, network analysis and optimization can be formulated and studied in the framework of (1) and (2), see [3-6, 15-24] and the references therein.

If φ is an indicator function of a closed convex set K in H, that is,

$$\varphi(u) = I_K(v) = \begin{cases} 0, & \text{if} \quad v \in K; \\ +\infty, & \text{otherwise.} \end{cases}$$

then the mixed variational inequalities (1) are equivalent to finding $u \in K$ such that

$$\langle Tu, v-u\rangle \geq 0, \quad \forall v \in K, \tag{3}$$

which is known as the classical variational inequality, introduced and studied by Stampacchia [26] in 1964. For the numerical methods, formulations and applications of the mixed variational inequalities, readers are advised to see [1-25] and the references therein.

We now recall some well known concepts and results.

Definition 2.1[3]. For any maximal operator T, the resolvent operator associated with T, for any $\rho > 0$, is defined as

$$J_T(u) = (I+\rho T)^{-1}(u), \qquad \forall u \in H.$$

It is well known that an operator T is maximal monotone if and only if its resolvent operator J_T is defined everywhere. It is single-valued and nonexpansive. that is,

$$\|J_T u - J_T v\| \leq \|u-v\|, \quad \forall u, v \in H.$$

If $\varphi(.)$ is a proper, convex and lower-semicontinuous function, then its subdifferential $\partial\varphi(.)$ is a maximal monotone operator. In this case, we can define the resolvent operator

$$J_\varphi(u) = (I+\rho T)^{-1}(u), \quad \forall u \in H$$

associated with the subdifferential $\partial\varphi(.)$. The resolvent operator J_φ has the following useful characterization, see[3,20].

Lemma 2.1. For a given $z \in H$, $u \in H$ satisfies the inequality

$$\langle u-z, v-u\rangle + \rho\varphi(v) - \rho\varphi(u) \geq 0, \quad \forall v \in H \tag{4}$$

if and only if $u = J_\varphi(z$, where $J_\varphi = (I+\rho\partial\varphi)^{-1}$ is the resolvent operator.

It is well-known that the resolvent operator J_φ is a nonexpansive operator, that is,

$$\|J_\varphi(u) - J_\varphi(v)\| \leq \|u-v\|, \quad \forall u, v \in H.$$

Lemma 2.1 plays a very important and significant role in the analysis of the mixed variational inequalities. If the proper, convex and semi-lowercontinuous function φ is an indicator function of a closed convex set K in H, then $J_\varphi \equiv P_K$, is the projection operator from H onto the closed convex set K. In this case, Lemma 2.1 reduces to the following well known result, which is known as the projection Lemma.

Lemma 2.2 . Let K be a closed convex set K in H. Then, for a given $z \in H$, $u \in K$ satisfies the inequality

$$\langle u - z, v - u \rangle \geq 0, \quad \forall v \in K,$$

if and only if

$$u = P_K z,$$

where P_K is the projection of H onto the closed convex set K. It is also known that the projection operator P_K is nonexpansive. For the applications of Lemma 2.2, see [1-25].

Definition 2.2. $\forall u, v \in H$, the operator $T : H \longrightarrow H$ with respect the function φ is said to be *pseudomonotone,* if

$$\langle Tu, v - u \rangle + \varphi(v) - \varphi(u) \geq 0 \quad \text{implies} \quad \langle Tv, v - u \rangle + \varphi(v) - \varphi(u) \geq 0.$$

Note that monotonicity implies pseudomonotonicity but the converse is not true [5].

3. Main results

In this section, we use the projection technique to suggest some iterative methods for solving the variational inequalities. For this purpose, we need the following result, which can be proved by invoking Lemma 2.1.

Lemma 3.1. The function $u \in H$ is a solution of the mixed variational inequality (1) if and only if $u \in H$ satisfies the relation

$$u = J_\varphi[u - \rho Tu], \tag{5}$$

where $\rho > 0$ is a constant and $J_\varphi u = (I + \rho\partial\varphi)^{-1}(u)$ is the resolvent operator.

Lemma 3.1 implies that problems (1) and (5) are equivalent. This alternative formulation is very important from the numerical analysis point of view and has played a significant part in suggesting several numerical methods for solving variational inequalities and complementarity problems, see [1-7,10-20].

We now define the projection residue vector by the relation

$$R(u) = u - J_\varphi[u - \rho Tu] = u - y, \quad y = J_\varphi[u - \rho Tu].$$

Invoking Lemma 3.1, one can easily show that $u \in H$ is a solution of (1) if and only if $u \in H$ is a zero of the equation

$$R(u) = 0.$$

For a positive constant α, we consider

$$u = u - \alpha R(u) = u - \alpha\{u - J_{\varphi}[u - \rho Tu]\},$$

which is another fixed point problem. We use alternative fixed point formulation to suggest and analyze the following iterative method for solving the mixed variational inequality (1).

Algorithm 3.1. For a given $u_0 \in H$, compute the approximate solution u_{n+1} by the iterative scheme

$$\begin{aligned} u_{n+1} &= J_{\varphi}[u_n - \gamma_n R(u_{n+1})] \\ &= J_{\varphi}[u_n - \gamma_n\{u_n - J_{\varphi}[u_n - \rho T u_{n+1}]\}], \quad n = 0, 1, 2, \ldots, \end{aligned}$$

or equivalently

$$\begin{aligned} y_n &= J_{\varphi}[u_n - \rho T u_{n+1}] \\ u_{n+1} &= J_{\varphi}[u_n - \gamma_n\{u_n - y_n\}], \qquad n = 0, 1, 2, \ldots \end{aligned}$$

which can be considered as a proximal point method and appears to be a new one.

If φ is the indicator function of a closed convex set K, then $J_{\varphi} \equiv P_K$, the projection of H onto K. Consequently, Algorithm 3.1 collapse to the following algorithm for solving classical variational inequalities (3).

Algorithm 3.2. For a given $u_0 \in H$, compute the approximate solution u_{n+1} by the iterative scheme

$$\begin{aligned} u_{n+1} &= P_K[u_n - \gamma_n R(u_{n+1})] \\ &= P_K[u_n - \gamma_n\{u_n - P_K[u_n - \rho T u_{n+1}]\}], \quad n = 0, 1, 2, \ldots, \end{aligned}$$

or equivalently

$$\begin{aligned} y_n &= P_K[u_n - \rho T u_{n+1}] \\ u_{n+1} &= P_K[u_n - \gamma_n\{u_n - y_n\}], \qquad n = 0, 1, 2, \ldots \end{aligned}$$

which can be considered as a proximal-extragradient method.

Note that for $\gamma_n = 1$, Algorithm 3.1 reduces to:

Algorithm 3.3. For a given $u_0 \in H$, compute the approximate solution u_{n+1} by the iterative scheme

$$u_{n+1} = J_{\varphi}[u_n - \rho T u_{n+1}], \quad n = 0, 1, 2 \ldots$$

which is known as the proximal method and convergence requires only pseudomonotonicity, see Noor [20]. In recent years, proximal methods have been considered and studied extensively. Several conditions have been studied which are easy to implement, see [9, 17-20].

We now use the technique of updating the solution to rewrite the fixed-point formulation (5) as:

$$\begin{aligned} y &= J_\varphi[u-\rho Tu] \\ u &= J_\varphi[y-\rho Ty], \end{aligned} \tag{6}$$

which can be written as

$$u = J_\varphi[J_\varphi[u-\rho Tu] - \rho T J_\varphi[u-\rho Tu]],$$

which is another fixed point formulation of the mixed variational inequalities (1). Here we use this equivalent alternative formulation to suggest the following method for solving mixed variational inequalities (1).

Algorithm 3.4. For a given $u_0 \in H$, find the approximate solution u_{n+1} by the iterative schemes:

$$\begin{aligned} y_n &= J_\varphi[u_n - \rho T u_{n+1}] \\ u_{n+1} &= J_\varphi[y_n - \rho T y_n], \quad n = 0,1,2,\ldots \end{aligned}$$

Algorithm 3.5. For a given $u_0 \in H$, find the approximate solution u_{n+1} by the iterative schemes:

$$u_{n+1} = J_\varphi[J_\varphi[u_n - \rho T u_{n+1}] - \rho T J_\varphi[u_n - \rho T u_{n+1}]], \quad n = 0,1,2,\ldots$$

Algorithms 3.4 and Algorithm 3.5 are called the two-step or predictor-corrector implicit iterative resolvent methods for solving the mixed variational inequalities (1) and appear to be new ones.

If φ is the indicator function of a closed convex set K, then Algorithm 3.5 is equivalent to the following implicit projection iterative method for solving the classical variational inequalities (3), which are mainly due to Noor [16-18].

Algorithm 3.6. For a given $u_0 \in H$, find the approximate solution u_{n+1} by the iterative schemes:

$$u_{n+1} = P_K[P_K[u_n - \rho T u_{n+1}] - \rho T P_K[u_n - \rho T u_{n+1}]], \quad n = 0,1,2,\ldots$$

Now we look at Algorithm 3.4 from a different angle. Consider y defined by (6) as an approximate solution of the mixed variational inequality (1) and define

$$\begin{aligned} w &= J_\varphi[u - \gamma(u-y)] \\ z &= u - \rho T w. \end{aligned}$$

We use this formulation to suggest the following iterative method

Algorithm 3.7. For a given $u_0 \in H$, calculate the approximate solution u_{n+1} by the iterative schemes;

$$\begin{aligned} y_n &= J_\varphi[u_n - \rho T u_{n+1}] \\ w_n &= J_\varphi[u_n - \gamma(u_n - y_n)] \\ u_{n+1} := z_n &= u_n - \rho T w_n, \qquad n = 0,1,2,\ldots \end{aligned}$$

which is called the modified extraresolvent method and appears to be a new one.

Note that for $\gamma = 1$, Algorithm 3.7 reduces to

Algorithm 3.8. For a given $u_0 \in H$, compute the approximate solution u_{n+1} by the iterative scheme

$$\begin{aligned} y_n &= J_\varphi[u_n - \rho T u_{n+1}] \\ u_{n+1} &= u_n - \rho T y_n, \quad n = 0, 1, 2, \ldots \end{aligned}$$

which is called the extraresolvent method for solving the mixed variational inequality (1).

For a positive constant α, consider

$$u = u - \alpha(u - z).$$

Here the positive constant α can be viewed as a step length along the direction $-(u - z)$.

We use this fixed-point formulation to suggest the following iterative method.

Algorithm 3.9. For a given $u_0 \in H$, compute the following iterative schemes:

$$\begin{aligned} y_n &= J_\varphi[u_n - \rho_n T u_{n+1}] \\ w_n &= J_\varphi[u_n - \gamma_n(u_n - y_n)] \\ z_n &= J_\varphi[u_n - \rho_n T w_n] && (7) \\ u_{n+1} &= u_n - \alpha(u_n - z_n), \qquad n = 0, 1, 2, \ldots && (8) \\ \alpha &= \frac{\|z_n - w_n\|^2 + \|u_n - z_n\|^2 - \triangle(w_n)}{2\|u_n - z_n\|^2} && (9) \end{aligned}$$

where

$$\begin{aligned} \triangle(w_n) &\leq \nu(\|z_n - w_n\|^2 + \|u_n - z_n\|^2), \qquad \nu < 1 \\ &= \nu\{2\langle w_n - z_n, w_n - u_n + \rho_n T w_n + \rho_n \varphi'(w_n)\rangle - \|w_n - z_n\|^2\}. && (10) \end{aligned}$$

Here $\triangle(w_n)$ is known as the inexactness criteria which can be viewed as stepsize and $\varphi'(.)$ is the differential of the convex function φ.

For $\alpha = 1$ and $z_n = w_n$, Algorithm 3.9 is exactly Algorithm 3.8. If $y = w$, then Algorithm 3.9 reduces to:

Algorithm 3.10. For a given $u_\in H$, compute the approximate solution u_{n+1} by the iterative schemes

$$\begin{aligned} y_n &= J_\varphi[u_n - \rho_n T u_{n+1}] \\ w_n &= J_\varphi[u_n - \gamma(u_n - y_n)] \\ u_{n+1} := z_n &= u_n - \alpha(u_n - w_n), \qquad n = 0, 1, 2, \ldots \\ \alpha &= \frac{\|u_n - y_n\|^2 + \|u_n - w_n\|^2 - \triangle(y_n)}{2\|u_n - w_n\|^2} \end{aligned}$$

which is an approximate extraresolvent method for solving (1).

Remark 3.1. Algorithms 3.5-3.10 are called the approximate proximal extraresolvent methods, which are new ones. We would like to point out that if the nonlinear term φ in the mixed variational inequality (1) is an indicator function of a closed convex set K, then the resolvent $J_\varphi = P_K$ is the projection operator of H onto the closed convex set K. Consequently, Algorithms 3.1-3.10 reduce to Algorithms for variational inequalities (3) which appear to be new ones for the variational inequalities (3). In a similar way, one can obtain several new and known algorithms as special cases of Algorithm 3.9. This shows that Algorithm 3.9 is more flexible and unifies several recently proposed (implicit or explicit) algorithms for solving the mixed variational inequalities.

We now study the convergence analysis of Algorithm 3.9. The analysis is in the spirit of He, Yang and Yuan [9] and Noor [19]. To convey the idea and for the sake of completeness, we include the details.

Theorem 3.1. Let the operator T be pseudomonotone. If $u \in K$ be a solution of the mixed variational inequality (1) and u_{n+1} be the approximate solution obtained from Algorithm 3.9, then

$$\|u_{n+1}(\alpha) - u\|^2 \leq \|u_n - u\|^2 - \frac{(1-\nu)^2}{4}\{\|u_n - w_n\|^2 + \|u_n - z_n\|^2\}. \tag{11}$$

Proof. Let $u \in K$ be a solution of (1). Then

$$\langle Tu, v-u\rangle + \varphi(v) - \varphi(u) \geq 0, \quad \forall v \in K,$$

implies that

$$\langle Tv, v-u\rangle + \varphi(v) - \varphi(u) \geq 0, \tag{12}$$

since T is pseudomonotone.

Taking $v = w_n$ in (12), we have

$$\langle Tw_n, w_n - u\rangle + \varphi(w_n) - \varphi(u) \geq 0,$$

which can be written as

$$\langle Tw_n, z_n - u\rangle \geq \langle Tw_n, z_n - w_n\rangle + \varphi(u) - \varphi(w_n). \tag{13}$$

Taking $z = [u_n - \rho_n Tw_n], \quad u = z_n$ and $v = u$ in (4), we have

$$\langle u_n - \rho_n Tw_n - z_n, u_n - u\rangle + \rho_n\varphi(u) - \rho_n\varphi(z_n) \geq 0,$$

from which we have

$$\langle u_n - z_n, u_n - u\rangle \geq \langle u_n - u, \rho_n Tw_n\rangle + \rho_n\varphi(z_n) - \rho_n\varphi(u). \tag{14}$$

From (13) and (14), we have

$$\begin{aligned}\langle u_n - z_n, z_n - w_n\rangle &\geq \langle \rho_n Tw_n, z_n - w_n\rangle + \rho_n(\varphi(z_n) - \varphi(w_n)) \\ &\geq \rho_n\langle Tw_n + \varphi'(w_n), z_n - w_n\rangle. \end{aligned} \tag{15}$$

Consider

$$\begin{aligned}\|u_n-u\|^2 &- \|u_{n+1}(\alpha)-u\|^2 = \|u_n-u\|^2-\|u_n-\alpha(u_n-z_n)-u\|^2 \\ &\geq \|u_n-u\|^2-\|u_n-u-\alpha(u_n-z_n)\|^2 \\ &= 2\alpha\langle u_n-u,u_n-z_n\rangle-\alpha^2\|u_n-z_n\|^2 \\ &= 2\alpha\|u_n-z_n\|^2+2\alpha\langle z_n-u,u_n-z_n\rangle-\alpha^2\|u_n-z_n\|^2. \qquad (16)\end{aligned}$$

Combining (10), (15) and (16), we obtain

$$\|u_n-u\|^2-\|u_{n+1}(\alpha)-u\|^2 \geq \alpha\{\|z_n-w_n\|^2+\|u_n-z_n\|^2-\triangle(w_n)\}-\alpha^2\|u_n-z_n\|^2, \qquad (17)$$

which is a quadratic in α and has a maximum at

$$\alpha^*=\frac{\|z_n-w_n\|^2+\|u_n-z_n\|^2-\triangle(w_n)}{2\|u_n-w_n\|^2}. \qquad (18)$$

From (10), (17) and (18), we have the required result (11). □

Theorem 3.2. Let H be a finite dimensional subspace. If $u\in K$ be a solution of (1) and u_{n+1} be the approximate solution obtained from Algorithm 3.9, then $\lim_{n\longrightarrow\infty}(u_n)=u$.

Proof. Let $u\in H$ be a solution of (1). From (11), it follows that the sequence $\{\|u-u_n\|\}$ is nonincreasing and consequently $\{u_n\}$ is bounded. Furthermore, we have

$$\sum_{n=1}^{\infty}\frac{(1-\nu)^2}{4}\{\|z_n-w_n\|^2+\|u_n-z_n\|^2\}\leq\|u_0-u\|^2,$$

which implies that

$$\lim_{n\longrightarrow\infty}\|z_n-w_n\| = 0 \qquad (19)$$

$$\lim_{n\longrightarrow\infty}\|u_n-z_n\| = 0. \qquad (20)$$

Thus we see that the sequences $\{w_n\}$ and $\{z_n\}$ are also bounded. Also from (19) and (20), we have

$$\begin{aligned}\|R(w_n)\| &= \|w_n-J_\varphi[w_n-\rho Tw_n]\| = \|w_n-z_n+z_n-J_\varphi[w_n-\rho Tw_n]\| \\ &\leq \|w_n-z_n\|+\|J_\varphi[u_n-\rho Tw_n]-J_\varphi[w_n-\rho Tw_n]\| \\ &\leq \|w_n-z_n\|+\|u_n-w_n\|=0.\end{aligned}$$

Thus

$$\lim_{n\longrightarrow\infty}R(w_n)=0. \qquad (21)$$

Let $\hat{u}$ be a cluster point of $\{w_n\}$ and the subsequence $\{w_{n_i}\}$ converges to $\hat{u}$. Since $R(u)$ is a continuous function of u, it follows that

$$\lim_{n\longrightarrow\infty}R(w_{n_i})=R(\hat{u})=0,$$

which shows that $\hat{u}$ is a solution of the mixed variational inequality (1). From (19) and (20), we know that $\lim_{n\longrightarrow\infty}(y_{n_i}) = \hat{u} = \lim_{n\longrightarrow\infty}(z_{n_i})$. Hence from (11), we have

$$\|u_{n+1} - \hat{u}\|^2 \leq \|u_n - \hat{u}\|^2, \qquad \forall n \geq 0,$$

which shows that the sequence $\{u_n\}$ converges to $\hat{u}$, the required result. □

Conclusion. In this paper, we have suggested and analyzed some new proximal extraresolvent methods for pseudomonotone mixed variational inequalities and related optimization problems. The convergence of the new methods require only the pseudomonotonicity of the operator, which is a weaker condition than monotonicity. It has been shown [9] that a special case of Algorithm 3.9 is numerically efficient and robust in solving traffic equilibrium problems. The results obtained are encouraging. The comparison of these new methods with the other methods is an interesting open problem for further research.

Acknowledgement. I wish to express my deepest gratitude to Prof. Dr. Claudio Cuevas, Editor-in-Chief, Cubo Journal, for this invitation. I am also grateful to Dr. S. M. Junaid Zaidi, Rector, CIIT, for the excellent research facilities.

Received: January 2008. Revised: April 2008.

References

[1] A. Bnouhachem, *A self-adaptive method for mixed quasi variational inequality*, J. Math. Anal. Appl., **309** (2005), 136–150.

[2] A. Bnouhachem, M. Aslam Noor and T. M. Rassias, *Three-step iterative algorithms for mixed variational inequalities*, Appl. Math. Computation, **183** (2006), 436–446.

[3] H. Brezis, *Operateurs Maximaux Monotone et Semigroupes de Contractions dans les Espace d'Hilbert*, North-Holland, Amsterdam, Holland, 1973.

[4] P. Daniele, F. Giannessi and A. Maugeri, *Equilibrium Problems and Variational Models*, Kluwer Academic Publishers, United Kingdom, 2003.

[5] F. Giannessi and A. Maugeri, *Variational Inequalities and Network Equilibrium Problems*, Plenum Press, New York, 1995.

[6] F. Giannessi, A. Maugeri and P. M. Pardalos, *Equilibrium Problems, Nonsmooth Optimization and Variational Inequalities Problems*, Kluwer Academic Publishers, Dordrecht Holland, 2001.

[7] R. Glowinski, J.L. Lions and R. Tremolieres, *Numerical Analysis of Variational Inequalities*, North-Holland, Amsterdam, Holland 1981.

[8] B.S. He and L.Z. Liao, *Improvement of some projection methods for monotone nonlinear variational inequalities*, J. Optim. Theory Appl., **112** (2002), 111–128.

[9] B.S. He, Z. Yang and X.M. Yuan, *An approximate proximal-extragradient method for monotone variational inequalities*, J. Math. Anal. Appl., **300** (2004), 362–374.

[10] J.L. Lions and G. Stampacchia, *Variational inequalities*, Commun. Pure Appl. Math., **20** (1967), 493–512.

[11] U. Mosco, *Implicit variational methods and quasi variational inequalities, in: Nonlinear Operators and the Calculus of Variations*, Lecture Notes in Mathematics, Springer-Verlag, Berlin, Germany, **543** (1976), 83–126.

[12] M. Aslam Noor, *On Variational Inequalities*, PhD Thesis, Brunel University, London, UK, 1975.

[13] M. Aslam Noor, *Some recent advances in variational inequalities, Part I, basic concepts*, New Zealand J. Math., **26** (1997), 53–80.

[14] M. Aslam Noor, *Some recent advances in variational inequalities, Part II, other concepts*, New Zealand J. Math., **26** (1997), 229–255.

[15] M. Aslam Noor, *Some algorithms for general monotone mixed variational inequalities*, Math. Comput. Modell., **29** (1999), 1–9.

[16] M. Aslam Noor, *New approximation schemes for general variational inequalities*, J. Math. Anal. Appl., **251** (2000), 217–229.

[17] M. Aslam Noor, *New extragradient-type methods for general variational inequalities*, J. Math. Anal. Appl., **277** (2003), 379–395.

[18] M. Aslam Noor, *Some developments in general variational inequalities*, Appl. Math. Computation, **152** (2004), 199–277.

[19] M. Aslam Noor, *Projection-proximal methods for general variational inequalities*, J. Math. Anal. Appl., **318** (2006), 53–62.

[20] M. Aslam Noor, *Fundamentals of mixed quasi variational inequalities*, Int. J. Pure Appl. Math., **15** (2004), 137–258.

[21] M. Aslam Noor and A. Bnouhachem, *Proximal extragradient methods for pseudomonotone variational inequalities*, Tamkang J. Math., **37** (2006), 109–116.

[22] M. Aslam Noor, K. Inayat Noor and T.M. Rassias, *Some aspects of variational inequalities*, J. Comput. Appl. Math., **47** (1993), 285–312.

[23] M. Aslam Noor, K. Inayat Noor and T.M. Rassias, *Set-valued resolvent equations and mixed variational inequalities*, J. Math. Anal. Appl., **220** (1998), 741–759.

[24] M. Patriksson, *Nonlinear Programming and Variational Inequalities: A Unified Approach*, Kluwer Academic Publishers, Dordrecht, 1998.

[25] M.V. Solodov and B.F. Svaiter, *Error bounds for proximal point subproblems and associated inexact proximal point algorithm*, Math. Program., **88** (2000),371–389.

[26] G. Stampacchia, *Formes bilineaires coercivities sur les ensembles coercivities sur les ensembles convexes*, C. R. Acad. Sci. Paris, **258** (1964), 4413–4416.

CUBO A Mathematical Journal
Vol.10, N$^{\underline{o}}$03, (13–20). October 2008

Poincaré Type Inequalities for Linear Differential Operators

George A. Anastassiou
Department of Mathematical Sciences,
University of Memphis,
Memphis, TN 38152 U.S.A.
email: ganastss@memphis.edu

ABSTRACT

Various L_p form Poincaré type inequalities [1], forward and reverse, are given for a Linear Differential Operator L, involving its related initial value problem solution y, Ly, the associated Green's function H and initial conditions at point $x_0 \in \mathbb{R}$.

RESUMEN

Varias L_p desigualdes de tipo Poincaré [1], hacia adelante o atrás, son dadas para un operador diferencial linear L, envolviendo la solución y de un problema de valor inicial asociado, Ly, la función Green asociada H y las condiciones iniciales en un punto $x_0 \in \mathbb{R}$.

Key words and phrases: *Poincaré inequality, linear differential operator.*

Math. Subj. Class.: *26D10.*

1. Background

Here we follow [2], pp. 145-154.

Let $[a,b] \subset \mathbb{R}$, $a_i(x)$, $i = 0,1,\ldots,n-1$ $(n \in \mathbb{N})$, $h(x)$ be continuous functions on $[a,b]$ and let $L = D^n + a_{n-1}(x)D^{n-1} + \ldots + a_0(x)$ be a fixed linear differential operator on $C^n([a,b])$. Let $y_1(x),\ldots,y_n(x)$ be a set of linear independent solutions to $Ly = 0$. Here the associated Green's functions for L is

$$H(x,t) := \frac{\begin{vmatrix} y_1(t) & \ldots & y_n(t) \\ y_1'(t) & \ldots & y_n'(t) \\ \ldots & \ldots & \ldots \\ y_1^{(n-2)}(t) & \ldots & y_n^{(n-2)}(t) \\ y_1(x) & \ldots & y_n(x) \end{vmatrix}}{\begin{vmatrix} y_1(t) & \ldots & y_n(t) \\ y_1'(t) & \ldots & y_n'(t) \\ \ldots & \ldots & \ldots \\ y_1^{(n-2)}(t) & \ldots & y_n^{(n-2)}(t) \\ y_1^{(n-1)}(t) & \ldots & y_n^{(n-1)}(t) \end{vmatrix}}, \tag{1}$$

which is a continuous function on $[a,b]^2$.

Consider a fixed $x_0 \in [a,b]$, then

$$y(x) = \int_{x_0}^{x} H(x,t)\,h(t)\,dt, \quad \forall x \in [a,b], \tag{2}$$

is the unique solution to the initial value problem

$$Ly = h; \quad y^{(i)}(x_0) = 0, \quad i = 0,1,\ldots,n-1. \tag{3}$$

Next we assume all of the above.

2. Results

We present the following Poincaré type inequalities.

Theorem 1. *Let* $x_0 < b$ *and* $x \in [x_0,b]$, *and* p, $q > 1 : \frac{1}{p} + \frac{1}{q} = 1$; $\nu > 0$.

Then

$$1)\ \|y\|_{L_\nu(x_0,b)} \le \left(\int_{x_0}^{b}\left(\int_{x_0}^{x}|H(x,t)|^p\,dt\right)^{\nu/p}dx\right)^{1/\nu}\|Ly\|_{L_q(x_0,b)}. \tag{4}$$

When $\nu = q$ *we have*

$$2)\ \|y\|_{L_q(x_0,b)} \le \left(\int_{x_0}^{b}\left(\int_{x_0}^{x}|H(x,t)|^p\,dt\right)^{q/p}dx\right)^{1/q}\|Ly\|_{L_q(x_0,b)}. \tag{5}$$

When $\nu = p = q = 2$ *we get*

$$3)\ \|y\|_{L_2(x_0,b)} \le \left(\int_{x_0}^{b}\left(\int_{x_0}^{x}H^2(x,t)\,dt\right)dx\right)^{1/2}\|Ly\|_{L_2(x_0,b)}. \tag{6}$$

Proof. From (2) we have

$$|y(x)| \le \int_{x_0}^{x}|H(x,t)|\,|h(t)|\,dt \le$$

$$\left(\int_{x_0}^{x}|H(x,t)|^p\,dt\right)^{1/p}\left(\int_{x_0}^{x}|h(t)|^q\,dt\right)^{1/q} \le$$

$$\left(\int_{x_0}^{x}|H(x,t)|^p\,dt\right)^{1/p}\left(\int_{x_0}^{b}|h(t)|^q\,dt\right)^{1/q}. \tag{7}$$

That is

$$|y(x)|^\nu \le \left(\int_{x_0}^{x}|H(x,t)|^p\,dt\right)^{\nu/p}\|Ly\|^{\nu}_{L_q(x_0,b)}, \tag{8}$$

Therefore

$$\int_{x_0}^{b}|y(x)|^\nu\,dx \le \left(\int_{x_0}^{b}\left(\int_{x_0}^{x}|H(x,t)|^p\,dt\right)^{\nu/p}dx\right)\|Ly\|^{\nu}_{L_q(x_0,b)}, \tag{9}$$

proving the claim. □

We continue with

Theorem 2. *Let* $x_0 > a$ *and* $x \in [a, x_0]$, *and* $p, q > 1 : \frac{1}{p} + \frac{1}{q} = 1;\ \nu > 0$.

Then

$$1)\ \|y\|_{L_\nu(a,x_0)} \le \left(\int_{a}^{x_0}\left(\int_{x}^{x_0}|H(x,t)|^p\,dt\right)^{\nu/p}dx\right)^{1/\nu}\|Ly\|_{L_q(a,x_0)}. \tag{10}$$

When $\nu = q$ *we have*

$$2)\ \|y\|_{L_q(a,x_0)} \le \left(\int_{a}^{x_0}\left(\int_{x}^{x_0}|H(x,t)|^p\,dt\right)^{q/p}dx\right)^{1/q}\|Ly\|_{L_q(a,x_0)}. \tag{11}$$

When $\nu = p = q = 2$ we get

$$3)\ \|y\|_{L_2(a,x_0)} \leq \left(\int_a^{x_0}\left(\int_x^{x_0} H^2(x,t)\,dt\right)dx\right)^{1/2}\|Ly\|_{L_2(a,x_0)}. \tag{12}$$

Proof. From (2) we have

$$|y(x)| \leq \int_x^{x_0}|H(x,t)|\,|h(t)|\,dt \leq$$

$$\left(\int_x^{x_0}|H(x,t)|^p\,dt\right)^{1/p}\left(\int_x^{x_0}|h(t)|^q\,dt\right)^{1/q} \leq$$

$$\left(\int_x^{x_0}|H(x,t)|^p\,dt\right)^{1/p}\left(\int_a^{x_0}|h(t)|^q\,dt\right)^{1/q}. \tag{13}$$

That is

$$|y(x)|^\nu \leq \left(\int_x^{x_0}|H(x,t)|^p\,dt\right)^{\nu/p}\|Ly\|^\nu_{L_q(a,x_0)}, \tag{14}$$

Therefore

$$\int_a^{x_0}|y(x)|^\nu\,dx \leq \left(\int_a^{x_0}\left(\int_x^{x_0}|H(x,t)|^p\,dt\right)^{\nu/p}dx\right)\|Ly\|^\nu_{L_q(a,x_0)}, \tag{15}$$

proving the claim. □

Extreme cases follow

Proposition 3. *Here $x_0 < b$, $x \in [x_0, b]$, and $p = 1$, $q = \infty$.*

Then

$$1)\ \|y\|_{L_\nu(x_0,b)} \leq \left(\int_{x_0}^{b}\left(\int_{x_0}^{x}|H(x,t)|\,dt\right)^\nu dx\right)^{1/\nu}\|Ly\|_{L_\infty(x_0,b)}. \tag{16}$$

When $\nu = 1$ we have

$$2)\ \|y\|_{L_1(x_0,b)} \leq \left(\int_{x_0}^{b}\left(\int_{x_0}^{x}|H(x,t)|\,dt\right)dx\right)\|Ly\|_{L_\infty(x_0,b)}. \tag{17}$$

Proof. From (2) we have

$$|y(x)| \leq \int_{x_0}^{x}|H(x,t)|\,|h(t)|\,dt \leq$$

$$\left(\int_{x_0}^{x}|H(x,t)|\,dt\right)\|h\|_{L_\infty(x_0,b)}. \tag{18}$$

That is

$$|y(x)|^{\nu} \leq \left(\int_{x_0}^{x} |H(x,t)| \, dt \right)^{\nu} \|Ly\|^{\nu}_{L_\infty(x_0,b)}, \tag{19}$$

and

$$\int_{x_0}^{b} |y(x)|^{\nu} \, dx \leq \left(\int_{x_0}^{b} \left(\int_{x_0}^{x} |H(x,t)| \, dt \right)^{\nu} dx \right) \|Ly\|^{\nu}_{L_\infty(x_0,b)}, \tag{20}$$

proving the claim. □

We continue with

Proposition 4. *Here $x_0 > a$, $x \in [a, x_0]$, and $p = 1$, $q = \infty$.*

Then

$$1) \ \|y\|_{L_\nu(a,x_0)} \leq \left(\int_{a}^{x_0} \left(\int_{x}^{x_0} |H(x,t)| \, dt \right)^{\nu} dx \right)^{1/\nu} \|Ly\|_{L_\infty(a,x_0)}. \tag{21}$$

When $\nu = 1$ we get

$$2) \ \|y\|_{L_1(a,x_0)} \leq \left(\int_{a}^{x_0} \left(\int_{x}^{x_0} |H(x,t)| \, dt \right) dx \right) \|Ly\|_{L_\infty(a,x_0)}. \tag{22}$$

Proof. From (2) we have

$$|y(x)| \leq \int_{x}^{x_0} |H(x,t)| \, |h(t)| \, dt \leq$$

$$\left(\int_{x}^{x_0} |H(x,t)| \, dt \right) \|h\|_{L_\infty(a,x_0)}. \tag{23}$$

That is

$$|y(x)|^{\nu} \leq \left(\int_{x}^{x_0} |H(x,t)| \, dt \right)^{\nu} \|Ly\|^{\nu}_{L_\infty(a,x_0)}, \tag{24}$$

and

$$\int_{a}^{x_0} |y(x)|^{\nu} \, dx \leq \left(\int_{a}^{x_0} \left(\int_{x}^{x_0} |H(x,t)| \, dt \right)^{\nu} dx \right) \|Ly\|^{\nu}_{L_\infty(a,x_0)}, \tag{25}$$

proving the claim. □

Next we give reverse Poincaré type inequalities.

Theorem 5. *Let $x_0 < b$, $x \in [x_0, b]$, and $0 < p < 1$, $q < 0 : \frac{1}{p} + \frac{1}{q} = 1$, $\nu > 0$.*

Assume $H(x,t) \geq 0$ for $x_0 \leq t \leq x$, and $Ly = h$ is of fixed sign and nowhere zero.

Then

$$1) \ \|y\|_{L_\nu(x_0,b)} \geq \left(\int_{x_0}^{b} \left(\int_{x_0}^{x} (H(x,t))^{p} \, dt \right)^{\nu/p} dx \right)^{1/\nu} \|Ly\|_{L_q(x_0,b)}. \tag{26}$$

When $\nu = p$ we get

$$2)\ \|y\|_{L_p(x_0,b)} \geq \left(\int_{x_0}^{b} \left(\int_{x_0}^{x} (H(x,t))^p \, dt \right) dx \right)^{1/p} \|Ly\|_{L_q(x_0,b)} . \tag{27}$$

When $\nu = 1$ we obtain

$$3)\ \|y\|_{L_1(x_0,b)} \geq \left(\int_{x_0}^{b} \left(\int_{x_0}^{x} (H(x,t))^p \, dt \right)^{1/p} dx \right) \|Ly\|_{L_q(x_0,b)} . \tag{28}$$

Proof. By (2) we have

$$|y(x)| = \int_{x_0}^{x} H(x,t) \, |h(t)| \, dt, \text{ for all } x_0 \leq x \leq b. \tag{29}$$

From (29) by reverse Hölder's inequality we obtain

$$\begin{aligned} |y(x)| &\geq \left(\int_{x_0}^{x} (H(x,t))^p \, dt \right)^{1/p} \left(\int_{x_0}^{x} |h(t)|^q \, dt \right)^{1/q} \\ &\geq \left(\int_{x_0}^{x} (H(x,t))^p \, dt \right)^{1/p} \left(\int_{x_0}^{b} |h(t)|^q \, dt \right)^{1/q}, \end{aligned} \tag{30}$$

for all $x_0 < x \leq b$.

I.e. it holds

$$|y(x)|^\nu \geq \left(\int_{x_0}^{x} (H(x,t))^p \, dt \right)^{\nu/p} \|h\|_{L_q(x_0,b)}^{\nu} , \tag{31}$$

for all $x_0 \leq x \leq b$, and

$$\int_{x_0}^{b} |y(x)|^\nu \, dx \geq \left(\int_{x_0}^{b} \left(\int_{x_0}^{x} (H(x,t))^p \, dt \right)^{\nu/p} dx \right) \|h\|_{L_q(x_0,b)}^{\nu} , \tag{32}$$

proving the claim. □

We continue with

Theorem 6. *Let $x_0 > a$, $x \in [a, x_0]$, and $0 < p < 1$, $q < 0 : \frac{1}{p} + \frac{1}{q} = 1$, $\nu > 0$.*

Assume $H(x,t) \leq 0$ for $x \leq t \leq x_0$, and $Ly = h$ is of fixed sign and nowhere zero. Then

$$1)\ \|y\|_{L_\nu(a,x_0)} \geq \left(\int_{a}^{x_0} \left(\int_{x}^{x_0} (-H(x,t))^p \, dt \right)^{\nu/p} dx \right)^{1/\nu} \|Ly\|_{L_q(a,x_0)} . \tag{33}$$

When $\nu = p$ *we get*

$$2)\ \ \|y\|_{L_p(a,x_0)} \geq \left(\int_a^{x_0}\left(\int_x^{x_0}(-H(x,t))^p\,dt\right)dx\right)^{1/p}\|Ly\|_{L_q(a,x_0)}\,. \tag{34}$$

When $\nu = 1$ *we have*

$$3)\ \ \|y\|_{L_1(a,x_0)} \geq \left(\int_a^{x_0}\left(\int_x^{x_0}(-H(x,t))^p\,dt\right)^{1/p}dx\right)\|Ly\|_{L_q(a,x_0)}\,. \tag{35}$$

Proof. From (2) we have

$$|y(x)| = \left|\int_{x_0}^{x} H(x,t)\,h(t)\,dt\right| =$$
$$\left|\int_x^{x_0} H(x,t)\,h(t)\,dt\right| =$$
$$\left|\int_x^{x_0}(-H(x,t))\,h(t)\,dt\right| =$$
$$\int_x^{x_0}(-H(x,t))\,|h(t)|\,dt. \tag{36}$$

From (36) by reverse Hölder's inequality we obtain

$$|y(x)| \geq \left(\int_x^{x_0}(-H(x,t))^p\,dt\right)^{1/p}\left(\int_x^{x_0}|h(t)|^q\,dt\right)^{1/q}$$
$$\geq \left(\int_x^{x_0}(-H(x,t))^p\,dt\right)^{1/p}\left(\int_a^{x_0}|h(t)|^q\,dt\right)^{1/q}. \tag{37}$$

for all $a \leq x < x_0$.

I.e. it holds

$$|y(x)|^{\nu} \geq \left(\int_x^{x_0}(-H(x,t))^p\,dt\right)^{\nu/p}\|Ly\|^{\nu}_{L_q(a,x_0)}\,, \tag{38}$$

for all $a \leq x \leq x_0$, and

$$\int_a^{x_0}|y(x)|^{\nu}\,dx \geq \left(\int_a^{x_0}\left(\int_x^{x_0}(-H(x,t))^p\,dt\right)^{\nu/p}dx\right)\|Ly\|^{\nu}_{L_q(a,x_0)}\,, \tag{39}$$

proving the claim. □

Received: December 2007. Revised: April 2008.

References

[1] G. Acosta and R. G. Durán, *An optimal Poincaré inequality in L^1 for convex domains*, Proc. AMS, **132** (1)2003, pp.195–202.

[2] D. Kreider, R. Kuller, D. Ostberg and F. Perkins, *An Introduction to Linear Analysis*, Addison-Wesley Publishing Company, Inc., Reading, Mass., USA, 1966.

CUBO A Mathematical Journal
Vol.10, Nº 03, (21–41). October 2008

Multiple Solutions for Doubly Resonant Elliptic Problems Using Critical Groups

Ravi P. Agarwal
Department of Mathematical Sciences, Florida Institute of Technology, Melbourne 32901-6975, FL, U.S.A
email: agarwal@fit.edu

Michael E. Filippakis
Department of Mathematics, National Technical University, Zografou Campus, Athens 15780, Greece
email: mfilip@math.ntua.gr

Donal O'Regan
Department of Mathematics, National University of Ireland, Galway, IRELAND
email: donal.oregan@nuigalway.ie

and

Nikolaos S. Papageorgiou
Department of Mathematics, National Technical University, Zografou Campus, Athens 15780, Greece
email: npapg@math.ntua.gr

ABSTRACT

We consider a semilinear elliptic equation, with a right hand side nonlinearity which may grow linearly. Throughout we assume a double resonance at infinity in the spectral interval $[\lambda_1, \lambda_2]$. In this paper, we can also have resonance at zero or even double resonance in the order interval $[\lambda_m, \lambda_{m+1}]$, $m \geq 2$. Using Morse theory and in particular critical groups, we prove two multiplicity theorems.

RESUMEN

Nosotros consideramos una ecuación semilinear eliptica con una no-linealidad la cual puede crecer linealmente. Asumimos una doble resonancia en infinito en el intervalo espectral $[\lambda_1, \lambda_2]$. En este artículo, podemos también tener resonancia en cero o incluso doble resonancia en el intervalo ordenado $[\lambda_m, \lambda_{m+1}]$, $m \geq 2$. Usando teoria de Morse y en particular grupos críticos, provamos dos teoremas de mulplicidad.

Key words and phrases: *Double resonance, C-condition, critical groups, critical point of mountain pass-type, Poincare-Hopf formula.*

Math. Subj. Class.: *35J20, 35J25.*

1 Introduction

Let $Z \subseteq \mathbb{R}^{\mathbb{N}}$ be a bounded domain with a C^2-boundary ∂Z. We consider the following semilinear elliptic problem:

$$\left\{ \begin{array}{l} -\triangle x(z) = \lambda_1 x(z) + f(z, x(z)) \text{ a.e. on } Z, \\ x|_{\partial Z} = 0. \end{array} \right\} \tag{1.1}$$

Here $\lambda_1 > 0$ is the principal eigenvalue of $(-\triangle, H_0^1(Z))$. Assume that

$$\lim_{|x| \to \infty} \frac{f(z,x)}{x} = 0 \text{ uniformly for a.a. } z \in Z. \tag{1.2}$$

The problem (1.1) is resonant at infinity with respect to the principal eigenvalue $\lambda_1 > 0$. Resonant problems, were first studied by Landesman-Lazer [7], who assumed a bounded nonlinearity and introduced the well-known sufficient asymptotic solvability conditions, which carry their name (the LL-conditions for short). We can be more general and instead of (1.2), assume only that

$$\liminf_{|x| \to \infty} \frac{f(z,x)}{x} \text{ and } \limsup_{|x| \to \infty} \frac{f(z,x)}{x}$$

belong in the interval $[0, \lambda_2 - \lambda_1]$ uniformly for a.a. $z \in Z$, with λ_2 ($\lambda_2 > \lambda_1$) being the second eigenvalue of $(-\triangle, H_0^1(Z))$. In this more general setting, the nonlinearity $f(z,x)$ need not be bounded. This more general situation was examined by Berestycki-De Figueiredo [2], Landesman-Robinson-Rumbos [8], Nkashama [11], Robinson [13],[14], Rumbos [15] and Su [16]. From these works, Berestycki-De Figueiredo [2], Nkashama [11], Robinson [13] and Rumbos [15], prove existence theorems in a double resonance setting (i.e. asymptotically at $\pm\infty$, we have complete interaction of the ”slope” $\frac{f(z,x)}{x}$ with both ends of the spectral interval $[0, \lambda_2 - \lambda_1]$; see Berestycki-De Figueiredo [2] who coined the term ”double resonance” and Robinson [13]) or in a one-sided resonance setting (i.e. the ”slope” $\frac{f(z,x)}{x}$ is not allowed to cross $\lambda_2 - \lambda_1$; see Nkashama [11] and Rumbos [15]). Multiplicity results were proved by Landesman-Robinson-Rumbos [8] (one-sided resonant problems) and by Robinson [14] and Su [16] (doubly resonant problems).

In this paper, we extend the work of Landesman-Robinson-Rumbos [8] and partially extend and complement the works of Robinson [14] and Su [16], by covering cases which are not included in their multiplicity results.

2 Mathematical background

We start by recalling some basic facts about the following weighted linear eigenvalue problem:

$$\left\{ \begin{array}{l} -\triangle u(z) = \widehat{\lambda} m(z) u(z) \text{ a.e. on } Z, \\ u|_{\partial Z} = 0, \ \ \widehat{\lambda} \in \mathbb{R}. \end{array} \right\} \tag{2.1}$$

Here $m \in L^\infty(Z)_+ = \{m \in L^\infty(Z) : m(z) \geq 0 \text{ a.e. on } Z\}$, $m \neq 0$ (the weight function). By an eigenvalue of (2.1), we mean a real number $\widehat{\lambda}$, for which problem (2.1) has a nontrivial solution $u \in H_0^1(Z)$. It is well-known (see for example Gasinski-Papageorgiou [5]), that problem (2.1) (or equivalently that $(-\triangle, H_0^1(Z), m)$), has a sequence $\{\widehat{\lambda}_k(m)\}_{k\geq 1}$ of distinct eigenvalues, $\widehat{\lambda}_1(m) > 0$ and $\widehat{\lambda}_k(m) \to +\infty$ as $k \to +\infty$. Moreover, $\widehat{\lambda}_1(m) > 0$ is simple (i.e. the corresponding eigenspace $E(\widehat{\lambda}_1)$ is one-dimensional). Also we can find an orthonormal basis $\{u_n\}_{n\geq 1} \subseteq H_0^1(Z) \cap C^\infty(Z)$ for the Hilbert space $L^2(Z)$ consisting of eigenfunctions corresponding to the eigenvalues $\{\widehat{\lambda}_k(m)\}_{k\geq 1}$. Note that $\{u_n\}_{n\geq 1}$ is also an orthogonal basis for the Hilbert space $H_0^1(Z)$. Moreover, since by hypothesis ∂Z is a C^2-manifold, then $u_n \in C^2(\overline{Z})$ for all $n \geq 1$. For every $k \geq 1$, by $E(\widehat{\lambda}_k)$ we denote the eigenspace corresponding to the eigenvalue $\widehat{\lambda}_k(m)$. This space has the so-called ”unique continuation property”, namely, if $u \in E(\widehat{\lambda}_k)$ is such that it vanishes on a set of positive measure, then $u(z) = 0$ for all $z \in \overline{Z}$. We set

$$\overline{H}_k = \mathop{\oplus}_{i=1}^{k} E(\widehat{\lambda}_i)$$
$$\text{and } \ \widehat{H}_{k+1} = \overline{\mathop{\oplus}_{i\geq k+1} E(\lambda_i)} = \overline{H}_k^{\perp}, \quad k \geq 1.$$

We have the orthogonal direct sum decomposition

$$H_0^1(Z) = \overline{H}_k \oplus \widehat{H}_{k+1}.$$

Using these spaces, we can have useful variational characterizations of the eigenvalues $\{\widehat{\lambda}_k(m)\}_{k\geq 1}$ using the Rayleigh quotient. Namely we have:

$$\widehat{\lambda}_1(m) = \min \left[\frac{\|Du\|_2^2}{\int_Z mu^2 dz} : u \in H_0^1(Z), u \neq 0 \right]. \tag{2.2}$$

In (2.2) the minimum is attained on $E(\widehat{\lambda}_1)\backslash\{0\}$. By $u_1 \in C_0^2(\overline{Z})$, we denote the principal eigenfunction satisfying $\int_Z m u_1^2 dz = 1$. For $k \geq 2$, we have

$$\widehat{\lambda}_k(m) = \max\left[\frac{\|D\overline{u}\|_2^2}{\int_Z m\overline{u}^2 dz} : \overline{u} \in \overline{H}_k, \overline{u} \neq 0\right] \tag{2.3}$$

$$= \min\left[\frac{\|D\widehat{u}\|_2^2}{\int_Z m\widehat{u}^2 dz} : \widehat{u} \in \widehat{H}_k, \widehat{u} \neq 0\right]. \tag{2.4}$$

In (2.3) (resp.(2.4)), the maximum (resp.minimum) is attained on $E(\widehat{\lambda}_k)$. From these variational characterizations of the eigenvalues and the unique continuation property of the eigenspaces $E(\widehat{\lambda}_k)$, we see that the eigenvalues $\{\widehat{\lambda}_k(m)\}_{k\geq 1}$ have the following strict monotonicity property:

"If $m_1, m_2 \in L^\infty(Z)_+$, $m_1(z) \leq m_2(z)$ a.e. on Z and $m_1 \neq m_2$, then $\widehat{\lambda}_k(m_2) < \widehat{\lambda}_k(m_1)$ for all $k \geq 1$."

If $m \equiv 1$, then we simply write λ_k for all $k \geq 1$ and we have the full-spectrum of $(-\triangle, H_0^1(Z))$.

Let H be a Hilbert space and $\varphi \in C^1(H)$. We say that φ satisfies the "Cerami condition" (the C-condition for short), if the following is true:"every sequence $\{x_n\}_{n\geq 1} \subseteq H$ such that $|\varphi(x_n)| \leq M_1$ for some $M_1 > 0$, all $n \geq 1$ and $(1 + \|x_n\|)\varphi'(x_n) \to 0$ in H^* as $n \to \infty$, has a strongly convergent subsequence".

This condition is a weakened version of the well-known Palais-Smale condition (PS-condition for short). Bartolo-Benci-Fortunato [1], showed that the C-condition suffices to prove a deformation theorem and from this produce minimax expressions for the critical values of the functional φ.

For every $c \in \mathbb{R}$, let

$$\begin{aligned} \varphi^c &= \{x \in X : \varphi \leq c\} \text{ (the sublevel set at } c \text{ of } \varphi), \\ K &= \{x \in x : \varphi'(x) = 0\} \text{ (the set of critical points of } \varphi) \\ \text{and } K_c &= \{x \in K : \varphi(x) = c\} \text{ (the critical points of } \varphi \text{ at level } c). \end{aligned}$$

If X is a Hausdorff topological space and Y a subspace of it, for every integer $n \geq 0$, by $H_n(X,Y)$ we denote the n^{th}-relative singular homology group with integer coefficients. The critical groups of φ at an isolated critical point $x_0 \in H$ with $\varphi(x_0) = c$, are defined by

$$C_n(\varphi, x_0) = H_n(\varphi^c \cap U, (\varphi^c \cap U)\backslash\{x_0\}),$$

where U is a neighborhood of x_0 such that $K \cap \varphi^c \cap U = \{x_0\}$. By the excision property of singular homology theory, we see that the above definition of critical groups, is independent of U (see for example Mawhin-Willem [10]).

Suppose that $-\infty < \inf \varphi(K)$. Choose $c < \inf \varphi(K)$. The critical groups at infinity, are defined by

$$C_k(\varphi, \infty) = H_k(H, \varphi^c) \text{ for all } k \geq 0.$$

If K is finite, then the Morse-type numbers of φ, are defined by

$$M_k = \sum_{x\in K} \mathrm{rank} C_k(\varphi, x).$$

The Betti-type numbers of φ, are defined by

$$\beta_k = \mathrm{rank} C_k(\varphi, \infty).$$

By Morse theory (see Chang [4] and Mawhin-Willem [10]), we have

$$\sum_{k=0}^{m}(-1)^{m-k}M_k \geq \sum_{k=0}^{m}(-1)^{m-k}\beta_k$$
$$\text{and} \quad \sum_{k\geq 0}(-1)^k M_k = \sum_{k\geq 0}(-1)^k \beta_k.$$

From the first relation, we deduce that $\beta_k \leq M_k$ for all $k \geq 0$. Therefore, if $\beta_k \neq 0$ for some $k \geq 0$, then φ must have a critical point $x \in H$ and the critical group $C_k(\varphi, x)$ is nontrivial. The second relation (the equality), is known as the "Poincare-Hopf formula". Finally, if $K = \{x_0\}$, then $C_k(\varphi, \infty) = C_k(\varphi, x_0)$ for all $k \geq 0$.

3 Multiplicity of solutions

The hypotheses on the nonlinearity $f(z, x)$ are the following:

<u>$H(f)$</u>: $f : Z \times \mathbb{R} \to \mathbb{R}$ is a function such that $f(z, 0) = 0$ a.e. on Z and

- **(i)** for all $x \in \mathbb{R}$, $z \to f(z, x)$ is measurable;
- **(ii)** for almost all $z \in Z$, $f(z, \cdot) \in C^1(\mathbb{R})$;
- **(iii)** $|f'_x(z, x)| \leq c(1 + |x|^r),\ r < \frac{4}{N-2}, c > 0.$
- **(iv)** $0 \leq \liminf_{|x|\to\infty} \frac{f(z,x)}{x} \leq \limsup_{|x|\to\infty} \frac{f(z,x)}{x} \leq \lambda_2 - \lambda_1$ uniformly for a.a. $z \in Z$;
- **(v)** suppose that $\|x_n\| \to \infty$,
  - (i) if $\frac{\|x_n^0\|}{\|x_n\|} \to 1$, $x_n = x_n^0 + \widehat{x}_n$ with $x_n^0 \in E(\lambda_1) = \overline{H}_1$, $\widehat{x}_n \in \widehat{H}_2$, then there exist $\gamma_1 > 0$ and $n_1 \geq 1$ such that

$$\int_Z f(z, x_n(z))x_n^0(z)dz \geq \gamma_1 \ \text{ for all } \ n \geq n_1;$$

  - (ii) if $\frac{\|x_n^0\|}{\|x_n\|} \to 1$, $x_n = x_n^0 + \widehat{x}_n$ with $x_n^0 \in E(\lambda_2)$, $\widehat{x}_n \in W = E(\lambda_2)^\perp$, then there exist $\gamma_2 > 0$ and $n \geq 1$ such that

$$\int_Z (f(z, x_n(z)) - (\lambda_2 - \lambda_1)x_n(z))x_n^0(z)dz \leq -\gamma_2 \ \text{ for all } \ n \geq n_2;$$

(vi) if $F(z,x)=\int_0^x f(z,s)ds$, then there exist $\eta\in L^\infty(Z)$ and $\delta>0$, such that $\eta(z)\le 0$ a.e. on Z with strict inequality on a set of positive measure and

$$F(z,x)\le \frac{\eta(z)}{2}x^2 \text{ for a.a. } z\in Z \text{ and all } |x|\le\delta.$$

Remark 3.1. *Hypothesis $H(f)(iv)$ implies that asymptotically at $\pm\infty$, we have double resonance. Hypothesis $H(f)(v)$ is a generalized LL-condition. Similar conditions can be found in the works of Landesman-Robinson-Rumbos [8], Robinson [13],[14] and Su [16]. Consider a C^2-function $x\to F(x)$ which in a neighborhood of zero equals x^4-sinx^2, while for $|x|$ large (say $|x|\ge M>0$), $F(x)=c|x|^{\frac{3}{2}}, c>0$. If $f(x)=F'(x)$, then $f\in C^1(\mathbb{R})$ satisfies hypothesis $H(f)$ above. To verify the generalized LL-condition in hypothesis $H(f)(v)$, we use Lemma 2.1 of Su-Tang [17]. Similarly we can consider if near the origin, $F(x)=\frac{1}{2}x^2-tan^{-1}x^2$ or $F(x)=-cosx^2$. This second case is interesting because then $f(x)=2xsinx^2$ and $f'(x)=2sinx^2+4x^2cosx^2$. So $f'(0)=0$. This example, which is covered by hypotheses $H(f)$, illustrates that our framework of analysis incorporates also problems with resonance at zero with respect to $\lambda_1>0$ (double-double resonance). This is not possible in the setting of Landesman-Robinson-Rumbos [8] (see Theorem 2 in [8]). Also such a potential function is not covered by the multiplicity results of Robinson [14] (theorem 2) and Su [16] (Theorem 2).*

We consider the Euler functional for problem (1.1), $\varphi: H_0^1(Z)\to\mathbb{R}$ defined by

$$\varphi(x)=\frac{1}{2}\|Dx\|_2^2-\frac{\lambda_1}{2}\|x\|_2^2-\int_Z F(z,x(z))dz \text{ for all } x\in H_0^1(Z).$$

It is well-known that $\varphi\in C^2(H_0^1(Z))$ and if by $\langle\cdot,\cdot\rangle$ we denote the duality brackets for the pair $(H_0^1(Z), H^{-1}(Z)=H_0^1(Z)^*)$, we have

$$\langle\varphi'(x),y\rangle=\int_Z(Dx,Dy)_{\mathbb{R}^N}dz-\lambda_1\int_Z xydz-\int_Z f(z,x(z))y(z)dz$$

$$\text{and } \varphi''(x)(u,v)=\int_Z(Du,Dv)_{\mathbb{R}^N}dz-\lambda_1\int_Z uvdz-\int_Z f'(z,x(z))u(z)v(z)dz$$

for all $x,y,u,v\in H_0^1(Z)$.

Proposition 3.2. *If hypotheses $H(f)$ hold then φ satisfies the C-condition.*

Proof. Let $\{x_n\}_{n\ge1}\subseteq H_0^1(Z)$ be a sequence such that

$$(1+\|x_n\|)\varphi'(x_n)\to 0 \text{ as } n\to\infty.$$

We will show that $\{x_n\}_{n\ge1}\subseteq H_0^1(Z)$ is bounded. We argue indirectly. Suppose that $\{x_n\}_{n\ge1}\subseteq H_0^1(Z)$ is unbounded. We may assume that $\|x_n\|\to\infty$. Let $y_n=\frac{x_n}{\|x_n\|}$, $n\ge1$. By passing to a suitable subsequence if necessary, we may assume that

$$y_n\xrightarrow{w} y \text{ in } H_0^1(Z),\ y_n\to y \text{ in } L^2(Z),\ y_n(z)\to y(z) \text{ a.e. on } Z$$

$$\text{and } |y_n(z)|\le k(z) \text{ a.e. on } Z, \text{ for all } n\ge1, \text{ with } k\in L^2(Z)_+.$$

Hypotheses $H(f)(iii)$ and (iv), imply that

$$|f(z,x)| \leq a(z) + c|x| \text{ for a.a. } z \in Z, \text{ all } x \in \mathbb{R}, \text{ with } a \in L^{\infty}(Z)_+, c > 0,$$

$$\Rightarrow \frac{|f(z,x_n(z))|}{\|x_n\|} \leq \frac{a(z)}{\|x_n\|} + c|y_n(z)| \text{ for a.a. } z \in Z, \text{ all } n \geq 1, \tag{3.1}$$
$$\Rightarrow \left\{ \frac{f(\cdot,x_n(\cdot))}{\|x_n\|} \right\}_{n\geq 1} \subseteq L^2(Z) \text{ is bounded.}$$

Thus we may assume that

$$\frac{f(\cdot,x_n(\cdot))}{\|x_n\|} \xrightarrow{w} h \text{ in } L^2(Z) \text{ as } n \to \infty.$$

For every $\varepsilon > 0$ and $n \geq 1$, we set

$$C^+_{\varepsilon,n} = \{z \in Z : x_n(z) > 0, \ -\varepsilon \leq \frac{f(z,x_n(z))}{x_n(z)} \leq \lambda_2 - \lambda_1 + \varepsilon\}$$
$$\text{and } C^-_{\varepsilon,n} = \{z \in Z : x_n(z) < 0, \ -\varepsilon \leq \frac{f(z,x_n(z))}{x_n(z)} \leq \lambda_2 - \lambda_1 + \varepsilon\}$$

Note that $x_n(z) \to +\infty$ a.e. on $\{y > 0\}$ and $x_n(z) \to -\infty$ a.e. on $\{y < 0\}$. Then by virtue of hypothesis $H(f)(iv)$, we have

$$\chi_{C^+_{\varepsilon,n}}(z) \to \chi_{\{y>0\}}(z) \text{ and } \chi_{C^-_{\varepsilon,n}}(z) \to \chi_{\{y<0\}}(z) \text{ a.e. on } Z.$$

Using the dominated convergent theorem, we see that

$$\|(1-\chi_{C^+_{\varepsilon,n}})\frac{f(\cdot,x_n(\cdot))}{\|x_n\|}\|_{L^2(\{y>0\})} \to 0$$
$$\text{and } \|(1-\chi_{C^-_{\varepsilon,n}})\frac{f(\cdot,x_n(\cdot))}{\|x_n\|}\|_{L^2(\{y<0\})} \to 0 \text{ as } n \to \infty.$$

It follows that

$$\chi_{C^+_{\varepsilon,n}}(\cdot)\frac{f(\cdot,x_n(\cdot))}{\|x_n\|} \xrightarrow{w} h \text{ in } L(\{y>0\})$$
$$\text{and } \chi_{C^-_{\varepsilon,n}}(\cdot)\frac{f(\cdot,x_n(\cdot))}{\|x_n\|} \xrightarrow{w} h \text{ in } L(\{y<0\}) \text{ as } n \to \infty.$$

From the definitions of the sets $C^+_{\varepsilon,n}$ and $C^-_{\varepsilon,n}$ we have

$$-\varepsilon y_n(z) \leq \frac{f(z,x_n(z))}{\|x_n\|} = \frac{f(z,x_n(z))}{x_n(z)} y_n(z) \leq (\lambda_2 - \lambda_1 + \varepsilon) y_n(z) \text{ a.e. on } C^+_{\varepsilon,n}$$

and

$$-\varepsilon y_n(z) \geq \frac{f(z,x_n(z))}{\|x_n\|} = \frac{f(z,x_n(z))}{x_n(z)} y_n(z) \geq (\lambda_2 - \lambda_1 + \varepsilon) y_n(z) \text{a.e. on } C^-_{\varepsilon,n}.$$

Passing to the limit as $n \to \infty$, using Mazur's lemma and recalling that $\varepsilon > 0$ is arbitrary, we obtain

$$0 \le h(z) \le (\lambda_2 - \lambda_1)y(z) \text{ a.e. on } \{y > 0\} \tag{3.2}$$
$$\text{and } 0 \ge h(z) \ge (\lambda_2 - \lambda_1)y(z) \text{ a.e. on } \{y < 0\}. \tag{3.3}$$

Moreover, from (3.1) it is clear that

$$h(z) = 0 \text{ a.e. on } \{y = 0\}. \tag{3.4}$$

From (3.2), (3.3) and (3.4), it follows that

$$h(z) = g(z)y(z) \text{ a.e. on } Z,$$

where $g \in L^\infty(Z)_+$, $0 \le g(z) \le \lambda_2 - \lambda_1$ a.e. on Z.

Recall that by $\langle \cdot, \cdot \rangle$ we denote the duality brackets for the pair $(H_0^1(Z), H^{-1}(Z))$.

Let $A \in \mathcal{L}(H_0^1(Z), H^{-1}(Z))$ be defined by

$$\langle A(x), y \rangle = \int_Z (Dx, Dy)_{\mathbb{R}^N} dz \text{ for all } x, y \in H_0^1(Z).$$

Also let $N : L^2(Z) \to L^2(Z)$ be the Nemitskii operator corresponding to the nonlinearity $f(z, x)$, i.e.

$$N(x)(\cdot) = f(\cdot, x(\cdot)) \text{ for all } x \in L^2(Z).$$

Because of (3.1), by Krasnoselskii's theorem, we know that N is continuous and bounded. Moreover, exploiting the compact embedding of $H_0^1(Z)$ into $L^2(Z)$, we see that N is completely continuous (hence compact too) as a map from $H_0^1(Z)$ into $L^2(Z)$ (see for example Gasinski-Papageorgiou [5], pp.267-268). We have

$$\varphi'(x_n) = A(x_n) - \lambda_1 x_n - N(x_n) \text{ for all } n \ge 1.$$

From the choice of the sequence $\{x_n\}_{n\ge1} \subseteq H_0^1(Z)$, we know that

$$|\langle \varphi'(x_n), v \rangle| \le \varepsilon_n \text{ for all } v \in H_0^1(Z) \text{ with } \varepsilon_n \downarrow 0,$$
$$\Rightarrow \left| \langle A(y_n) - \lambda_1 y_n - \frac{N(x_n)}{\|x_n\|}, v \rangle \right| \le \frac{\varepsilon_n}{\|x_n\|} \text{ for all } n \ge 1. \tag{3.5}$$

Let $v = y_n - y \in H_0^1(Z)$, $n \ge 1$. Then

$$\left| \langle A(y_n), y_n - y \rangle - \lambda_1 \int_Z y_n(y_n - y)dz - \int_Z \frac{N(x_n)}{\|x_n\|}(y_n - y)dz \right| \le \frac{\varepsilon_n}{\|x_n\|} \text{ for all } n \ge 1. \tag{3.6}$$

Evidently

$$\int_Z y_n(y_n - y)dz \to 0 \text{ and } \int_Z \frac{N(x_n)}{\|x_n\|}(y_n - y) \to 0 \text{ as } n \to \infty.$$

So from (3.6), we infer that

$$\langle A(y_n), y_n - y\rangle \to 0. \tag{3.7}$$

We have $A(y_n) \xrightarrow{w} A(y)$ in $H^{-1}(Z)$. From (3.7) it follows that

$$\begin{aligned} &\langle A(y_n), y_n\rangle \to \langle A(y), y\rangle, \\ \Rightarrow &\|Dy_n\|_2 \to \|Dy\|_2. \end{aligned}$$

Also $Dy_n \xrightarrow{w} Dy$ in $L^2(Z, \mathbb{R}^{\mathbb{N}})$. Since the Hilbert space $L^2(Z, \mathbb{R}^{\mathbb{N}})$ has the Kadec-Klee property, we deduce that

$$Dy_n \to Dy \text{ in } L^2(Z, \mathbb{R}^{\mathbb{N}}) \Rightarrow y_n \to y \text{ in } H_0^1(Z), \text{ i.e. } \|y\| = 1,\ y \neq 0.$$

We return to (3.5) and we pass to the limit as $n \to \infty$. We obtain

$$\begin{aligned} &\langle A(y) - \lambda_1 y - gy, v\rangle = 0 \text{ for all } v \in H_0^1(Z), \\ \Rightarrow &A(y) = (\lambda_1 + g)y \text{ in } H^{-1}(Z), \\ \Rightarrow &-\triangle y(z) = (\lambda_1 + g(z))y(z) \text{ a.e. on } Z,\ y|_{\partial Z} = 0. \end{aligned} \tag{3.8}$$

We distinguish three cases for problem (3.8) depending on where the function $g \in L^\infty(Z)_+$ stands in the interval $[0, \lambda_2 - \lambda_1]$.

Case 1: $g(z) = 0$ a.e. on Z.

Then from (3.8), we have

$$\begin{aligned} &-\triangle y(z) = \lambda_1 y(z) \text{ a.e. on } Z,\ \ y|_{\partial Z} = 0, \\ \Rightarrow &y \in E(\lambda_1),\ y \neq 0. \end{aligned}$$

We consider the orthogonal direct sum decomposition $H_0^1(Z) = E(\lambda_1) \oplus \widehat{H}_2$, $\widehat{H}_2 = E(\lambda_1)^\perp$.

Then for every $n \geq 1$, we have

$$x_n = x_n^0 + \widehat{x}_n \text{ and } x_n^0 \in E(\lambda_1),\ \widehat{x}_n \in \widehat{H}_2.$$

We have $y_n = y_n^0 + \widehat{y}_n$, with

$$y_n^0 = \frac{x_n^0}{\|x_n\|} \in E(\lambda_1) \text{ and } \widehat{y}_n = \frac{\widehat{x}_n}{\|x_n\|} \in \widehat{H}_2 \text{ for all } n \geq 1.$$

Since $y \in E(\lambda_1)$, $\|y\| = 1$, we have

$$\frac{\|x_n^0\|}{\|x_n\|} \to 1 \text{ as } n \to \infty.$$

Recall that

$$\left| \langle A(x_n), v \rangle - \lambda_1 \int_Z x_n v dz - \int_Z N(x_n) v dz \right| \leq \varepsilon_n \text{ for all } v \in H_0^1(Z).$$

Let $v = x_n^0 \in H_0^1(Z)$. We have

$$\begin{aligned} &\left| \|Dx_n^0\|_2^2 - \lambda_1 \|x_n^0\|_2^2 - \int_Z f(z, x_n(z)) x_n^0(z) dz \right| \leq \varepsilon_n, \\ \Rightarrow &\int_Z f(z, x_n(z)) x_n^0(z) dz \leq \varepsilon_n \text{ (see (2.2)) for all } n \geq 1. \end{aligned} \tag{3.9}$$

But by virtue of hypothesis $H(f)(v)$

$$0 < \gamma_1 \leq \int_Z f(z, x(z)) x_n^0(z) dz \text{ for all } n \geq n_1. \tag{3.10}$$

Comparing (3.9) and (3.10), we reach a contradiction.

<u>Case 2:</u> $g(z) = \lambda_2 - \lambda_1$ a.e. on Z.

In this case, from (3.8) we have

$$\begin{aligned} &-\triangle y(z) = \lambda_2 y(z) \text{ a.e. on Z, } y|\partial Z = 0, \\ \Rightarrow &y \in E(\lambda_2), \ y \neq 0. \end{aligned}$$

Now we consider the orthogonal direct sum decomposition $H_0^1(Z) = E(\lambda_2) \oplus W$, with $W = E(\lambda_2)^\perp$. Then

$$x_n = x_n^0 + \widehat{x}_n \text{ with } x_n^0 \in E(\lambda_2), \ \widehat{x}_n \in W, \ n \geq 1.$$

Since $y \in E(\lambda_2)$, $\|y\| = 1$, we have

$$\frac{\|x_n^0\|}{\|x_n\|} \to 1 \text{ as } n \to \infty. \tag{3.11}$$

We have

$$\begin{aligned} &|\langle A(x_n), v \rangle - \lambda_1 \int_Z x_n v dz - \int_Z f(z, x_n(z)) v(z) dz| \leq \varepsilon_n \\ &\text{for all } v \in H_0^1(Z), \text{ with } \varepsilon_n \downarrow 0. \end{aligned}$$

Let $v = x_n^0$. Then

$$\left| \|Dx_n^0\|_2^2 - \lambda_1 \|x_n^0\|_2^2 - \int_Z f(z, x_n(z)) x_n^0(z) dz \right| \leq \varepsilon_n,$$

$$\Rightarrow \left| \|Dx_n^0\|_2^2 - \lambda_2 \|x_n^0\|_2^2 - \int_Z (f(z, x_n(z)) - (\lambda_2 - \lambda_1) x_n(z)) x_n^0(z) dz \right| \leq \varepsilon_n,$$

$$\Rightarrow \int_Z (f(z, x_n(z)) - (\lambda_2 - \lambda_1) x_n(z)) x_n^0(z) dz \geq -\varepsilon_n \text{ (see (2.3) and (2.4)).} \tag{3.12}$$

But again hypothesis $H(f)(v)$ implies

$$0 > -\gamma_2 \geq \int_Z (f(z, x_n(z)) - (\lambda_2 - \lambda_1) x_n(z)) x_n^0(z) dz \text{ for all } n \geq n_2. \tag{3.13}$$

Comparing (3.12) and (3.13) we reach a contradiction.

<u>Case 3:</u> $0 \leq g(z) \leq \lambda_2 - \lambda_1$ a.e. on Z with $g \neq 0$, $g \neq \lambda_2 - \lambda_1$.

Note that

$$\lambda_1 \leq \lambda_1 + g(z) \leq \lambda_2 \text{ a.e. on } Z$$

and the inequalities are strict on sets (in general different) of positive measure. Exploiting the strict monotonicity property of the eigenvalues of $(-\triangle, H_0^1(Z), m)$ on the weight function m (see Section 2), we have

$$\widehat{\lambda}_1(\lambda_1 + g) < \widehat{\lambda}_1(\lambda_1) = 1$$
$$\text{and } \widehat{\lambda}_2(\lambda_1 + g) > \widehat{\lambda}_2(\lambda_2) = 1.$$

Combining this with (2.2), we see that $y = 0$, a contradiction to the fact that $\|y\| = 1$.

So in all these cases we have reached a contradiction. This means that $\{x_n\}_{n \geq 1}$ is bounded and so we may assume (at least for a subsequence) that

$$x_n \xrightarrow{w} x \text{ in } H_0^1(Z), \ x_n \to x \text{ in } L^2(Z), \ x_n(z) \to x(z) \text{ a.e. on } Z$$
$$\text{and } |x_n(z)| \leq k(z) \text{ a.e. on } Z \text{ for all } n \geq 1, \text{ with } k \in L^2(Z)_+.$$

Recall that

$$\left| \langle A(x_n), x_n - x \rangle - \lambda_1 \int_Z x_n(x_n - x) dz - \int_Z f(z, x_n(z))(x_n - x) dz \right| \leq \varepsilon_n.$$

Since

$$\int_Z x_n(x_n - x) dz \to 0 \text{ and } \int_Z f(z, x_n(z))(x_n - x)(z) dz \to 0 \text{ as } n \to \infty,$$

we obtain

$$\langle A(x_n), x_n - x \rangle \to 0 \text{ as } n \to \infty.$$

We know that $A(x_n) \xrightarrow{w} A(x)$ in $H^{-1}(Z)$. So as before, via the Kadec-Klee property of $H_0^1(Z)$, we conclude that $x_n \to x$ in $H_0^1(Z)$. This proves that φ satisfies the C-condition. □

In the sequel, we will need the following simple lemma:

Lemma 3.3. *If $\beta \in L^\infty(Z)$, $\beta(z) \le \lambda_1$ a.e. on Z and the inequality is strict on a set of positive measure, then there exists $\xi_1 > 0$ such that*

$$\psi(x) = \|Dx\|_2^2 - \int_Z \beta(z)x(z)^2 dz \ge \xi_1 \|Dx\|_2^2 \text{ for all } x \in H_0^1(Z).$$

Proof. From (2.2), we see that $\psi \ge 0$. Suppose that the lemma is not true. Exploiting the 2-homogeneity of ψ, we can find $\{x_n\}_{n\ge1} \subseteq H_0^1(Z)$ such that

$$\|Dx_n\|_2 = 1 \text{ for all } n \ge 1 \text{ and } \psi(x_n) \downarrow 0 \text{ as } n \to \infty.$$

By Poincare's inequality $\{x_n\}_{n\ge1} \subseteq H_0^1(Z)$ is bounded. So we may assume that

$$\begin{gathered} x_n \xrightarrow{w} x \text{ in } H_0^1(Z),\ x_n \to x \text{ in } L^2(Z),\ x_n(z) \to x(z) \text{ a.e. on } Z \\ \text{and } |x_n(z)| \le k(z) \text{ a.e. on } Z \text{ for all } n \ge 1, \text{ with } k \in L^2(Z)_+. \end{gathered}$$

From the weak lower semicontinuity of the norm functional, we have

$$\|Dx\|_2^2 \le \liminf_{n\to\infty} \|Dx_n\|_2^2,$$

while from the dominated convergence theorem, we have

$$\int_Z \beta(z)x_n(z)^2 dz \to \int_Z \beta(z)x(z)^2 dz \text{ as } n \to \infty.$$

Hence

$$\begin{aligned} &\psi(x) \le \liminf_{n\to\infty} \psi(x_n) = 0, && (3.14) \\ \Rightarrow &\|Dx\|_2^2 \le \int_Z \beta(z)x(z)^2 dz \le \lambda_1 \|x\|_2^2, \\ \Rightarrow &\|Dx\|_2^2 = \lambda_1 \|x\|_2^2 \text{ (see (2.2))}, \\ \Rightarrow &x = 0 \text{ or } x = \pm u_1 \text{ with } u_1 \in E(\lambda_1). \end{aligned}$$

If $x = 0$, then $\|Dx_n\|_2 \to 0$, a contradiction to the fact that $\|Dx_n\|_2 = 1$ for all $n \ge 1$.

If $x = \pm u_1$, then $|x(z)| > 0$ for all $z \in Z$ and so from the first inequality in (3.9) and the hypothesis on β, we have

$$\|Dx\|_2^2 < \lambda_1 \|x\|_2^2,$$

a contradiction to (2.2). ∎

Using this lemma, we prove the following proposition.

Proposition 3.4. *If hypotheses $H(f)$ hold, then the origin is a local minimizer of φ.*

Proof. Let $\delta > 0$ be as in hypothesis $H(f)(vi)$ and consider the closed ball

$$\overline{B}_{\delta}^{C_0^1} = \{x \in C_0^1(\overline{Z}) : \|x\|_{C_0^1(\overline{Z})} \leq \delta\}.$$

By virtue of hypothesis $H(f)(vi)$, for every $x \in \overline{B}_{\delta}^{C_0^1}$, we have

$$F(z, x(z)) \leq \frac{\eta(z)}{2} x(z)^2 \ \text{ for a.a. } \ z \in Z. \tag{3.15}$$

Thus, for all $x \in \overline{B}_{\delta}^{C_0^1}$, we have

$$\begin{aligned} \varphi(x) &= \frac{1}{2}\|Dx\|_2^2 - \frac{\lambda_1}{2}\|x\|_2^2 - \int_Z F(z, x(z))dz \\ &\geq \frac{1}{2}\|Dx\|_2^2 - \frac{1}{2}\int_Z (\lambda_1 + \eta(z))x(z)^2 dz \ \text{ (see (3.15))} \\ &\geq \frac{\xi_1}{2}\|Dx\|_2^2 \ \text{ (apply Lemma 3.3 with } g = \lambda_1 + \eta \in L^\infty(Z)) \\ &\geq 0 = \varphi(0). \end{aligned} \tag{3.16}$$

From (3.16) we see that $x = 0$ is a local $C_0^1(\overline{Z})$-minimizer of φ. But then from Brezis-Nirenberg [3], we have that $x = 0$ is a local $H_0^1(Z)$-minimizer of φ. □

We may assume that the origin is an isolated critical point of φ or otherwise we have a sequence of nontrivial solutions for problems (1.1). Then from the description of the critical groups at an isolated local minimizer (see Chang [4], p.33 and Mawhin-Willem [10], p.175), we have:

Corollary 3.5. *If hypotheses $H(f)$ hold, then $C_k(\varphi, 0) = \delta_{k,0}\mathbb{Z}$ for all $k \geq 0$.*

In the next proposition, we produce the first nontrivial solution for problem (1.1).

Proposition 3.6. *If hypotheses $H(f)$ hold then problem (1.1) has a nontrivial solution $x_0 \in C_0^1(\overline{Z})$ and x_0 is a critical point of φ of mountain pass-type.*

Proof. Recall that $x = 0$ is an isolated local minimum of φ. So we can find $\rho_0 > 0$ such that

$$\varphi|_{\partial B_{\rho_0}} > 0. \tag{3.17}$$

Let $u_1 \in C_0^1(\overline{Z})$ be the $L^2(Z)$-normalized principal eigenfunction of $(-\triangle,\ H_0^1(Z))$ and let $t > 0$. For $0 < \beta_0 < t$, via the mean value theorem, we have

$$F(z, tu_1(z)) = F(z, \beta_0 u_1(z)) + \int_{\beta_0}^t f(z, \mu u_1(z))u_1(z)d\mu \ \text{ a.e. on } \ Z. \tag{3.18}$$

Integrating over Z and using Fubini's theorem, we obtain

$$\int_Z F(z, tu_1(z))dz = \int_Z F(z, \beta_0 u_1(z))dz + \int_{\beta_0}^t \frac{1}{\mu}\int_Z f(z, \mu u_1(z))\mu u_1(z)dzd\mu.$$

Choosing $\beta_0 > 0$ large, because of hypothesis $H(f)(v)$, we have

$$\int_Z f(z,\mu u_1(z))\mu u_1(z)dz \geq \gamma_1 > 0 \ \text{ for all } \ \mu \in [\beta_0, t]. \tag{3.19}$$

From (3.18) and (3.19), we obtain

$$\begin{aligned} &\int_Z F(z,tu_1(z))dz \geq \int_Z F(z,\beta_0 u_1(z))dz + \int_{\beta_0}^t \frac{\gamma_1}{\mu} d\mu \ \text{ for } \ \beta_0 > 0 \ \text{ large,} \\ \Rightarrow &\int_Z F(z,tu_1(z))dz \geq \int_Z F(z,\beta_0 u_1(z))dz + \gamma_1(\ln t - \ln\beta_0). \end{aligned} \tag{3.20}$$

So from (3.20) it follows that

$$-\int_Z F(z,tu_1(z))dz \to -\infty \ \text{ as } \ t \to +\infty.$$

Hence

$$\varphi(tu_1) = -\int_Z F(z,tu_1(z))dz \to -\infty \ \text{ as } \ t \to +\infty \ \text{ (see (2.2)).}$$

Therefore for $t > 0$ large, we have

$$\varphi(tu_1) < \varphi(0) = 0 < \inf_{\partial B_{\rho_0}} \varphi = c.$$

This fact together with Proposition 3.2, permit the use of the mountain pass theorem (see Bartolo-Benci-Fortunato [1]), which gives $x_0 \in H_0^1(Z)$ such that

$$\varphi'(x_0) = 0 \ \text{ and } \ \varphi(0) = 0 < c \leq \varphi(x_0). \tag{3.21}$$

From (3.21), we deduce that $x_0 \neq 0$. From the equality in (3.21), we have

$$\begin{aligned} &A(x_0) = \lambda_1 x_0 + N(x_0), \\ \Rightarrow &-\triangle x_0(z) = \lambda_1 x_0(z) + f(z,x_0(z)) \ \text{ a.e. on } \ Z, \ x_0|_{\partial Z} = 0. \end{aligned}$$

Thus $x_0 \in H_0^1(\overline{Z})$ is a nontrivial solution of problem (1.1) and from regularity theory (see for example Gasinski-Papageorgiou [5], pp.737-738), we have $x_0 \in C_0^1(\overline{Z})$. Let $d = \varphi(x_0)$ and assume without loss of generality that K_d is discrete (otherwise we have a whole sequence of nontrivial solutions for problem (1.1)). Then invoking Theorem 1 of Hofer [6], we can say that $x_0 \in C_0^1(\overline{Z})$ is a critical point of φ which is of mountain pass-type. □

From the description of the critical groups for a critical point of a mountain pass-type (see Chang [4], p.91 and Mawhin-Willem [10], pp.195-196), we have:

Corollary 3.7. *If hypotheses $H(f)$ hold and $x_0 \in C_0^1(\overline{Z})$ is the nontrivial solution of (1.1) obtained in Proposition 3.6, then $C_k(\varphi, x_0) = \delta_{k,1}\mathbb{Z}$ for all $k \geq 0$.*

In the next proposition, we determine the critical groups of φ at infinity.

To do this, we will need the following slight generalization of Lemma 2.4 of Perera-Schechter [12].

Lemma 3.8. *If H is a Hilbert space, $\{\varphi_t\}_{t\in[0,1]}$ is a one-parameter family of $C^1(H)$-functions such that φ_t' and $\partial_t\varphi_t$ are both locally Lipschitz in $u \in H$ and there exists $R > 0$ such that*

$$\begin{aligned}\inf[(1+\|u\|)\|\varphi_t'(u)\| : t \in [0,1], \|u\| > R] > 0 \\ \textit{and} \;\; \inf[\varphi_t(u) : t \in [0,1], \|u\| \leq R] > -\infty,\end{aligned}$$

then $C_k(\varphi_0, \infty) = C_k(\varphi_1, \infty)$ for all $k \geq 0$.

Proof. Let $\xi < \inf[\varphi_t(u) : t \in [0,1], \|u\| \leq R]$. Let $h(t;u)$ $(t \in [0,1], u \in \varphi_0^\xi)$ be the flow generated by the Cauchy problem

$$\dot{h}(t) = -\frac{\partial_t\varphi_t(h(t))}{\|\varphi_t'(h(t))\|^2}\varphi_t'(h(t)) \;\text{ a.e. on }\; \mathbb{R}_+,\; h(0) = u.$$

We have

$$\begin{aligned}&\frac{d}{dt}\varphi_t(h(t)) = \langle\varphi_t'(h(t)), \dot{h}(t)\rangle + \partial_t\varphi_t(h(t)) = 0 \;\text{ for all }\; t \geq 0, \\ \Rightarrow &\varphi_t(h(t)) = \varphi_0(u) \;\text{ for all }\; t \geq 0.\end{aligned}$$

Since $u \in \varphi_0^a$, we have $\varphi_t(h(t)) \leq \xi$ and so $\|h(t)\| > R$ for all $t \geq 0$. This then by virtue of the hypothesis of the lemma, implies that this flow exists for all $t \geq 0$ (see Bartolo-Benci-Fortunato [1]).

It can be reversed, if we replace φ_t with φ_{1-t}. Therefore $h(1)$ is a homeomorphism of φ_0^ξ and φ_1^ξ and so

$$C_k(\varphi_0, \infty) = H_k(H, \varphi_0^\xi) \cong H_k(H, \varphi_1^\xi) = C_k(\varphi_1, \infty).$$

□

Proposition 3.9. *If hypotheses $H(f)(i) \to (v)$ hold, then $C_k(\varphi, \infty) = \delta_{k,1}\mathbb{Z}$ for all $k \geq 0$.*

Proof. Let $0 < \sigma < \lambda_2 - \lambda_1$ and consider the following one-parameter C^2-functions on the Hilbert space $H_0^1(Z)$:

$$\varphi_t(x) = \frac{1}{2}\|Dx\|_2^2 - \frac{\lambda_1 + \sigma}{2}\|x\|_2^2 - t\int_Z (F(z, x(z)) - \sigma x(z))dz \;\text{ for all }\; x \in H_0^1(Z).$$

We claim that we can find $R > 0$ such that

$$\inf[(1+\|u\|)\|\varphi_t'(u)\| : t \in [0,1],\; \|u\| > R] > 0. \tag{3.22}$$

Suppose that this is not possible. Then we can find $t_n \to t \in [0,1]$ and $\|u_n\| \to \infty$ such that $\varphi'_{t_n}(u_n) \to 0$ in $H^{-1}(Z)$ as $n \to \infty$. Let $y_n = \frac{u_n}{\|u_n\|}$, $n \geq 1$. By passing to a suitable subsequence if necessary, we may assume that

$$\begin{aligned} &y_n \xrightarrow{w} y \text{ in } H^{-1}(Z),\ y_n \to y \text{ in } L^2(Z),\ y_n(z) \to y(z) \text{ a.e. on } Z, \\ &\text{and } |y_n(z)| \leq k(z) \text{ for a.a. } z \in Z, \text{ all } n \geq 1, \text{ with } k \in L^2(Z). \end{aligned}$$

We have

$$\begin{aligned} &\left| \langle \frac{\varphi'_{t_n}(u_n)}{\|u_n\|}, v \rangle \right| \leq \varepsilon_n \text{ for all } v \in H_0^(Z), \text{ with } \varepsilon_n \downarrow 0 \text{ (see (3.22))} \\ \Rightarrow &\left| \langle A(y_n), v \rangle - (\lambda_1 + \sigma) \int_Z y_n v dz - t_n \int_Z \frac{N(u_n)}{\|u_n\|} v dz + t_n \sigma \int_Z y_n v dz \right| \leq \varepsilon_n \end{aligned} \tag{3.23}$$

From the proof of Proposition 3.2, we know that

$$\frac{N(u_n)}{\|u_n\|} \xrightarrow{w} h = gy \text{ in } L^2(Z)$$

with $g \in L^\infty(Z)_+$, $0 \leq g(z) \leq \lambda_2 - \lambda_1$ a.e. on Z. Moreover, arguing as in that proof, we can also show that

$$y_n \to y \text{ in } H_0^1(Z), \text{ hence } \|y\| = 1, \text{ i.e. } y \neq 0.$$

So, if we pass to the limit as $n \to \infty$ in (3.23), we obtain

$$\begin{aligned} &\langle A(y), v \rangle = (\lambda_1 + \sigma) \int_Z yv dz + t \int_Z (g + \sigma) yv dz \text{ for all } v \in H_0^1(Z), \\ \Rightarrow &A(y) = (\lambda_1 + (1-t)\sigma + tg) y. \end{aligned} \tag{3.24}$$

As in the proof of Proposition 3.2, we consider three distinct possibilities for the weight function $m = \lambda_1 + (1-t)\sigma + tg \in L^\infty(Z)_+$.

<u>Case 1:</u> $t = 1$ and $g = 0$.

From (3.24), we have

$$\begin{aligned} &A(y) = \lambda_1(y), \\ \Rightarrow &-\triangle y(z) = \lambda_1 y(z) \text{ a.e. on } Z,\ y|_{\partial Z} = 0, \\ \Rightarrow &y \in E(\lambda_1),\ \ y \neq 0. \end{aligned}$$

So, if $u_n = u_n^0 + \widehat{u}_n$ with $u_n^0 \in E(\lambda_1)$, $\widehat{u}_n \in \widehat{H}_2 = E(\lambda_1)^\perp$, $n \geq 1$, then

$$\frac{\|u_n^0\|}{\|u_n\|} \to 1 \text{ as } n \to \infty. \tag{3.25}$$

We have

$$\left|\langle A(u_n), v\rangle - (\lambda_1+\sigma)\int_Z u_n v dz - t_n\int_Z N(u_n)v dz + t_n\sigma\int_Z u_n v dz\right| \leq \varepsilon_n$$

for all $v \in H_0^1(Z)$.

Let $v = u_n^0 \in E(\lambda_1)$. We obtain

$$\left|\|Du_n^0\|_2^2 - (\lambda_1+\sigma)\|u_n^0\|_2^2 - t_n\int_Z f(z,u_n(z))u_n^0(z)dz + t_n\sigma\|u_n^0\|_2^2\right| \leq \varepsilon_n. \tag{3.26}$$

Since $u_n^0 \in E(\lambda_1)$, we know that $\|Du_n^0\|_2^2 = \lambda_1\|u_n^0\|_2^2$. Also because of (3.25) and hypothesis $H(f)(v)$, we have

$$\int_Z f(z,u_n(z))u_n^0(z)dz \geq \gamma_1 \text{ for all } n \geq n_1.$$

Then from (3.26), we obtain

$$\begin{aligned}&(1-t_n)\sigma\|u_n^0\|_2^2 + t_n\gamma_1 \leq \varepsilon_n \text{ for all } n \geq n_1,\\ \Rightarrow & t_n\gamma_1 \leq \varepsilon_n \text{ for all } n \geq n_1.\end{aligned}$$

Since $t_n \to t = 1$ and $\varepsilon_n \downarrow 0$, in the limit as $n \to \infty$, we obtain

$$0 < \gamma_1 \leq 0,$$

a contradiction.

Case 2: $t = 1$ and $g = \lambda_2 - \lambda_1$.

From (3.24), we have

$$\begin{aligned}&A(y) = \lambda_2 y,\\ \Rightarrow & -\triangle y(z) = \lambda_2 y(z) \text{ a.e. on } Z, \;\; y|_{\partial Z} = 0,\\ \Rightarrow & y \in E(\lambda_2), \;\; y \neq 0.\end{aligned}$$

Now we write $u_n = u_n^0 + \widehat{u}_n$ with $u_n^0 \in E(\lambda_2)$ and $\widehat{u}_n \in W = E(\lambda_2)^\perp$. We have

$$\frac{\|u_n^0\|}{\|u_n\|} \to 1 \text{ as } n \to \infty. \tag{3.27}$$

Recall that

$$\left|\langle A(u_n), v\rangle - (\lambda_1+\sigma)\int_Z u_n v dz - t_n\int_Z N(u_n)v dz + t_n\sigma\int_Z u_n v dz\right| \leq \varepsilon_n$$

for all $v \in H_0^1(Z)$.

Let $v = u_n^0 \in E(\lambda_2)$. We obtain

$$\begin{aligned}&\Big| \|Du_n^0\|_2^2 - t_n\lambda_2\|u_n^0\|_2^2 - (1-t_n)(\lambda_1+\sigma)\|u_n^0\|_2^2 \\ &\quad - t_n \textstyle\int_Z (f(z,u_n(z)) - (\lambda_2-\lambda_1)u_n(z))u_n^0(z)dz \Big| \le \varepsilon_n.\end{aligned} \tag{3.28}$$

Note that $t_n\lambda_2 + (1-t_n)(\lambda_1+\sigma) < \lambda_2$ and so

$$0 < \|Du_n^0\|_2^2 - (t_n\lambda_2 + (1-t_n)(\lambda_1+\sigma))\|u_n^0\|_2^2. \tag{3.29}$$

In addition because of (3.27) and hypothesis $H(f)(v)$, we have

$$\int_Z (f(z,u_n(z)) - (\lambda_2-\lambda_1)u_n(z))u_n^0(z)dz \le -\gamma_2 < 0 \ \text{ for all } \ n \ge n_2. \tag{3.30}$$

Using (3.29) and (3.30) in (3.28), we obtain

$$t_n\gamma_2 \le \varepsilon_n \ \text{ for all } \ n \ge n_2.$$

Passing to the limit as $n \to \infty$ and recalling that $t_n \to 1$ and $\varepsilon \downarrow 0$, we get

$$0 < \gamma_2 \le 0,$$

again a contradiction.

Case 3: $t \ne 1$ or $0 \le g(z) \le \lambda_2 - \lambda_1$ a.e. on Z with $g \ne 0$ and $g \ne \lambda_2 - \lambda_1$.

From (3.24), we have

$$\begin{aligned}&A(y) = (\lambda_1 + \widehat{\xi})y,\ y \ne 0 \ \text{ with } \ \widehat{\xi} = (1-t)\sigma + tg \in L^\infty(Z)_+, \\ &\Rightarrow -\triangle y(z) = (\lambda_1 + \widehat{\xi}(z))y(z) \ \text{ a.e. on } \ Z,\ \ y|_{\partial Z} = 0.\end{aligned} \tag{3.31}$$

Note that since $t \ne 1$ or ($g \ne 0$ and $g \ne \lambda_2 - \lambda_1$), we have

$$\lambda_1 \le \lambda_1 + \widehat{\xi}(z) \le \lambda_2 \ \text{ a.e. on } \ Z,\ \ \lambda_1 \ne \lambda_1 + \widehat{\xi} \ \text{ and } \ \lambda_2 \ne \lambda_1 + \widehat{\xi}.$$

Hence from the strict monotonicity of the eigenvalues on the weight function, we infer that

$$\widehat{\lambda}_1(\lambda_1 + \widehat{\xi}) < \widehat{\lambda}_1(\lambda_1) = 1 \ \text{ and } \ \widehat{\lambda}_2(\lambda_1 + \widehat{\xi}). \tag{3.32}$$

Using (3.32) in (3.31), we infer that $y = 0$, a contradiction to the fact that $\|y\| = 1$.

So in all three cases we have reached a contradiction and this means that there exists $R > 0$ for which (3.22) is valid.

Also it is clear, that due to hypotheses $H(f)(iii), (iv)$, we have

$$\inf[\varphi_t(u) : t \in [0,1],\ \|u\| \le R] > -\infty.$$

So we can apply Lemma 3.8 and have that

$$C_k(\varphi_0, \infty) = C_k(\varphi, \infty) \text{ for all } k \geq 0. \tag{3.33}$$

Note that

$$\varphi_0(x) = \frac{1}{2}\|Dx\|_2^2 - \frac{\lambda_1 + \sigma}{2}\|x\|_2^2 \text{ and } \varphi_1(x) = \varphi(x) \text{ for all } x \in H_0^1(Z).$$

Since $0 < \sigma < \lambda_2 - \lambda_1$, the only critical point of φ_0 is $u = 0$. Hence

$$C_k(\varphi_0, \infty) = C_k(\varphi, 0) \text{ for all } k \geq 0. \tag{3.34}$$

Moreover, from Proposition 2.3 of Su [16], we have

$$C_k(\varphi_0, 0) = \delta_{k,1}\mathbb{Z} \text{ for all } k \geq 0. \tag{3.35}$$

From (3.33), (3.34) and (3.35), we conclude that

$$C_k(\varphi, \infty) = \delta_{k,1}\mathbb{Z} \text{ for all } k \geq 0.$$

□

Now we are ready for the first multiplicity theorem.

Theorem 3.10. *If hypotheses $H(f)$ hold, then problem (1.1) has at least two nontrivial solutions $x_0, v_0 \in C_0^1(\overline{Z})$.*

Proof. One nontrivial solution $x_0 \in C_0^1(\overline{Z})$, exists by virtue of Proposition 3.6.

Suppose that $\{0, x_0\}$ are the only critical points of φ. Then using Corollaries 3.5, 3.7, 3.9 and the Poincare-Hopf formula, we have

$$(-1)^0 + (-1)^1 = (-1)^1,$$

a contradiction. So there exists a third critical point $v_0 \neq x_0$, $v_0 \neq 0$. Evidently v_0 is a solution of (1.1) and by regularity theory, we have $v_0 \in C_0^1(\overline{Z})$. □

We have another multiplicity result by modifying hypothesis $H(f)(vi)$. So the new hypotheses on the nonlinearity $f(z, x)$ are the following:

$\underline{H(f)'}$: $f : Z \times \mathbb{R} \to \mathbb{R}$ is a function such that $f(z, 0) = 0$ a.e. on Z, hypotheses $H(f)'(i) \to (v)$ are the same as hypotheses $H(f)(i) \to (v)$ respectively and

(vi) there exist $m \geq 2$ and $\delta > 0$ such that

$$\lambda_m - \lambda_1 \leq \frac{f(z, x)}{x} \leq \lambda_{m+1} - \lambda_1 \text{ for a.a. } z \in Z \text{ and all } 0 < |x| \leq \delta.$$

Remark 3.11. *Hypotheses $H(f)'(iv)$ and (vi) imply that we can have double resonance both at infinity and at zero. A double-double resonance situation.*

Theorem 3.12. *If hypotheses $H(f)'$ hold, then problem (1.1) has at least two nontrivial solutions $x_0, v_0 \in C_0^1(\overline{Z})$.*

Proof. Because of hypothesis $H(f)'(vi)$ and Proposition 1.1 of Li-Perera-Su [9], we have

$$C_k(\varphi, 0) = \delta_{k,d}\mathbb{Z}, \tag{3.36}$$

where d =sum of multiplicities of $\{\lambda_k\}_{k=1}^m = \dim \overline{H}_m \geq 2$, since $m \geq 2$.

Also from Proposition 3.9, we know that

$$C_k(\varphi, \infty) = \delta_{k,1}\mathbb{Z}. \tag{3.37}$$

So there exists a critical point x_0 of φ such that

$$C_1(\varphi, x_0) \neq 0. \tag{3.38}$$

Comparing this with (3.36), we infer that $x_0 \neq 0$. Moreover, due to (3.38) x_0 is of mountain pass type and so

$$C_1(\varphi, x_0) = \delta_{k,1}\mathbb{Z}. \tag{3.39}$$

If $\{0, x_0\}$ are the only critical points of φ, then from (3.36), (3.37) and (3.39) and the Poincare-Hopf formula, we have

$$(-1)^d + (-1)^1 = (-1)^1,$$
$$\Rightarrow (-1)^d = 0, \quad \text{a contradiction.}$$

So there exists a second nontrivial critical point v_0 of φ. Evidently $x_0, v_0 \in H_0^1(Z)$ are nontrivial solutions of problem (1.1). From regularity theory, we conclude that $x_0, v_0 \in C_0^1(\overline{Z})$. □

Remark 3.13. *Theorem 3.12 above partially extends Theorem 3 of Robinson [14] and also Theorem 2 of Su [16].*

Received: February 2008. Revised: April 2008.

References

[1] P. Bartolo, V. Benci and D. Fortunato, *Abstract critical point theorems to some nonlinear problems with strong resonance at infinity*, Nonlin. Anal., **7** (1983), 981–1012.

[2] H. Berestycki and D. De Figueiredo, *Double resonance and semilinear elliptic problems*, Comm. PDE, **6** (1981), 91–120.

[3] H. Brezis and L. Nirenberg, *H^1 versus C^1 local minimizers*, CRAS Paris, t. **317** (1993), 465–472.

[4] K-C. Chang, *Infinite Dimensional Morse Theory and Multiple Solution Problems*, Boston, (1993).

[5] L. Gasinski and N.S. Papageorgiou, *Nonlinear Analysis*, Chapman and Hall/CRC Press, Boca Raton, (2006).

[6] H. Hofer, *A note on the topological degree at a critical point of mountain-pass type*, Proc. AMS, **90** (1984), 309–315.

[7] E. Landesman and A. Lazer, *Nonlinear perturbations of linear elliptic boundary value problems*, J. Math. Mech., **19** (1969/1970), 609–623.

[8] E. Landesman, S. Robinson and A. Rumbos, *Multiple solutions of semilinear elliptic problems at resonance*, Nonlin. Anal., **24** (1995), 1049–1059.

[9] S-J. Li, K. Perera and J-B. Su, *Computation of critical groups in elliptic boundary-value problems where the asymptotic limits may not exist*, Proc. Royal Soc. Edin, **131A** (2001), 721–732.

[10] J. Mawhin and M. Willem, *Critical Point Theory and Hamiltonian Systems*, Springer-Verlag, New York, (1989).

[11] M. Nkashama, *Density condition at infinity and resonance in nonlinear elliptic partial differential equations*, Nonlin. Anal., **22** (1994), 251–265.

[12] K. Perera and M. Schechter, *Solution of nonlinear equations having asymptotic limits at zero and infinity*, Calc. Var., **12** (2001), 359–369.

[13] S. Robinson, *Double resonance in semilinear elliptic boundary value problems over bounded and unbounded domains*, Nonlin. Anal., **21** (1993), 407–424.

[14] S. Robinson, *Multiple solutions for semilinear elliptic boundary value problems at resonance*, Electr. J. Diff. Eqns, No.**1** (1995), pp. 14.

[15] A. Rumbos, *A semilinear elliptic boundary value problem at resonance where the nonlinearity may grow linearly*, Nonlin. Anal., **16** (1991), 1159–1168.

[16] J-B. Su, *Semilinear elliptic boundary value problems with double resonance between two consecutive eigenvalues*, Nonlin. Anal., **48** (2002), 881–895.

[17] J-B.Su and C.L. Tang, *Multiplicity results for semilinear eliptic equations with resonance at higher eigenvalues*, Nonlin. Anal., **44** (2001), 311–321.

CUBO A Mathematical Journal
Vol.10, Nº 03, (43–55). October 2008

The Modulo Two Homotopy Groups of the L_2-Localization of the Ravenel Spectrum

Ippei Ichigi and Katsumi Shimomura
Department of Mathematics, Faculty of Science,
Kochi University, Kochi, 780-8520, Japan
email: `95sm004@math.kochi-u.ac.jp`
email: `katsumi@math.kochi-u.ac.jp`

ABSTRACT

The Ravenel spectra $T(m)$ for non-negative integers m interpolate between the sphere spectrum and the Brown-Peterson spectrum. Let L_2 denote the Bousfield-Ravenel localization functor with respect to $v_2^{-1}BP$. In this paper, we determine the homotopy groups $\pi_*(L_2T(m):\mathbb{Z}/2) = [M_2, L_2T(m)]_*$ for $m > 1$, where M_2 denotes the modulo two Moore spectrum.

RESUMEN

El espectro de Ravenel $T(m)$ para enteros no negativos m interpola entre el espectro esferico y el espectro de Brown-Peterson. Denotemos por L_2 el funtor de localización de Bousfield-Ravenel con respecto a $v_2^{-1}BP$. En este artículo, determinamos el grupo de homotopia $\pi_*(L_2T(m):\mathbb{Z}/2) = [M_2, L_2T(m)]_*$ para $m > 1$, donde M_2 denota el espectro de Moor modulo dos.

Key words and phrases: *homotopy groups, Bousfield-Ravenel localization, Ravenel spectrum.*

Math. Subj. Class.: *55Q99, 55Q51, 20J06.*

1 Introduction

Let $\mathcal{S}_{(2)}$ denote the stable homotopy category of 2-local spectra, and $BP \in \mathcal{S}_{(2)}$ denote the Brown-Peterson ring spectrum. Then, $BP_* = \pi_*(BP) = \mathbb{Z}_{(2)}[v_1, v_2, \dots]$ and $BP_*(BP) = \pi_*(BP \wedge BP) = BP_*[t_1, t_2, \dots]$, which form a Hopf algebroid. The Adams-Novikov spectral sequence for computing the homotopy groups $\pi_*(X)$ of a spectrum X has the E_2-term $E_2^*(X) = \mathrm{Ext}^*_{BP_*(BP)}(BP_*, BP_*(X))$. Let $L_2 \colon \mathcal{S}_{(2)} \to \mathcal{S}_{(2)}$ be the Bousfield-Ravenel localization functor with respect to $v_2^{-1}BP$. Then, the E_2-term $E_2^*(L_2S^0)$ for the sphere spectrum S^0 is determined in [12], but the homotopy groups $\pi_*(L_2S^0)$ stay undetermined. The Ravenel spectrum $T(m)$ for $m > 0$ is a ring spectrum characterized by $BP_*(T(m)) = BP_*[t_1, t_2, \dots t_m] \subset BP_*(BP)$ as a $BP_*(BP)$-comodule. The spectrum $T(m)$ interpolates between the sphere spectrum and the Brown-Peterson spectrum, and so the homotopy groups $\pi_*(L_2T(m))$ seem accessible if m is sufficiently large. Indeed, $\pi_*(L_2T(\infty)) = \pi_*(L_2BP)$ is determined by Ravenel [8]. Let M_k denote the mod k Moore spectrum defined by the cofiber sequence

$$S^0 \xrightarrow{2} S^0 \xrightarrow{i} M_k \xrightarrow{j} S^1. \tag{1.1}$$

For $m = 1$, $T(1) \wedge M_2$ is the Mahowald spectrum $X\langle 1\rangle$ and the homotopy groups of $L_2X\langle 1\rangle$ are determined in [11]. But even the homotopy groups of $L_2T(1) \wedge M_4$ are too complicated to be determined completely (*cf.* [2], [3]). Consider a spectrum $T(m)/(v_1^a)$ defined as a cofiber of the self-map $v_1^a \colon \Sigma^{2a}T(m) \to T(m)$ defined by the generator $v_1 \in \pi_2(T(m))$. We use the notation:

$$V_m(0) = T(m) \wedge M_2 \quad \text{and} \quad V_m(1)_a = T(m)/(v_1^a) \wedge M_2, \tag{1.2}$$

and abbreviate $V_m(1)_1$ to $V_m(1)$. In this paper, we consider the case where $m > 1$, and determine $\pi_*(L_2V_m(1))$ and $\pi_*(L_2V_m(0))$. The Adams-Novikov E_2-term $E_2^*(L_2V_m(1))$ for $m > 1$ is determined by Ravenel [10] as follows:

$$E_2^*(L_2V_m(1)) = K_m(2)_* \otimes \wedge(h_{1,0}, h_{1,1}, h_{2,0}, h_{2,1}) \tag{1.3}$$

for generators $h_{i,j} \in E_2^{1,2^{m+i+j+1}-2^{j+1}}(L_2V_m(1))$ and $K_m(2)_* = v_2^{-1}\mathbb{Z}/2[v_2, v_3, \dots, v_{m+2}]$. We show that $V_m(1)$ is a $T(m)$-module spectrum with M_2-action, and then that all additive generators of the E_2-term are permanent cycles and the extension problem of the spectral sequence is trivial.

Theorem 1.4. *$\pi_*(L_2V_m(1)) = K_m(2)_* \otimes \wedge(h_{1,0}, h_{1,1}, h_{2,0}, h_{2,1})$ as a $\mathbb{Z}/2$-module.*

Let $\alpha \colon \Sigma^8M_2 \to M_2$ denote the Adams map such that $BP_*(\alpha) = v_1^4$, and K_2^a denote a cofiber of α^a. Then, we show that $V_m(1)_{4a} = T(m) \wedge K_2^a$ in Lemma 2.4 and denote the telescope of $V_m(1)_4 \xrightarrow{\alpha} \cdots \xrightarrow{\alpha} V_m(1)_{4a} \xrightarrow{\alpha} V_m(1)_{4a+4} \xrightarrow{\alpha} \cdots$ by $V_m(1)_\infty$. By the v_1-Bockstein spectral sequence, we determine the Adams-Novikov E_2-term $E_2^*(L_2V_m(1)_\infty)$, whose structure is given in [4] without proof. Here we give a proof of it. Consider the integers e_n and a_n defined by

$$e_n = \frac{8^n - 1}{7} \quad \text{and} \quad a_n = \begin{cases} 1 & n = 0 \\ 3e_{k+1} - 1 & n = 3k+1 \\ 6e_{k+1} & n = 3k+2 \\ 12e_{k+1} & n = 3k+3. \end{cases} \tag{1.5}$$

We introduce modules

$$\begin{aligned} E_m(2)_* &= v_2^{-1}\mathbb{Z}_{(2)}[v_1, v_2, \dots, v_{m+2}], \\ Q(k) &= E_{m-1}(2)_*/(2, v_1^{a_k})[x_{k+1}]\langle x_k/v_1^{a_k}\rangle, \end{aligned}$$

where $x_n \in E_m(2)_*$ is an element defined in (4.1) such that $x_n \equiv v_{m+2}^{2^n}$ modulo $(2, v_1)$, and $x_n/v_1^{a_n} \in E_2^0(L_2V_m(1)_\infty)$ by Proposition 4.3. We also introduce homology classes ζ and ζ_n of $E_2^1(V_m(0))$, which correspond to elements $v_{m+2}h_{1,1}$ and $v_{m+2}^{2^l e_k}\zeta_l \in E_2^1(L_2V_m(1))$ for $n = 3k+l$ with $l \in \{1, 2, 3\}$, respectively, where ζ_l corresponds to $h_{1,0}$ if $l = 1$, and $h_{2,l-2}$ if $l = 2, 3$.

Proposition 1.6. (*cf.* [4]) *The E_2-term of Adams-Novikov spectral sequence for computing $\pi_*(L_2V_m(1)_\infty)$ is isomorphic to the direct sum of $Q(0) \otimes \wedge(h_{1,0}, h_{2,0}, h_{2,1})$ and the tensor product of $\wedge(\zeta)$ and*

$$E_{m-1}(2)_*/(2, v_1^\infty) \oplus \bigoplus_{k>0} Q(k) \otimes \wedge(\zeta_{k+1}, \zeta_{k+2})$$

as a $\mathbb{Z}/2[v_1]$-module.

By noticing that $x_n \in E_2^0(L_2V_m(1)_{a_n})$ survives to $\pi_*(L_2V_m(1)_{a_n})$ in Lemma 5.1, we see that all additive generators of Proposition 1.6 are permanent cycles.

Theorem 1.7. *The homotopy groups $\pi_*(L_2V_m(1)_\infty)$ are isomorphic to the Adams-Novikov E_2-term given in Proposition 1.6.*

Consider the cofiber sequence

$$V_m(0) \xrightarrow{\eta} v_1^{-1}V_m(0) \xrightarrow{p} V_m(1)_\infty \longrightarrow \Sigma V_m(0) \tag{1.8}$$

for the localization map η. Here, we introduce algebras

$$k_m(1)_* = \mathbb{Z}/2[v_1, v_2, \dots, v_{m+1}] \quad \text{and} \quad K_m(1)_* = v_1^{-1}k_m(1)_*.$$

Ravenel showed the following

Proposition 1.9. (*cf.* [10]) *The homotopy groups $\pi_*(v_1^{-1}V_m(0))$ are isomorphic to $K_m(1)_* \otimes \wedge(h_{1,0})$.*

There is a relation between $h_{1,0}$ and ζ, which is shown in section four:

Lemma 1.10. *The induced homomorphism p_* from p in (1.8) assigns $h_{1,0}/v_1^j \in E_2^1(v_1^{-1}V_m(0))$ to $\zeta/v_1^{j-2} \in E_2^1(L_2V_m(1)_\infty)$.*

Observing the correspondence in the Adams-Novikov E_2-terms, we obtain

Corollary 1.11. *The homotopy groups $\pi_*(L_2V_m(0))$ are isomorphic to the direct sum of $\Sigma^{-1}Q(0)\otimes \wedge(h_{1,0}, h_{2,0}, h_{2,1})$ and the tensor product of $\wedge(\zeta)$ and*

$$k_m(1)_* \oplus \Sigma^{-1}k_m(1)_*/(2, v_1^\infty, v_2^\infty) \oplus \bigoplus_{k>0} \Sigma^{-1}Q(k) \otimes \wedge(\zeta_{k+1}, \zeta_{k+2})$$

as a $\mathbb{Z}/2[v_1]$-module.

In the next section, we observe about an action of the Moore spectrum M_2 on $V_m(1)_t$ and a ring structure of $V_m(1)_{4t}$, in order to study the Adams-Novikov differential and the extension problem of the spectral sequence in the following sections. We prove Theorem 1.4 in section three. Section four is devoted to show Proposition 1.6. We end by proving Theorem 1.7 in the last section.

2 The spectrum $T(m) \wedge K_k^t$

We work in the stable homotopy category of spectra localized at the prime two. Let BP denote the Brown-Peterson spectrum. Then, we have the Adams-Novikov spectral sequence

$$E_2^{s,t}(X) = \operatorname{Ext}_\Gamma^{s,t}(A, BP_*(X)) \Longrightarrow \pi_*(X).$$

Here (A, Γ) is the associated Hopf algebroid such that

$$(A, \Gamma) \;=\; (BP_*,\; BP_*(BP)) \;=\; (\mathbb{Z}_{(2)}[v_1, v_2, \dots], BP_*[t_1, t_2, \dots])$$

for the Hazewinkel generators $v_k \in BP_{2^{k+1}-2}$ and the generators $t_k \in BP_{2^{k+1}-2}(BP)$.

Let M_k and K_k^t for $k = 2, 4$ and $t > 0$ denote spectra defined by the cofiber sequences

$$S^0 \xrightarrow{2} S^0 \xrightarrow{i} M_k \xrightarrow{j} S^1 \quad \text{and} \quad \Sigma^{8t}M_k \xrightarrow{\alpha^t} M_k \xrightarrow{i_k^t} K_k^t \xrightarrow{j_k^t} \Sigma^{8t+1}M_k.$$

Here α denotes the Adams map such that $BP_*(\alpha) = v_1^4$. Note that M_4 and K_4^t are ring spectra (*cf.* [5]). The Ravenel spectrum $T(m)$ is characterized by $BP_*(T(m)) = A[t_1, \dots, t_m] \subset \Gamma$ as Γ-comodules, and is a ring spectrum, whose multiplication and unit map we denote by μ and ι, respectively. Throughout the paper, we fix a positive integer m. Let $(A, \Gamma_m) = (A, \Gamma/(t_1, t_2, \dots, t_m))$ be the Hopf algebroid associated with (A, Γ), and consider a spectrum X such that $BP_*(X) = M \otimes_A A[t_1, \dots, t_m]$ for a Γ-comodule M. Then, we have an isomorphism

$$E_2^*(X) = \operatorname{Ext}_{\Gamma_m}^*(A, M) \tag{2.1}$$

by the change of rings theorem (*cf.* [10]). By observing the reduced cobar complex for the Ext group, we have

Lemma 2.2. *The E_2-term has the vanishing line of the slope $1/(q_m - 1)$ if M is (-1)-connected.*

Hereafter, we put

$$q_m = 2^{m+2} - 2 \tag{2.3}$$

which is the degree of $u_1 = v_{m+1}$ and $s_1 = t_{m+1}$. This shows $\pi_2(T(m)) = BP_2 = \mathbb{Z}_{(2)}\{v_1\}$ if $m > 0$. Let $T(m)/(v_1^a)$ for an integer $a > 0$ denote the cofiber of $\widetilde{v}_1^a \colon \Sigma^{8a}T(m) \to T(m)$, where $\widetilde{v}_1 : \Sigma^8 T(m) \to T(m)$ is the composite

$$\widetilde{v}_1 : \quad \Sigma^8 T(m) = S^8 \wedge T(m) \xrightarrow{v_1 \wedge T(m)} T(m) \wedge T(m) \quad \xrightarrow{\mu} \quad T(m).$$

Lemma 2.4. *For $k = 2, 4$ and $a > 0$, $T(m)/(v_1^{4a}) \wedge M_k = T(m) \wedge K_k^a$. In particular, $T(m) \wedge K_2^a \wedge M_4 = T(m)/(v_1^{4a}) \wedge M_2 \wedge M_4 = T(m) \wedge M_2 \wedge K_4^a$.*

Proof. Since $\pi_8(T(m) \wedge M_k) = BP_8/(k) = \mathbb{Z}/k\{v_1^4, v_1 v_2\}$ by Lemma 2.2, we see that $v_1^4 \wedge M_k = \iota \wedge \alpha i \in \pi_8(T(m) \wedge M_k)$. Indeed, both of these elements are assigned to $v_1^4 \in BP_8(T(m) \wedge M_i)$ under the homomorphism induced from the unit map of BP. It extends to $v_1^4 \wedge M_k = \iota \wedge \alpha : M_k \to T(m) \wedge M_k$, since $[M_k, T(m) \wedge M_k]_8 = \pi_8(T(m) \wedge M_k)$. Indeed, $\pi_9(T(m) \wedge M_k) = BP_9/(k) = 0$. We further extend it to a self-map $A = \widetilde{v}_1^4 \wedge M_k = T(m) \wedge \alpha : T(m) \wedge M_k \to T(m) \wedge M_k$ by the ring structure of $T(m)$. Now the cofiber of A^a is $T(m)/(v_1^{4a}) \wedge M_k = T(m) \wedge K_k^a$.

□

This lemma implies

$$V_m(1)_{4a} = T(m) \wedge K_2^a \tag{2.5}$$

for the spectrum $V_m(1)_{4a}$ in (1.2).

Lemma 2.6. *Let F denote one of the spectra M_k and K_k^a for $k = 2, 4$ and $a > 0$. Then, there is a pairing $\nu_F : F \wedge F \to T(m) \wedge F$ such that $\nu_F \circ (F \wedge i_F) = \iota \wedge F \colon F \to T(m) \wedge F$ for $m > 0$. Here $i_F : S^0 \to F$ denotes the inclusion to the bottom cell.*

Proof. The pairing for $F = M_4$ or K_4^a is the composite $(\iota \wedge F \wedge F)(T(m) \wedge \mu_F)$ for the multiplication μ_F of the ring spectrum of F (see [5]).

For $F = M_2$, we see that $\pi_0(T(m) \wedge M_2) = BP_0/(2) = \mathbb{Z}/2$ and $\pi_1(T(m) \wedge M_2) = BP_1/(2) = 0$ by Lemma 2.2, and so $[M_2, T(m) \wedge M_2]_0 = \mathbb{Z}/2$.

Note that $M_2 \wedge M_4 = M_2 \vee \Sigma M_2$. Then, by Lemma 2.4,

$$\begin{aligned} T(m) \wedge M_2 \wedge K_4^a &= T(m)/(v_1^{4a}) \wedge M_2 \wedge M_4 = T(m)/(v_1^{4a}) \wedge (M_2 \vee \Sigma M_2) \\ &= T(m)/(v_1^{4a}) \wedge M_2 \vee \Sigma T(m)/(v_1^{4a}) \wedge M_2 = T(m) \wedge K_2^a \vee \Sigma T(m) \wedge K_2^a. \end{aligned}$$

We also see that $T(m) \wedge K_2^a \wedge K_4^a = T(m)/(v_1^{4a}) \wedge K_2^a \wedge M_4 = T(m)/(v_1^{4a}) \wedge (K_2^a \vee \Sigma K_2^a)$, and so $T(m) \wedge K_2^a \wedge K_4^a \wedge M_2 = T(m) \wedge K_2^a \wedge K_2^a \vee \Sigma T(m) \wedge K_2^a \wedge K_2^a$. Then,

$$\begin{aligned} T(m) \wedge M_2 \wedge K_4^a \wedge K_4^a \wedge M_2 &= T(m) \wedge K_2^a \wedge K_4^a \wedge M_2 \vee \Sigma T(m) \wedge K_2^a \wedge K_4^a \wedge M_2 \\ &= T(m) \wedge K_2^a \wedge K_2^a \vee \Sigma T(m) \wedge K_2^a \wedge K_2^a \vee \Sigma T(m) \wedge K_2^a \wedge K_2^a \wedge M_2. \end{aligned}$$

Let $\mu_K : K_4^a \wedge K_4^a \to K_4^a$ denote the multiplication of the ring spectrum K_4^a, and $\widetilde{\nu}$ be the composite $T(m) \wedge M_2 \wedge M_2 \xrightarrow{T(m)\wedge\nu_{M_2}} T(m) \wedge T(m) \wedge M_2 \xrightarrow{\mu\wedge M_2} T(m) \wedge M_2$. Then the desired pairing is a composite

$$K_2^a \wedge K_2^a \xrightarrow{\iota\wedge K\wedge K} T(m) \wedge K_2^a \wedge K_2^a \xrightarrow{inc\wedge K_2^a} T(m) \wedge M_2 \wedge K_4^a \wedge K_4^a \wedge M_2 \xrightarrow{switch}$$

$$T(m) \wedge M_2 \wedge M_2 \wedge K_4^a \wedge K_4^a \xrightarrow{\widetilde{\nu}} T(m) \wedge M_2 \wedge K_4^a \wedge K_4^a \xrightarrow{T(m)\wedge M_2\wedge\mu_K} T(m) \wedge M_2 \wedge K_4^a \xrightarrow{prj} T(m) \wedge K_2^a.$$

□

Corollary 2.7. *The spectra $V_m(0)$ and $V_m(1)_{4a}$ for $a > 0$ are ring spectra.*

We say that a spectrum X has M_2-action, if there is a pairing $\varphi_X : X \wedge M_2 \to X$ such that $\varphi_X(X \wedge i) = id_X$. Here $i : S^0 \to M_2$ is the inclusion of (1.1) and $id_X : X \to X$ denotes the identity map.

Lemma 2.8. *$V_m(1)_t$ has M_2-action.*

Proof. Since $T(m)$ is an associative ring spectrum, $T(m)/(v_1^t)$ is a $T(m)$-module spectrum. The action $\varphi_{V_m(1)_t}$ is defined by the composite $V_m(1)_t \wedge M_2 = T(m)/(v_1^t) \wedge M_2 \wedge M_2 \xrightarrow{T(m)/(v_1^t)\wedge\nu_{M_2}}$ $T(m)/(v_1^t) \wedge T(m) \wedge M_2 \longrightarrow T(m)/(v_1^t) \wedge M_2 = V_m(1)_t$. □

Since $V_m(1)_t$ is a $T(m)$-module spectrum, it implies the following

Corollary 2.9. *$V_m(1)_t$ is a $V_m(0)$-module spectrum.*

3 The homotopy groups of $L_2V_m(1)$

Note that if $BP_*(X)$ is $(2, v_1)$-nil, then $BP_*(L_2X) = v_2^{-1}BP_*(X)$, since L_2 is smashing (*cf.* [8], [9]). Therefore, the Adams-Novikov E_2-term $E_2^*(L_2V_m(1)_t)$ is $\mathrm{Ext}_\Gamma^*(A, v_2^{-1}BP_*/(2, v_1^t)[t_1, \ldots, t_m])$, which is isomorphic to

$$E_2^*(L_2V_m(1)_t) = \mathrm{Ext}_{\Gamma_m}^*(A, v_2^{-1}BP_*/(2, v_1^t))$$

by (2.1). Consider a spectrum

$$E_m(2) = v_2^{-1}BP\langle m+2\rangle$$

for the Johnson-Wilson spectrum $BP\langle m+2\rangle$. Then we obtain a Hopf algebroid

$$(E_m(2)_*, \Sigma_m(2)) = (v_2^{-1}\mathbb{Z}_{(2)}[v_1, v_2, \dots, v_{m+2}], E_m(2)_* \otimes_A \Gamma_m \otimes_A E_m(2)_*).$$

Since

$$v_2^{-1}BP_*/J \xrightarrow{1\otimes\eta_R} E_m(2)_*/J \otimes_A \Gamma_m$$

for an invariant regular ideal $J = (2^b, v_1^a)$ is a faithfully flat extension, we have an isomorphism

$$\mathrm{Ext}^*_{\Gamma_m}(A, BP_*/J) \cong \mathrm{Ext}^*_{\Sigma_m(2)}(E_m(2)_*, E_m(2)_*/J)$$

by a theorem of Hopkins' (*cf.* [1, Th. 3.3]). Note that $m+2$ is the smallest number n, for which $v_2^{-1}BP_*/J \xrightarrow{1\otimes\eta_R} v_2^{-1}BP\langle n\rangle_*/J \otimes_A \Gamma_m$ is a faithfully flat extension. We use the abbreviation

$$H^*M = \mathrm{Ext}^*_{\Sigma_m(2)}(E_m(2)_*, M) \tag{3.1}$$

for a $\Sigma_m(2)$-comodule M. We compute the Ext group H^*M by the reduced cobar complex $\widetilde{\Omega}^*_{\Sigma_m(2)}M$ (*cf.* [10]). Since the differentials of the cobar complex are defined by the right unit $\eta_R : E_m(2)_* \to \Sigma_m(2)$ and the diagonal $\Delta : \Sigma_m(2) \to \Sigma_m(2) \otimes_{E_m(2)_*} \Sigma_m(2)$, we write down here some formulas on them based on the Hazewinkel and the Quillen formulas:

$$\begin{array}{l} v_n = 2\ell_n - \sum_{k=1}^{n-1} \ell_k v_{n-k}^{2^k} \in \mathbb{Q}\otimes A = \mathbb{Q}[\ell_1, \ell_2, \dots], \\ \eta_R(\ell_n) = \sum_{k=0}^{n} \ell_k t_{n-k}^{2^k} \in \mathbb{Q}\otimes\Gamma = \mathbb{Q}\otimes A[t_1, t_2, \dots] \quad \text{and} \\ \sum_{i+j=n} \ell_i \Delta(t_j^{2^i}) = \sum_{i+j+k=n} \ell_i t_j^{2^i} \otimes t_k^{2^{i+j}} \in \mathbb{Q}\otimes\Gamma\otimes_A\Gamma. \end{array} \tag{3.2}$$

Hereafter, we put $v_2 = 1$ and use the following notation:

$$u_i = v_{m+i} \quad \text{and} \quad s_i = t_{m+i}.$$

Since the structure maps preserve degrees, we may recover v_2's from its degrees. Then, we obtain the following two lemmas immediately from (3.2) by a routine computation:

Lemma 3.3. *The right unit $\eta_R : A \to \Gamma_m/(2)$ acts as follows:*

$$\begin{array}{rcl} \eta_R(v_n) &=& v_n \quad \textit{for } n \le m+1, \\ \eta_R(u_2) &=& u_2 + v_1 s_1^2 + v_1^{2^{m+1}} s_1, \\ \eta_R(u_3) &\equiv& u_3 + s_1^4 + s_1 + v_1 r_1 \mod (2, v_1^{2^{m+2}}), \\ \eta_R(u_4) &\equiv& u_4 + s_2^4 + s_2 + v_3 s_1^8 + v_3^{2^{m+1}} s_1 \mod (2, v_1) \end{array}$$

for $r_1 = s_2^2 + v_1 u_2 s_1^2$.

This yields the relations in $\Sigma_m(2)$:

$$s_1^4 + s_1 \equiv v_1 r_1 \mod (2, v_1^{2^{m+2}}) \quad \text{and} \quad s_2^4 + s_2 + v_3 s_1^8 + v_3^{2^{m+1}} s_1 \equiv 0 \mod (2, v_1). \tag{3.4}$$

Lemma 3.5. *The diagonal Δ behaves on the generators s_i as follows:*

$$\begin{aligned}\Delta(s_1) &= s_1 \otimes 1 + 1 \otimes s_1,\\ \Delta(s_2) &= s_2 \otimes 1 + 1 \otimes s_2 + v_1 s_1 \otimes s_1,\\ \Delta(s_3) &\equiv s_3 \otimes 1 + 1 \otimes s_3 + v_2 s_1^2 \otimes s_1^2 \quad \bmod (2, v_1).\end{aligned}$$

Lemma 3.6. *Let z denote an element defined by $r_1^4 + r_1 + v_3^2 s_1^4 + v_3^{2^{m+2}} s_1^2 = v_1 z$. Then the cochains $r_1, z \in \widetilde{\Omega}^1_{\Sigma_m(2)} E_m(2)_*/(2)$ are cocycles. Besides, $z \equiv u_2 s_1^2$ modulo (v_1^2).*

Proof. Since $v_1 \in \widetilde{\Omega}^0_{\Sigma_m(2)} E_m(2)_*/(2)$ and $s_1 \in \widetilde{\Omega}^1_{\Sigma_m(2)} E_m(2)_*/(2)$ are both cocycles, so is r_1 by the relation $v_1 r_1 = s_1^4 + s_1 \in \Sigma_m(2)$ in (3.4). Furthermore, $v_3 \in \widetilde{\Omega}^0_{\Sigma_m(2)} E_m(2)_*/(2)$ is a cocycle. It follows similarly from its definition that z is a cocycle. By the definition of r_1, $r_1^4 + r_1 \equiv s_2^8 + s_2^2 + v_1 u_2 s_1^2 \equiv v_1 u_2 s_1^2 + v_3^2 s_1^{16} + v_3^{2^{m+2}} s_1^2$ modulo $(2, v_1^2)$ by (3.4). □

We now work as [6].

Lemma 3.7. *$u_2^t \in E_2^0(V_m(1))$ and $u_2^t h_{2,0} \in E_2^1(V_m(1))$ for each $t > 0$ are permanent cycles.*

Proof. For $t = 1$, the lemma is seen by Lemma 2.2. Consider the cofiber sequence $\Sigma^2 V_m(0) \xrightarrow{v_1} V_m(0) \xrightarrow{i_1} V_m(1) \xrightarrow{j_1} \Sigma^3 V_m(0)$. Put $d(u_2^t) = v_1 k_t' \in \widetilde{\Omega}^1_{\Sigma_m(2)} E_m(2)_*/(2)$ by virtue of Lemma 3.3, and let $k_t \in E_2^1(V_m(0))$ be the homology class of the cocycle k_t'. Then, $k_1 = h_{1,1}$, $v_1 k_t = 0$ and $k_{t+1} = \langle k_1, v_1, k_t \rangle$. Indeed, $\langle k_1, v_1, k_t \rangle$ is the class of $k_1' \eta_R(u_2^t) + u_2 k_t' = d(u_2^{t+1})/v_1 = k_{t+1}'$. Besides, $\delta(u_2^t) = k_t$ for the connecting homomorphism associated to the cofiber sequence. Let $\xi_1 \in \pi_{q_m - 1}(V_m(0))$ denote the homotopy element detected by k_1. Then, $v_1 \xi_1 = \xi_1 v_1 = 0$.

Suppose now that $u_2^t \in E_2^0(V_m(1))$ is a permanent cycle. Then, k_t is a permanent cycle that detects the element $\xi_t = j_1 u_2^t$ by the Geometric Boundary Theorem. Since $v_1 \xi_t = 0$, the Toda bracket $\{\xi_1, v_1, \xi_k\}$ is defined, which is detected by the Massey product $\langle k_1, v_1, k_t \rangle$. Note here that the Toda bracket is defined since $V_m(0)$ is a ring spectrum. It follows that k_{t+1} is a permanent cycle and detects a homotopy element, which we denote by ξ_{t+1}. Since the Massey product $\langle v_1, k_1, v_1 \rangle$ is zero in the E_2-term $E_2^{0, q_m + 4}(V_m(0))$, we see that $\{v_1, \xi_1, v_1\} = 0$ by Lemma 2.2. Now we compute $v_1 \{\xi_1, v_1, \xi_k\} = \{v_1, \xi_1, v_1\} \xi_k = 0$, and ξ_{t+1} is pulled back to u_2^{t+1} under the map j_1.

Turn to $u_2^t h_{2,0}$. In this case a similar argument works. For the connecting homomorphism δ, $\delta(u_2^t h_{2,0}) = \langle h_{1,0}^2, v_1, k_t \rangle$, which detects a homotopy element $\{\eta_0^2, v_1, \xi_t\}$, where η_0 denotes an element detected by $h_{1,0}$. Applying v_1 shows $\{v_1, \eta_0^2, v_1\} \xi_t = 0$. Indeed, $\{v_1, \eta_0^2, v_1\}$ is detected by $E_2^{s, 2q_m + 4 + s}(V_m(0))$ for $s > 2$. □

Lemma 3.8. *The elements $h_{1,0}, h_{1,1} \in E_2^1(V_m(0))$ and $h_{2,1} \in E_2^1(L_2 V_m(0))$ are permanent cycles.*

Proof. $h_{1,0}, h_{1,1}$ are seen immediately by Lemma 2.2.

The cobar module $\widetilde{\Omega}^{4,4q_m+6}_{\Gamma_m}BP_*/(2)$ is generated by $v_1^3 s_1^{\otimes 4}$ and $v_2 s_1^{\otimes 4}$ by degree reason. The first generator cobounds $v_1^2 s_2 \otimes s_1 \otimes s_1$, and we obtain $E_2^{4,4q_m+6}(V_m(0)) = \mathbb{Z}/2\{v_2 h_{1,0}^4\}$. Put $d_3(h_{2,1}) = av_2h_{1,0}^4 \in E_2^{4,4q_m+6}(V_m(0))$ for $a \in \mathbb{Z}/2$. Let w be an element fit in $d(s_3) = v_2 s_1^2 \otimes s_1^2 + v_1 w$ by virtue of Lemma 3.5. Then, $d(w) = 0$ in the cobar complex $\widetilde{\Omega}^3_{\Sigma_m(2)}E_m(2)_*/(2)$, and we see that $s_1^{\otimes 4}$ cobounds $s_3^2 \otimes s_1 \otimes s_1 + v_1 w^2 \otimes s_2 + (r_1 \otimes s_1 + s_1 \otimes r_1 + v_1 r_1 \otimes r_1) \otimes s_2$ (in which we set $v_2 = 1$). It follows that $d_3(h_{2,1}) = av_2h_{1,0}^4 = 0 \in E_2^4(L_2V_m(0))$ as desired. Indeed, $v_2h_{1,0}^4 = v_1 g h_{1,0}^2 = 0$, since $v_2 h_{1,0}^2 = v_1 g$ for an element g and $v_1 h_{1,0}^2 = 0$ by $d(s_2) = v_1 s_1 \otimes s_1$.

□

Proof of Theorem 1.4. Every element $x \in E_2^s(L_2V_m(1))$ is decomposed as $x = x'x''$ for $x' \in \mathbb{Z}/2[u_2]\otimes\wedge(h_{2,0})$ and $x'' \in K_{m-1}(2)_*\otimes\wedge(h_{1,0}, h_{1,1}, h_{2,1})$. Note that $K_{m-1}(2)_*\otimes\wedge(h_{1,0}, h_{1,1}, h_{2,1}) \subset E_2^*(L_2V_m(0))$. Since x' (resp. x'') is a permanent cycle of the Adams-Novikov spectral sequence for computing $\pi_*(L_2V_m(1))$ (resp. $\pi_*(L_2V_m(0))$) by Lemma 3.7 (resp. 3.8), we obtain that the element x is a permanent cycle from Corollary 2.9. We see that the extension problem is trivial by Lemma 2.8. Indeed, $\mathbb{Z}/2 = \pi_0(M_2)$ acts on $\pi_*(L_2V_m(1))$. □

4 The elements x_n

We introduce the integer b_n for $n \geq 0$ by

$$b_n = \begin{cases} a_n - 8 & n \equiv 1\ (3) \\ a_n - 3 & n \equiv 2\ (3) \\ 0 & n \equiv 0\ (3), \end{cases}$$

and the elements $x_n \in E_m(2)_*$ defined by

$$x_n = x_{n-1}^2 + v_1^{b_n} y_{n-1}, \quad \text{where} \quad y_n = \begin{cases} 0 & n \leq 0 \text{ or } n \equiv 2\ (3) \\ x_0 & n = 1 \\ x_2 + v_1^2 v_3^4 x_1^2 + v_1^4 v_3^{2^{m+3}} x_1 & n = 3 \\ x_{n-2} y_{n-3} & n \equiv 0, 1\ (3) \text{ and } n \geq 4. \end{cases} \tag{4.1}$$

We also consider cocycles $z_n \in \Sigma_m(2)$:

$$z_n = \begin{cases} s_1^{2^{n+1}} & n = 0, 1 \\ r_1^{2^{n-1}} & n = 2, 3 \\ x_{n-3} z_{n-3} & n > 3. \end{cases} \tag{4.2}$$

Proposition 4.3. *For the differential $d : \Omega^0_{\Sigma_m(2)}E_m(2)_*/(2) \to \Omega^1_{\Sigma_m(2)}E_m(2)_*/(2)$ of the cobar complex,*

$$d(x_n) = v_1^{a_n} z_n.$$

Proof. For $n = 0$ and 1, it is immediate from Lemma 3.3, and the cases for $n = 2$ and 3 follow from the computation $d(x_2) = d(u_2^4 + v_1^3 u_2) = v_1^4 s_1^8 + v_1^4 s_1^2 = v_1^6 r_1^2$ by (3.4). For $n = 4$,

$$\begin{aligned} d(x_4) &\equiv d(x_2^4 + v_1^{18}x_2 + v_1^{20}v_3^4x_1^2 + v_1^{22}v_3^{2^{m+3}}x_1) \\ &\equiv v_1^{24}r_1^8 + v_1^{24}r_1^2 + v_1^{24}v_3^4s_1^8 + v_1^{24}v_3^{2^{m+3}}s_1^4 \equiv v_1^{26}z^2 \equiv v_1^{26}x_1z_1 \mod (2, v_1^{28}) \end{aligned}$$

by the definition of z.

Suppose inductively that $d(x_{3k+1}) = v_1^{a_{3k+1}} x_{3k-2} z_{3k-2} \mod (2, v_1^{a_{3k+1}+2})$ for $k > 0$.

$$\begin{aligned} d(x_{3k+1}^2) &\equiv v_1^{2a_{3k+1}} x_{3k-2}^2 z_{3k-2}^2 \mod (2, v_1^{2a_{3k+1}+4}) \\ d(v_1^{a_{3k+2}-3} y_{3k+1}) &\equiv d(v_1^{a_{3k+2}-3} x_{3k-1} y_{3k-2}) \\ &\equiv v_1^{a_{3k+2}-3} x_{3k-1}(v_1 z_{3k-2}^2 + v_1^3 z_{3k-1}) \mod (2, v_1^{a_{3k+2}-3+a_{3k-1}}) \end{aligned}$$

and the sum shows $d(x_{3k+2}) \equiv v_1^{a_{3k+2}} x_{3k-1} z_{3k-1} \mod (2, v_1^{a_{3k+2}+2})$. Similarly,

$$\begin{aligned} d(x_{3k+2}^4) &\equiv v_1^{4a_{3k+2}} x_{3k-1}^4 z_{3k-1}^4 \mod (2, v_1^{4a_{3k+2}+8}) \\ d(v_1^{a_{3k+4}-8} y_{3k+3}) &\equiv d(v_1^{a_{3k+4}-8} x_{3k+1} y_{3k}) \\ &\equiv v_1^{a_{3k+4}-8} x_{3k+1}(v_1^6 z_{3k}^2 + v_1^8 z_{3k+1}) \mod (2, v_1^{a_{3k+4}-8+a_{3k+1}}) \end{aligned}$$

and we have $d(x_{3k+4}) = v_1^{a_{3k+4}} x_{3k+1} z_{3k+1} \mod (2, v_1^{a_{3k+4}+2})$, which completes the induction. □

Proof of Lemma 1.10. It suffices to show that $h_{1,0}/v_1^j \in E_2^1(L_2V_m(1)_\infty)$ equals ζ/v_1^{j-2}. The element $h_{1,0}/v_1^j$ is represented by s_1/v_1^j. We make a computation in the cobar complex

$$\begin{aligned} d(u_2^2/v_1^{j+2}) &= s_1^4/v_1^j = s_1/v_1^j + r_1/v_1^{j-1} \\ d(v_3^2u_2^2/v_1^{j+1}) &= v_3^2 s_1^4/v_1^{j-1} \\ d(v_3^{2^{m+2}} u_2/v_1^j) &= v_3^{2^{m+2}} s_1^2/v_1^{j-1} \\ d(x_2^2/v_1^{j+11}) &= r_1^4/v_1^{j-1} \end{aligned}$$

by Lemma 3.3 and Proposition 4.3. The sum yields the homologous relation $s_1/v_1^j \sim z/v_1^{j-2}$ by Lemma 3.6, and so $h_{1,0}/v_1^j = \zeta/v_1^{j-2}$ in $E_2^1(L_2V_m(1)_\infty)$. □

Proof of Proposition 1.6. We consider the v_1-Bockstein spectral sequence given by the short exact sequence $0 \to E_m(2)_*(V_m(1)) \xrightarrow{\varphi} E_m(2)_*(V_m(1)_\infty) \xrightarrow{v_1} E_m(2)_*(V_m(1)_\infty) \to 0$ for φ given by $\varphi(x) = x/v_1$. Let B^* denote the $\mathbb{Z}/2[v_1]$-module of the proposition. Then, it is easy to see that B^s contains the image of $\varphi_*\colon E_2^s(L_2V_m(1)) \to E_2^s(L_2V_m(1)_\infty)$ and that Proposition 4.3 defines a homomorphism $f\colon B^s \to E_2^s(L_2V_m(1)_\infty)$. We also consider the composite $\partial = \delta \circ f\colon B^s \to E_2^{s+1}(L_2V_m(1))$, where $\delta\colon E_2^s(L_2V_m(1)_\infty) \to E_2^{s+1}(L_2V_m(1))$ denotes the connecting homomorphism associated to the short exact sequence. By [7, Remark 3.11], it suffices to show the sequence

$$0 \longrightarrow \mathrm{Coker}\ \partial \xrightarrow{\varphi_*} B^* \xrightarrow{v_1} B^* \xrightarrow{\partial} \mathrm{Im}\ \partial \longrightarrow 0 \tag{4.4}$$

is exact.

We decompose $E_2^*(L_2V_m(1))$ into the direct sum of $M_C = K_{m-1}(2)_*[u_2^2]\{u_2\}\otimes\wedge(h_{10},h_{20},h_{21})$, $M_I = K_{m-1}(2)_*[u_2^2]\{h_{11}\}\otimes\wedge(h_{10},h_{20},h_{21})$ and $N\otimes\wedge(\zeta) = K_{m-1}(2)_*[u_2^2]\otimes\wedge(h_{10},h_{20},h_{21},\zeta)$. We notice that for non-negative integers n and r with $r<8$, there exist uniquely non-negative integers t and q such that $n = 8^qt + re_q$. By this fact, we decompose summands of N as follows:

$$\begin{aligned}
&K_{m-1}(2)_*[u_2^2]\\
&\quad = K_{m-1}(2)_* \oplus \textstyle\bigoplus_{k\ge 1} \underline{x_kK_{m-1}(2)_*[x_{k+1}]}_A,\\
&K_{m-1}(2)_*[u_2^2]h_{10}\\
&\quad = \textstyle\bigoplus_{q\ge 0}\left(\left(\underline{x_{3q+2}K_{m-1}(2)_*[x_{3q+3}]}_a \oplus \underline{x_{3q+3}K_{m-1}(2)_*[x_{3q+4}]}_b\right)\zeta_{3q+4} \oplus \underline{\underline{K_{m-1}(2)_*[x_{3q+2}]\zeta_{3q+1}}}_A\right),\\
&K_{m-1}(2)_*[u_2^2]h_{20}\\
&\quad = \textstyle\bigoplus_{q\ge 0}\left(\underline{x_{3q+3}K_{m-1}(2)_*[x_{3q+4}]\zeta_{3q+5}}_c \oplus \left(\underline{x_{3q+1}K_{m-1}(2)_*[x_{3q+2}]}_d \oplus \underline{\underline{K_{m-1}(2)_*[x_{3q+3}]}}_A\right)\zeta_{3q+2}\right),\\
&K_{m-1}(2)_*[u_2^2]h_{21}\\
&\quad = \textstyle\bigoplus_{q\ge 0}\left(\underline{x_{3q+1}K_{m-1}(2)_*[x_{3q+2}]}_e \oplus \underline{x_{3q+2}K_{m-1}(2)_*[x_{3q+3}]}_f \oplus \underline{\underline{K_{m-1}(2)_*[x_{3q+4}]}}_A\right)\zeta_{3q+3},
\end{aligned}$$

$$\begin{aligned}
&K_{m-1}(2)_*[u_2^2]h_{10}h_{20}\\
&\quad = \textstyle\bigoplus_{q\ge 0}\Big(\underline{\underline{K_{m-1}(2)_*[x_{3q+3}]\zeta_{3q+4}\zeta_{3q+2}}}_a \oplus \underline{x_{3q+3}K_{m-1}(2)_*[x_{3q+4}]\zeta_{3q+4}\zeta_{3q+5}}_B\\
&\qquad \oplus \underline{\underline{K_{m-1}(2)_*[x_{3q+2}]\zeta_{3q+1}\zeta_{3q+2}}}_d\Big),\\
&K_{m-1}(2)_*[u_2^2]h_{20}h_{21}\\
&\quad = \textstyle\bigoplus_{q\ge 0}\left(\underline{\underline{K_{m-1}(2)_*[x_{3q+4}]\zeta_{3q+3}\zeta_{3q+5}}}_c \oplus \left(\underline{x_{3q+1}K_{m-1}(2)_*[x_{3q+2}]}_B \oplus \underline{\underline{K_{m-1}(2)_*[x_{3q+3}]}}_f\right)\zeta_{3q+2}\zeta_{3q+3}\right),\\
&K_{m-1}(2)_*[u_2^2]h_{10}h_{21}\\
&\quad = \textstyle\bigoplus_{q\ge 0}\left(\left(\underline{K_{m-1}(2)_*[x_{3q+3}]x_{3q+2}}_B \oplus \underline{\underline{K_{m-1}(2)_*[x_{3q+4}]}}_b\right)\zeta_{3q+4}\zeta_{3q+3} \oplus \underline{\underline{K_{m-1}(2)_*[x_{3q+2}]\zeta_{3q+1}\zeta_{3q+3}}}_e\right),\\
&K_{m-1}(2)_*[u_2^2]h_{10}h_{20}h_{21}\\
&\quad = \textstyle\bigoplus_{k\ge 1} \underline{\underline{K_{m-1}(2)_*[x_{k+1}]\zeta_k\zeta_{k+1}\zeta_{k+2}}}_B.
\end{aligned}$$

Here, $\underline{M}_X$ and $\underline{\underline{M'}}_X$ for modules M and M' mean that M and M' are isomorphic under a Bockstein differential d_r for some r so that $d_r(M) = M'$, which is seen by Proposition 4.3. Let N_C (resp. N_I) be the direct sum of single (resp. double) underlined submodules of N, and put $\widetilde{M} = Q(0)\otimes\wedge(h_{1,0},h_{2,0},h_{2,1})$, $\widetilde{N} = \bigoplus_{k>0} Q(k)\otimes\wedge(\zeta_{k+1},\zeta_{k+2})$. Then we have the three exact sequences

$$\begin{gathered}
0\to M_C \xrightarrow{\varphi_*} \widetilde{M} \xrightarrow{v_1} \widetilde{M} \to M_I \to 0, \quad 0\to N_C \xrightarrow{\varphi_*} \widetilde{N} \xrightarrow{v_1} \widetilde{N} \to N_I \to 0 \quad \text{and}\\
0\to K_{m-1}(2)_* \to E_{m-1}(2)_*/(2,v_1^\infty) \to E_{m-1}(2)_*/(2,v_1^\infty) \to 0,
\end{gathered}$$

the direct sum of which yields the sequence (4.4). □

5 The Adams-Novikov E_∞-term for $\pi_*(L_2T(m)\wedge M_2)$

We first show that all elements of the Adams-Novikov E_2-term for $\pi_*(L_2V_m(1)_\infty)$ are permanent cycles. Take an element $x/v_1^t \in E_2^0(L_2V_m(1)_\infty)$. Then $x\in E_2^0(L_2V_m(1)_t)$. Thus, if $x = y^2/v_1^t$ for some $y \in E_2^0(L_2V_m(1)_{4t})$, then x is a permanent cycle. So it is sufficient to show that $d_3(x_n) = 0 \in E_2^3(L_2V_m(1)_{a_n})$ for each $n \geq 0$. We consider the integer

$$\varepsilon_n = \begin{cases} 2 & n \not\equiv 0\ (3) \\ 0 & n \equiv 0\ (3) \end{cases}$$

so that $V_m(1)_{a_n+\varepsilon_n}$ is a ring spectrum by Corollary 2.7.

Lemma 5.1. *$d_3(x_n) = 0 \in E_2^3(L_2V_m(1)_{a_n})$ for $n \geq 0$.*

Proof. For $n = 0$, it is shown in Lemma 3.7.

Suppose that $d_3(x_n) = \xi \in E_2^3(L_2V_m(1)_{a_n})$ for $n > 0$. Send this to $E_2^3(L_2V_m(1)_{a_{n-1}})$, and we see that $\xi = d_3(x_n) = d_3(x_{n-1}^2) \in E_2^3(L_2V_m(1)_{a_{n-1}})$. Then, the map $v_1^{\varepsilon_{n-1}} : E_2^3(L_2V_m(1)_{a_{n-1}}) \to E_2^3(L_2V_m(1)_{a_{n-1}+\varepsilon_{n-1}})$ assigns $v_1^{\varepsilon_{n-1}}\xi$ to $v_1^{2\varepsilon_{n-1}}\xi = d_3((v_1^{\varepsilon_{n-1}}x_{n-1})^2)$, which is zero, since $v_1^{\varepsilon_{n-1}}x_{n-1} \in E_2^0(L_2V_m(1)_{a_{n-1}+\varepsilon_{n-1}})$ and $V_m(1)_{a_{n-1}+\varepsilon_{n-1}}$ is a ring spectrum. It follows that $\xi = v_1^{a_{n-1}-\varepsilon_{n-1}}\xi'$ for some $\xi' \in E_2^3(L_2V_m(1)_{a_n-a_{n-1}+\varepsilon_{n-1}})$. Note that this works even if $n = 1$, though $V_m(1)$ is not a ring spectrum. Consider the commutative diagram

$$\begin{array}{ccccccc}
V_m(1) & = & V_m(1) & \longrightarrow & * & \longrightarrow & V_m(1) \\
\downarrow{\scriptstyle v_1^{a_n-a}} & & \downarrow{\scriptstyle v_1^{a_n}} & & \downarrow & & \downarrow{\scriptstyle v_1^{a_n-a}} \\
V_m(1)_{a_n-a+1} & \xrightarrow{v_1^a} & V_m(1)_{a_n+1} & \xrightarrow{i_v} & V_m(1)_a & \xrightarrow{j_v} & V_m(1)_{a_n-a+1} \\
\downarrow & & \downarrow & & \| & & \downarrow{\scriptstyle p} \\
V_m(1)_{a_n-a} & \xrightarrow{v_1^a} & V_m(1)_{a_n} & \xrightarrow{i'_v} & V_m(1)_a & \xrightarrow{j'_v} & V_m(1)_{a_n-a},
\end{array}$$

$(a = a_{n-1} - \varepsilon_{n-1})$ in which rows and columns are cofiber sequences. Let $\langle x\rangle \in \pi_*(X)$ denote a homotopy element detected by $x \in E_2^*(X)$. Noticing that $x_n \in E_2^0(L_2V_m(1)_{a_{n-1}-\varepsilon_{n-1}})$ is a permanent cycle, we see that $j_{v*}(\langle x_n\rangle) = \langle v_1^{a_n-a_{n-1}+\varepsilon_{n-1}}\zeta_n\rangle$ and $j'_{v*}(\langle x_n\rangle) = \langle \xi'\rangle$, and so $p_*(\langle v_1^{a_n-a_{n-1}+\varepsilon_{n-1}}\zeta_n\rangle) = \langle\xi'\rangle$. Since $\langle\zeta_n\rangle \in \pi_*(L_2V_m(1))$ by Theorem 1.4, we obtain $\langle\xi'\rangle = 0$, and $\langle x_n\rangle$ is in the image under the map i'_{v*}. It follows that there is a permanent cycle $x'_n \in E_2^0(L_2V_m(1)_{a_n})$, whose leading term is x_n, such that $i_{v*}(\langle x'_n\rangle) = \langle x_n\rangle \in \pi_*(L_2V_m(1)_{a_{n-1}-\varepsilon_{n-1}})$. The lemma now follows by replacing x_n by x'_n. □

Received: February 2008. Revised: April 2008.

References

[1] M. Hovey and H. Sadofsky, *Invertible spectra in the $E(n)$-local stable homotopy category*, J. London Math. Soc., **60** (1999), 284–302.

[2] I. Ichigi and K. Shimomura, *Subgroups of $\pi_*(L_2T(1))$ at the prime two*, Hiroshima Math. J., **33** (2003), 359–369.

[3] I. Ichigi, K. Shimomura and X. Wang, *On subgroups of $\pi_*(L_2T(1) \wedge M(2))$ at the prime two*, to appear in Bulletin of the Mexican Mathematical Society., **13**, (2007).

[4] Y. Kamiya and K. Shimomura, *The homotopy groups $\pi_*(L_2V(0) \wedge T(k))$*, Hiroshima Math. J., **31** (2001), 391–408.

[5] M. Mahowald, *The construction of small ring spectra*, Geom. Appl. Homotopy Theory II, Proc. Conf., Evanston 1977, Lect. Notes Math., **658** (1978), 234–239.

[6] M. Mahowald and K. Shimomura, *The Adams-Novikov spectral sequence for the L_2 localization of a v_2 spectrum*, the Proceedings of the International Congress in Algebraic Topology, Editated by M. Tangora, 1991, Contemporary Math., **146** (1993), 237–250.

[7] H.R. Miller, D.C. Ravenel, and W.S. Wilson, *Periodic phenomena in the Adams-Novikov spectral sequence*, Ann. of Math., **106** (1977), 469–516.

[8] D.C. Ravenel, *Localization with respect to certain periodic homology theories*, Amer. J. Math., **106** (1984), 351–414.

[9] D.C. Ravenel, *Nilpotence and Periodicity in Stable Homotopy Theory*, Annals of Mathematics Studies, Number 128, Princeton, 1992.

[10] D.C. Ravenel, *Complex cobordism and stable homotopy groups of spheres*, AMS Chelsea Publishing, Providence, 2004.

[11] K. Shimomura, *The homotopy groups of the L_2-localized Mahowald spectrum $X\langle 1\rangle$*, Forum Math., **7** (1995), 685–707.

[12] K. Shimomura and X. Wang, *The Adams-Novikov E_2-term for $\pi_*(L_2S^0)$ at the prime 2*, Math. Z., **241** (2002), 271–311.

CUBO A Mathematical Journal
Vol.10, Nº 03, (57–64). October 2008

The Extension of the Formula by Dupire

Shunsuke Kaji
Department of Mathematics, Graduate School of Science,
Osaka University, Machikaneyamachou 1-1,
Toyonaka, Osaka Japan 560-0043,
Osaka Japan 560-0043
email: kaji@math.sci.osaka-u.ac.jp

ABSTRACT

We provide the extension of Dupire's PDE, as the partial integro-differential equations of market prices of call options with many maturities and strike prices for jump diffusion model.

RESUMEN

Nosotros damos la extensión de Dupire PDE, como las ecuaciones parciales integro-diferenciales de precios de mercado de opciones de llamada con muchos vencimientos y golpe de precios para modelos de difusión con saltos.

Key words and phrases: *Dupire, PDE, Jump-diffusion model.*

Math. Subj. Class.: *35K15, 60H30.*

1 Introduction

Let $(\Omega, \mathcal{F}, P)$ be a probability space. On the space $(\Omega, \mathcal{F}, P)$ we set a standard Brownian motion $W = \{W_t\}_{t\in[0,T]}$ from $W_0 = 0$ and a Poission random measure $N(dtdx)$ on $(0, T]\times\mathbf{R}$ with intensity measure $dt\nu(dx)$, where $T \in (0, \infty)$ and the measure ν on $\mathbf{R}$ satisfies

$$\int_{\mathbf{R}} (1 + e^{2z}) \wedge z^2 \nu(dz) < \infty. \tag{1}$$

We consider a risk-neutral price process $\{S^x_t\}_{t\in[0,T]}$ of a risk asset satisfing

$$\begin{aligned} dS^x_t &= \sigma(t, S^x_t)S^x_t dW_t + (r - \delta)S^x_t dt; \\ S^x_0 &= x \in (0, \infty), \end{aligned}$$

where $r \geq 0$ denotes the interest rate and $\delta \geq 0$ the dividend rate. The function $\sigma : [0, T]\times(0, \infty) \to [0, \infty)$ has the Lipschiz condition and is often called the volatility of the asset's price. According to the well-known discussion of option pricing model, if for each $T, K \in (0, \infty)$ we have a unique solusion $u(t, x, T, K)$ to the parabolic equation and boundary condition

$$\begin{aligned} \frac{\partial u}{\partial t} + \frac{1}{2}\sigma(t, x)^2 x^2 \frac{\partial^2 u}{\partial x^2} + (r - \delta)x\frac{\partial u}{\partial x} - ru = 0, &\quad (t, x) \in [0, T) \times (0, \infty); \\ u(t, x, T, K)|_{t=T} = (x - K)^+, &\quad x \in (0, \infty), \end{aligned}$$

then a price of a call option with maturity T and strike price K is given by

$$u(t, x, T, K)|_{t=0} = e^{-rT}E[(S^x_T - K)^+].$$

Dupire[1] found that $u(t, x, T, K)$ as a function of (T, K) satisfies the following dual equation to the last parabolic equation:

$$\frac{\partial u}{\partial T} = \frac{1}{2}\sigma(T, K)^2 K^2 \frac{\partial^2 u}{\partial K^2} - (r - \delta)K\frac{\partial u}{\partial K} - \delta u, \quad (T, K) \in (t, \infty) \times (0, \infty).$$

But his approach is not enough mathematically. There are some works justifing rigorously his idea, for example, Klebaner[4] etc. Klebaner[4] gives the last equation by the Meyer-Tanaka formula. On the other hand, there are also works on option pricing model for jump-diffusion processes, for example, geometric Lévy processes by Fujiwara and Miyahara[2]. Recently, Jourdain[3] provides the extension of Dupire's work for jump-diffusion processes by stochastic flow approch.

Now, we consider the following risk-neutral evolusion $\{X^x_t\}_{t\in[0,T]}$ for the underlying risk asset's prices:

$$\begin{aligned} X^x_t &= x + \int_0^t a(u, X^x_u)X^x_u dW_u + (r - \delta)\int_0^t X^x_u du \\ &\quad + \int_{(0,t]\times\mathbf{R}} X^x_{u-}(e^z - 1)\{N(dudz) - du\nu(dz)\}, \quad t \in [0, T] \end{aligned}$$

where $a(t,x) : [0,T]\times(0,\infty) \to [0,\infty)$ satisfies the Lipschiz condition and has the second derivative with respect to x. Then $\{X_t^x\}_{t\in[0,T]}$ has the extended diffusion operator(see Yoshida[5] p.408)

$$\begin{aligned}(A_t f)(x) &= \frac{1}{2}a(t,x)^2 x^2 f''(x) + (r-\delta)xf'(x) \\ &\quad + \int_{\mathbf{R}} f(xe^z) - f(x) - (e^z-1)xf'(x)\nu(dz).\end{aligned}$$

For each maturity T and strike price K we denote

$$C(x,T,K) = e^{-rT}E[(X_T^x - K)^+] \tag{2}$$

by a call option price with an asset price x. In particular, in the case $a(\cdot,\cdot)\equiv a$ the last definition (2) is justified by Fujiwara and Miyahara[2]. If we moreover assume that $a(\cdot,\cdot)$ belongs to the class

$$\mathcal{V} = \left\{ f : [0,T]\times(0,\infty) \to \mathbf{R} | \sup_{(t,x)\in[0,T]\times(0,\infty)} \sum_{k=0}^{3} |x^k \frac{\partial^k f}{\partial x^k}(t,x)| < \infty \right\},$$

then Jourdain[3] provides the following equation of (T,K):

$$-\frac{\partial C}{\partial T} + \mathcal{A}_T C = 0, \quad (T,K)\in(0,\infty)\times(0,\infty),$$

where

$$\begin{aligned}(\mathcal{A}_T f)(K) &= \frac{1}{2}a(T,K)^2 K^2 f''(K) - (r-\delta)Kf'(K) - \delta f(K) \\ &\quad + \int_{\mathbf{R}} \{f(Ke^{-z}) - f(K) - (e^{-z}-1)Kf'(K)\}e^z\nu(dz).\end{aligned}$$

Here notice that the assumption $a(\cdot,\cdot)\in\mathcal{V}$ satisfies the Lipschiz condition. In this note we provide the same result of the above without $a(\cdot,\cdot)\in\mathcal{V}$ by using not only stochastic flow approch but also another one.

2 Main result

We fix $x\in(0,\infty)$ as follows. We have the following main theorem.

Theorem 2.1. *$C(x,T,K)$ as a function of (T,K) satisfies*

$$-\frac{\partial C}{\partial T} + \mathcal{A}_T C = 0, \quad (T,K)\in(0,\infty)\times(0,\infty)$$

in weak sense; that is,

$$\int_0^\infty\int_0^\infty C(x,T,K)\left\{\frac{\partial\varphi}{\partial T}(T,K) + \mathcal{A}_T^*\varphi(T,K)\right\}dTdK = 0,\ \forall\varphi\in C_0^\infty((0,\infty)^2),$$

where

$$\int_0^\infty\int_0^\infty \psi(T,K)\mathcal{A}_T^*\varphi(T,K)dTdK = \int_0^\infty\int_0^\infty \mathcal{A}_T\psi(T,K)\varphi(T,K)dTdK, \quad \forall\varphi,\forall\psi\in C_0^\infty((0,\infty)^2).$$

2.1 Lemmas

Lemma 2.1. *It follows that*

$$0 \le C(x,T,K) \le e^{-\delta T}x, \quad (T,K) \in (0,\infty)\times(0,\infty). \tag{3}$$

For every $\varphi(\cdot) \in C_0^2((0,\infty))$

$$e^{-rT}E[\varphi(X_T^x)] = \int_0^\infty C(x,T,K)\varphi''(K)dK, \quad T \in (0,\infty) \tag{4}$$

holds.

Remark 2.1. *It follows from* (3) *that* $C(x,T,K)$ *as a function of* (T,K) *is locally integrable on* $(0,\infty)\times(0,\infty)$. *Thus the right-hand side of* (4) *is well-defined.*

proof: By (2) we have

$$0 \le e^{rT}C(x,T,K) \le E[X_T^x].$$

Moreover, since $\{e^{-(r-\delta)t}X_t^x\}_{t\in[0,T]}$ is a nonnegative local martingale with initial value x, the right-hand side of the last inequality is

$$\le e^{(r-\delta)T}x.$$

Hence we get (3). Finally, we compute from (2) that the right-hand side of (4) is

$$\begin{aligned} &= \int_0^\infty e^{-rT}E[(X_T^x-K)^+]\varphi''(K)dK \\ &= e^{-rT}E\left[\int_0^\infty (X_T^x-K)^+\varphi''(K)dK\right] \\ &= e^{-rT}E[\varphi(X_T^x)]. \end{aligned}$$

Hence we get (4).

Before we moreover introduce lemmas, for every $\varphi(\cdot,\cdot) \in C_0^\infty((0,\infty)^2)$ we set a family $\{\Phi_h\}_{h>0}$ of all functions

$$\Phi_h(T,x) = \frac{1}{h}\{E[\varphi(T,X_{T+h}^x)] - E[\varphi(T,X_T^x)]\}, \quad (T,K)\in(0,\infty)\times(0,\infty).$$

Lemma 2.2.

$$\lim_{h\downarrow 0}\int_0^\infty \Phi_h(T,x)dT = -\int_0^\infty\int_0^\infty e^{rT}C(x,T,K)\frac{\partial^3\varphi}{\partial T\partial K^2}(T,K)dTdK.$$

proof: First, we set $\tilde{C}(x,T,K) = e^{rT}C(x,T,K)$. By using (4), we have

$$\int_0^\infty \Phi_h(T,x)dT = \int_0^\infty\left\{\int_0^\infty \frac{\tilde{C}(x,T+h,K)-\tilde{C}(x,T,K)}{h}\frac{\partial^2\varphi}{\partial K^2}(T,K)dK\right\}dT,$$

where $h > 0$. Moreover we compute that the right-hand side of the last equality is

$$\begin{aligned} &= \int_0^\infty \left\{ \int_0^\infty \frac{\tilde{C}(x,T+h,K)-\tilde{C}(x,T,K)}{h} \frac{\partial^2 \varphi}{\partial K^2}(T,K)dT \right\} dK \\ &= \int_0^\infty \left\{ \int_0^\infty \tilde{C}(x,T,K)\frac{1}{h}\left(\frac{\partial^2 \varphi}{\partial K^2}(T-h,K) - \frac{\partial^2 \varphi}{\partial K^2}(T,K)\right) dT \right\} dK, \end{aligned}$$

and so we have

$$\int_0^\infty \Phi_h(T,x)dT = \int_0^\infty \int_0^\infty \tilde{C}(x,T,K)\frac{1}{h}\left(\frac{\partial^2 \varphi}{\partial K^2}(T-h,K) - \frac{\partial^2 \varphi}{\partial K^2}(T,K)\right) dTdK.$$

Then, by using the dominated convergence theorem, $\varphi(\cdot,\cdot) \in C_0^\infty((0,\infty)^2)$ and (3) imply that the right-hand side of the last equality converges to

$$-\int_0^\infty \int_0^\infty \tilde{C}(x,T,K)\frac{\partial^3 \varphi}{\partial T \partial K^2}(T,K)dTdK$$

as $h \downarrow 0$. Hence we get the desired result.

We denote by the following operator depended on time $t \in [0,\infty)$:

$$\begin{aligned} (\tilde{\mathcal{A}}_t f)(x) &= \frac{1}{2}a(t,x)^2x^2f''(x) + \{\frac{\partial}{\partial x}(a(t,x)^2x^2) + (r-\delta)x\}f'(x) \\ &\quad + \left\{ \frac{1}{2}\frac{\partial^2}{\partial x^2}(a(t,x)^2x^2) + (r-2\delta) \right\} f(x) \\ &\quad + \int_{\mathbf{R}} e^{2z}f(xe^z) - (2e^z-1)f(x) - (e^z-1)xf'(x)\nu(dz). \end{aligned}$$

Lemma 2.3.

$$\lim_{h\downarrow 0} \int_0^\infty \Phi_h(T,x)dT = \int_0^\infty \int_0^\infty e^{rT}C(x,T,K)\left(\tilde{\mathcal{A}}_T \frac{\partial^2 \varphi}{\partial K^2}(T,K)\right) dTdK.$$

proof: First, we divide $A_\cdot$ into two parts as follows:

$$\begin{aligned} (A_\cdot f)(x) &= \left\{ \frac{1}{2}a(\cdot,x)^2x^2f''(x) + (r-\delta)xf'(x) \right. \\ &\quad + \int_{|z|<1} f(xe^z) - f(x) - zxf'(x)\nu(dz) \\ &\quad \left. - \int_{|z|<1} (e^z-1-z)xf'(x)\nu(dz) - \int_{|z|\geq 1} f(x) + (e^z-1)xf'(x)\nu(dz) \right\} \\ &\quad + \int_{|z|\geq 1} f(xe^z)\nu(dz) \\ &= (A_\cdot^0 f)(x) + \int_{|z|\geq 1} f(xe^z)\nu(dz). \end{aligned}$$

Since $\varphi(\cdot,\cdot)\in C_0^\infty((0,\infty)^2)$ we can choose subintervals $I_1=[\alpha_1,\beta_1]$ and $I_2=[\alpha_2,\beta_2]$ of $(0,\infty)$ such that *supp* $\varphi\subset I_1\times I_2$. We pick $\delta>0$ and set $\tilde{I_1}=\{x|\alpha_1\le x\le\beta_1+\delta\}$ and $\tilde{I_2}=\{x|\alpha_2e^{-1}\le x\le\beta_2e\}$. We denote by $\parallel f\parallel_{C(\Gamma)}=\sup_{x\in\Gamma}|f(x)|$, where Γ is a compact subset of $(0,\infty)^2$ and $f\in C(\Gamma)=\{f$ is a real-valued continuous function on $\Gamma\}$. Then observe that $A_T^0\varphi(T,\cdot)$ belongs to $C^2((0,\infty))$, since $\varphi(\cdot,\cdot)\in C_0^\infty((0,\infty)^2)$ and $a(T,\cdot)$ has the second derivative, and

$$\begin{aligned}
|A_u^0\varphi(T,K)| \le\ & \frac{1}{2}\parallel a^2\parallel_{C(\tilde{I_1}\times I_2)}\parallel K^2\frac{\partial^2\varphi}{\partial K^2}\parallel_{C(I_1\times I_2)}1_{I_1\times I_2}(T,K)\\
&+|r-\delta|\parallel K\frac{\partial\varphi}{\partial K}\parallel_{C(I_1\times I_2)}1_{I_1\times I_2}(T,K)\\
&+\int_{|z|<1}z^2\,\nu(dz)\parallel K^2\frac{\partial^2\varphi}{\partial K^2}+K\frac{\partial\varphi}{\partial K}\parallel_{C(I_1\times\tilde{I_2})}1_{I_1\times\tilde{I_2}}(T,K)\\
&+\int_{|z|<1}e^z-1-z\,\nu(dz)\parallel K\frac{\partial\varphi}{\partial K}\parallel_{C(I_1\times I_2)}1_{I_1\times I_2}(T,K)\\
&+\nu(|z|\ge1)\parallel\varphi\parallel_{C(I_1\times I_2)}1_{I_1\times I_2}(T,K)\\
&+\int_{|z|\ge1}|e^z-1|\,\nu(dz)\parallel K\frac{\partial\varphi}{\partial K}\parallel_{C(I_1\times I_2)}1_{I_1\times I_2}(T,K)\\
\le\ & \Bigg\{\frac{1}{2}\parallel a^2\parallel_{C(\tilde{I_1}\times I_2)}\parallel K^2\frac{\partial^2\varphi}{\partial K^2}\parallel_{C(I_1\times I_2)}\\
&+\Big(|r-\delta|+\int_{|z|<1}e^z-1-z\,\nu(dz)+\int_{|z|\ge1}|e^z-1|\,\nu(dz)\Big)\parallel K\frac{\partial\varphi}{\partial K}\parallel_{C(I_1\times I_2)}\\
&+\int_{|z|<1}z^2\,\nu(dz)\parallel K^2\frac{\partial^2\varphi}{\partial K^2}+K\frac{\partial\varphi}{\partial K}\parallel_{C(I_1\times\tilde{I_2})}+\nu(|z|\ge1)\parallel\varphi\parallel_{C(I_1\times I_2)}\Bigg\}\\
&\times 1_{I_1\times\tilde{I_2}}(T,K)\\
=\ & C_1\times 1_{I_1\times\tilde{I_2}}(T,K),\quad \forall u\in\tilde{I_1},
\end{aligned}$$

where $C_1<\infty$ holds since $\varphi(\cdot,\cdot)\in C_0^\infty((0,\infty)^2)$, (1), and $a(\cdot,\cdot)$ is continuous. Moreover, it is easy that we have

$$\left|\int_{|z|\ge1}\varphi(T,Ke^z)\nu(dz)\right|\le C_2 1_{I_1}(T),$$

where C_2 is a positive constant not depending on T and K. Therefore the inequality of the observation and the last inequality imply

$$|A_u\varphi(T,K)| \le (C_1+C_2)1_{I_1}(T),\quad \forall u\in\tilde{I_1}.$$

Here, fix T and by using Appendix 3.2 it follows from $\varphi(\cdot,\cdot)\in C_0^\infty((0,\infty)^2)$ that

$$\begin{aligned}
\Phi_h(T,x) &= \frac{1}{h}\int_T^{T+h}E[(A_u\varphi(T,\cdot))(X_u^x)]du\\
&= \frac{1}{h}\int_T^{T+h}E[A_u\varphi(T,X_u^x)]du,
\end{aligned}$$

for all $0 < h < \delta$. Then the last two inequality and equality imply

$$\begin{aligned} \lim_{h\downarrow 0}\Phi_h(T,x) &= E[A_T\varphi(T,X_T^x)]; \\ |\Phi_h(T,x)| &\leq (C_1+C_2)1_{I_1}(T), \quad 0<\forall h<\delta. \end{aligned}$$

According to the dominated convergence theorem, the last two results imply

$$\lim_{h\downarrow 0}\int_0^\infty \Phi_h(T,x)dT = \int_0^\infty E[A_T\varphi(T,X_T^x)]dT. \tag{5}$$

On the other hand, by using (4) we have from the above observation

$$e^{-rT}E[A_T^0\varphi(T,X_T^x)] = \int_0^\infty C(x,T,K)\frac{\partial^2}{\partial K^2}(A_T^0\varphi(T,\cdot))(K)dK$$

Moreover we have

$$\begin{aligned} e^{-rT}E[\int_{|z|\geq 1}\varphi(T,X_T^x e^z)\nu(dz)] &= \int_{|z|\geq 1} e^{-rT}E[\varphi(T,X_T^x e^z)]\nu(dz) \\ &= \int_{|z|\geq 1}\int_0^\infty C(x,T,K)\frac{\partial^2}{\partial K^2}(\varphi(T,Ke^z))dK\nu(dz) \\ &= \int_0^\infty C(x,T,K)\int_{|z|\geq 1} e^{2z}\frac{\partial^2\varphi}{\partial K^2}(T,Ke^z)\nu(dz)dK, \end{aligned}$$

where the second line of the last equality holds by (4). Therefore the last two equalities imply

$$\begin{aligned} e^{-rT}E[A_T\varphi(T,X_T^x)] &= \int_0^\infty C(x,T,K)\{\frac{\partial^2}{\partial K^2}(A_T^0\varphi(T,K)) \\ &\quad + \int_{|z|\geq 1} e^{2z}\frac{\partial^2\varphi}{\partial K^2}(T,Ke^z)\nu(dz)\}dK, \end{aligned}$$

and so by computing the right-hand side of the last equality we have

$$e^{-rT}E[A_T\varphi(T,X_T^x)] = \int_0^\infty C(x,T,K)(\tilde{A}_T\frac{\partial^2\varphi}{\partial K^2}(T,K))dK.$$

Hence (5) and the last equality imply the desired result.

2.2 Proof of Theorem 2.1

First, pick any $\psi(T,K)\in C_0^\infty((0,\infty)^2)$. According to Lemma 2.2 and 2.3, for all $\varphi(T,K)\in C_0^\infty((0,\infty)^2)$ such that $e^{rT}\frac{\partial^2\varphi}{\partial K^2}=\psi$, we have

$$\int_0^\infty\int_0^\infty e^{rT}C(x,T,K)\left\{\frac{\partial^3\varphi}{\partial T\partial K^2}(T,K)+\tilde{A}_T\frac{\partial^2\varphi}{\partial K^2}(T,K)\right\}dTdK=0,$$

and so

$$\int_0^\infty\int_0^\infty C(x,T,K)\left\{\frac{\partial\psi}{\partial T}(T,K)+\tilde{A}_T\psi(T,K)\right\}dTdK=0$$

holds. On the other hand, we can compute the integral by parts

$$\int_0^\infty \int_0^\infty \psi(T,K)\tilde{\mathcal{A}}_T\varphi(T,K)dTdK = \int_0^\infty \int_0^\infty \mathcal{A}_T\psi(T,K)\varphi(T,K)dTdK, \quad \forall\varphi, \forall\psi \in C_0^\infty((0,\infty)^2).$$

Hence the last two equalities imply the desired conclusion.

3 Appendix

Appendix 3.1. *Let* $\mathbf{X} \subset \mathbf{R}^d$*, where* d *is a positive integer, be a domain and* $C^k(\mathbf{X})$*, where* $k = 0, 1, 2, \cdots, \infty$*, be a class of all real-valued functions on* $\mathbf{X}$ *which have continuous partial derivatives of order* $\leq k$ *if* $k < \infty$*; of order* $< \infty$ *if* $k = \infty$*. Let* $C_0^k(\mathbf{X})$ *be a class of all functions which belong to* $C^k(\mathbf{X})$ *and compact supports.*

Appendix 3.2. *(Dynkin's formula)*
For every $f \in C_0^2((0,\infty))$*,*

$$E[f(X_t^x)] = f(x) + E\left[\int_0^t (A_u f)(X_u^x)du\right], \; (t,x) \in [0,\infty) \times (0,\infty)$$

holds.

Acknowledgment: The author is grateful to J.Sekine, Assistant Professor for useful advices.

Received: February 2008. Revised: May 2008.

References

[1] Dupire, B., *Pricing with a smile*, Risk, **7** (1994), pp.18–20.

[2] Fujiwara, T. and Miyahara, Y., *The minimal entropy martingale measures for geometric Lévy processes*, Finance Stochast., **7** (2003), pp.509–531.

[3] Jourdain, B., *Stochastic flow approach to Dupire's formula*, Finance Stochast., (2007).

[4] Klebaner, F., *Option price when the stock is a semimartingale*, Elect. Comm. in Probab., **7** (2002), pp.79–83.

[5] Yoshida, K., *Functional analysis*, 4th ed.

CUBO A Mathematical Journal
Vol.10, Nº 03, (65–82). October 2008

Equilibrium Cycles in a Two-Sector Economy with Sector Specific Externality

Miki Matsuo, Kazuo Nishimura
Institute of Economic Research, Kyoto University
email: `nishimura@kier.kyoto-u.ac.jp`

Tomoya Sakagami
Department of Economics,
Kumamoto Gakuen University

and

Alain Venditti†
CNRS-GREQAM

ABSTRACT

In this paper, we study the two-sector CES economy with sector-specific externality (feedback effects). We characterize the equilibrium paths in the case that allows negative externality, and show how the degree of externality may generate equilibrium cycles around the steady state.

RESUMEN

En este artculo estudiamos economia de dos-sector CES con externalidad de sector-especifico (efecto de retroalimentacin). Nosotros caracterizamos la trajectoria de equi-

†This paper has been written while Alain Venditti was visiting the Institute of Economic Research of Kyoto University. He thanks Professor Kazuo Nishimura and all the staff of the Institute for their kind invitation.

librio en el caso que permite externalidad negativa, e demonstramos como el grado de externalidad puede generar ciclos de equilibrio alrededor del estado regular.

Key words and phrases: *Difference equations, nonlinear dynamics, bifurcation, two-periodic cycle, multiple equilibria.*

Math. Subj. Class.: *37G10, 39A11, 91B50, 91B62, 91B64, 91B66.*

1 Introduction

The aim of this paper is to show the existence of equilibrium cycles around the steady state in the two-sector model with CES production function and sector specific externality.[1] A representative agent has concrete expectations on the level of externality and make a decision assuming that the externality is not affected by his own choice of decision variables. However, externalities come from the average values of capital and labor on the market. Therefore, if a representative agent chooses values of decision variables, externalities also vary as everybody also takes the same decision.

Over the last decade, an important literature has focused on the existence of locally indeterminate equilibria in dynamic general equilibrium economies with technological external effects. Local indeterminacy means that there exists a continuum of equilibria starting from the same initial condition, all of which converging to the same steady state. It is now well-known that local indeterminacy is a sufficient condition for the existence of endogenous fluctuations generated by purely extrinsic belief shocks which do not affect the fundamentals, i.e. the preferences and technologies.[2] Indeed, in presence of local indeterminacy, by randomizing beliefs over the continuum of equilibrium paths, one may construct equilibria defined with respect to shocks on expectations, and thus provide an alternative to technology or taste shocks to get propagation mechanisms and to explain macroeconomic volatility.

Benhabib and Nishimura [3, 4] proved that indeterminacy may arise in a continuous time economy in which the production functions from the social perspective have constant return to scale. Benhabib, Nishimura and Venditti [5] studied the two-sector model with sector specific external effects in discrete time framework. They provided conditions in which indeterminacy may occur even if the production function is decreasing return to scale from the social perspective. Nishimura and Venditti [7] study the interplay between the elasticity of capital-labor substitution and the rate of depreciation of capital, and its influence on the local behavior of equilibrium paths in a neighborhood of the steady state. However, in all these contributions, the existence of local bifurcations as the degree of externalities is modified is not discussed.

In this paper, we study the model in Nishimura and Venditti [7], focusing on the external effect of capital-labor ratio in the pure capital good sector and characterize the equilibrium paths

[1] External effects are feedbacks from the other agents in the economy who also face identical maximizing problems. See Benhabib and Farmer [2] for a survey.

[2] See Cass and Shell [6].

in the case that allows negative externality, which was not discussed in their paper. We will focus on the existence of flip bifurcation, i.e. of period-two equilibrium cycles, through the existence of local indeterminacy.

In Section 2 we describe the model. We discuss the existence of a steady state and give the local characterization of the equilibrium paths around the steady state in Section 3. Section 4 contains some concluding comments.

2 The model

We consider a two-sector model with an infinitely-lived representative agent. We assume that its single period linear utility function is given by

$$u(c_t) = c_t.$$

We assume that the consumption good, c, and capital good are produced with a Constant Elasticity of Substitution (CES) production functions.

$$c_t = \left[\alpha_1 K_{ct}^{-\rho_c} + \alpha_2 L_{ct}^{-\rho_c}\right]^{-\frac{1}{\rho_c}} \tag{1}$$

$$y_t = \left[\beta_1 K_{yt}^{-\rho_2} + \beta_2 L_{yt}^{-\rho_2} + e_t\right]^{-\frac{1}{\rho_y}} \tag{2}$$

where $\rho_c,\ \rho_y > -1$ and e_t represents the time-dependent externality (feedback effects) in the capital good sector. Let the elasticity of capital-labor substitution in each sector be $\sigma_c = \frac{1}{1+\rho_c} \geq 0$ and $\sigma_y = \frac{1}{1+\rho_y} \geq 0$. We assume that the externalities are as follows:

$$e = b\bar{K}_{yt}^{-\rho_y} - b\bar{L}_{yt}^{-\rho_y}, \tag{3}$$

where $b > 0$, and $\bar{K}_y$ and $\bar{L}_y$ represents the economy-wide average values. The representative agent takes these economy-wide average values as given.

Definition 1 *We call* $y = \left[\beta_1 K_y^{-\rho_y} + \beta_2 L_y^{-\rho_y} + e\right]^{-\frac{1}{\rho_y}}$ *the production function from the private perspective, and* $y = \left[(\beta_1 + b) K_y^{-\rho_y} + (\beta_2 - b) L_y^{-\rho_y}\right]^{-\frac{1}{\rho_y}}$ *the production function from the social perspective.*

In the rest of the paper we will assume that $\alpha_1 + \alpha_2 = \beta_1 + \beta_2 = 1$ so that the consumption good sector does not have externalities. Notice then that denoting $\hat{\beta}_1 = \beta_1 + b$ and $\hat{\beta}_2 = \beta_2 - b$, we get also $\hat{\beta}_1 + \hat{\beta}_2 = 1$. The investment good sector has externalities but the technology is linearly homogeneous, i.e. has constant returns, from the social perspective.

Remark 1 Notice that the externality (3) may be expressed as follows

$$e = b\bar{L}_2^{-\rho_y}\left[\left(\frac{\bar{K}_y}{\bar{L}_y}\right)^{-\rho_y} - 1\right]. \tag{4}$$

Now consider the production function from the social perspective as given in Definition 1. Dividing both sides by L_y, we get denoting $k_y = K_y/L_y$ and $\tilde{y} = y/L_y$

$$\tilde{y} = \left[(\beta_1 + b)\, k_y^{-\rho_y} + (\beta_2 - b)\right]^{-\frac{1}{\rho_y}}. \tag{5}$$

From equations (4) and (5) we derive that the externality is given in terms of the capital-labor ratio in the investment good sector.

The aggregate capital is divided between sectors,

$$k_t = K_{ct} + K_{yt},$$

and the labor endowment is normalized to one and divided between sectors,

$$L_{ct} + L_{yt} = 1.$$

The capital accumulation equation is

$$k_{t+1} = y_t,$$

as the capital depreciates completely in one period. To simplify we assume that both technologies are characterized by the same properties of substitution, i.e. $\rho_c = \rho_y = \rho$.

The consumer optimization problem will be given by

$$\begin{aligned} \max \quad & \sum_{t=0}^{\infty} \delta^t \left[\alpha_1 K_{ct}^{-\rho} + \alpha_2 L_{ct}^{-\rho}\right]^{-\frac{1}{\rho}} \\ s.t. \quad & y_t = \left[\beta_1 K_{yt}^{-\rho} + \beta_2 L_{yt}^{-\rho} + e_t\right]^{-\frac{1}{\rho}} \\ & 1 = L_{ct} + L_{yt} \\ & k_t = K_{ct} + K_{yt} \\ & y_t = k_{t+1} \\ & k_0,\ \{e_t\}_{t=0}^{\infty} \quad given \end{aligned} \tag{6}$$

where $\delta \in (0,1)$ is the discount factor. p_t, r_t, and w_t respectively denote the price of capital goods, the rental rate of the capital goods and the wage rate of labor at time $t \geq 0$[3]. For any sequences $\{e_t\}_{t=0}^{\infty}$ of external effects that the representative agent considers given, the Lagrangian at time

[3]We normalize the price of consumption goods to one.

$t \geq 0$ is defined as follows:

$$\begin{aligned} \mathcal{L}_t &= \left[\alpha_1 K_{ct}^{-\rho} + \alpha_2 L_{ct}^{-\rho}\right]^{-\frac{1}{\rho}} + p_t \left[\left[\beta_1 K_{yt}^{-\rho} + \beta_2 L_{yt}^{-\rho} + e_t\right]^{-\frac{1}{\rho}} - k_{t+1}\right] \\ &+ r_t \left(k_t - K_{ct} - K_{yt}\right) + w_t \left(1 - L_{ct} - L_{yt}\right). \end{aligned} \tag{7}$$

Then the first order conditions derived from the Lagrangian are as follows:

$$\frac{\partial \mathcal{L}_t}{\partial K_{ct}} = \alpha_1 \left(\frac{c_t}{K_{ct}}\right) - r_t = 0, \tag{8}$$

$$\frac{\partial \mathcal{L}_t}{\partial L_{ct}} = \alpha_2 \left(\frac{c_t}{L_{ct}}\right) - w_t = 0, \tag{9}$$

$$\frac{\partial \mathcal{L}_t}{\partial K_{yt}} = p_t \beta_1 \left(\frac{y_t}{K_{yt}}\right) - r_t = 0, \tag{10}$$

$$\frac{\partial \mathcal{L}_t}{\partial L_{yt}} = p_t \beta_2 \left(\frac{y_t}{L_{yt}}\right) - w_t = 0. \tag{11}$$

From the above first order conditions, we derive the following equation,

$$\left(\frac{\alpha_1 / \alpha_2}{\beta_1 / \beta_2}\right) = \left(\frac{K_{ct} / L_{ct}}{K_{yt} / L_{yt}}\right)^{1+\rho}. \tag{12}$$

If $\alpha_1/\alpha_2 > (<) \, \beta_1/\beta_2$, the consumption (capital) good sector is capital intensive from the private perspective.

For any value of $(k_t,\ y_t)$, solving the first order conditions with respect to K_{ct}, K_{yt}, L_{ct}, L_{yt} gives these inputs as functions of capital stock at time t and $t+1$, and external effect, namely:

$$\begin{aligned} K_{ct} &= K_c\left(k_t,\ y_t, e_t\right), \quad L_{ct} = L_c\left(k_t,\ y_t, e_t\right), \\ K_{yt} &= K_y\left(k_t,\ y_t, e_t\right), \ L_{yt} = L_y\left(k_t,\ y_t, e_t\right). \end{aligned}$$

For any given sequence $\{e_t\}_{t=0}^{\infty}$, we define the efficient production frontier as follows:

$$T^*\left(k_t,\ k_{t+1}, e_t\right) = \left[\alpha_1 K_c\left(k_t,\ y_t, e_t\right)^{-\rho} + \alpha_2 L_c\left(k_t,\ y_t, e_t\right)^{-\rho}\right]^{-\frac{1}{\rho}}.$$

Using the envelope theorem we derive the equilibrium prices,[4]

$$T_2\left(k_t, k_{t+1}, e_t\right) = -p_t, \tag{13}$$

$$T_1\left(k_t, k_{t+1}, e_t\right) = r_t. \tag{14}$$

[4]See Takayama for the envelope theorem, pp160-165. Using the envelope theorem, we get $\frac{\partial \mathcal{L}_t}{\partial k_t} = \frac{\partial T}{\partial k_t}$ and $\frac{\partial \mathcal{L}_t}{\partial k_{t+1}} = \frac{\partial T}{\partial k_{t+1}}$. This is equivalent to (13) and (14).

Next we solve the intertemporal problem (6). In this model, lifetime utility function becomes

$$U = \sum_{t=0}^{\infty} \delta^t T^* \left(k_t, k_{t+1}, e_t\right).$$

From the first order conditions with respect to k_{t+1}, we obtain the Euler equation

$$T_2 \left(k_t, k_{t+1}, e_t\right) + \delta T_1 \left(k_{t+1}, k_{t+2}, e_{t+1}\right) = 0. \tag{15}$$

The solution of equation (15) also has to satisfy the following transversality condition

$$\lim_{t\to+\infty} \delta^t k_t T_1 \left(k_t, k_{t+1}, e_t\right) = 0. \tag{16}$$

We denote the solution of this problem $\{k_t\}_{t=0}^{\infty}$. This path depends on the choice of sequence $\{e_t\}_{t=0}^{\infty}$. If the sequence $\{e_t\}_{t=0}^{\infty}$ satisfies

$$e_t = bK_y \left(k_t,\ y_t, e_t\right)^{-\rho} - bL_y \left(k_t,\ y_t, e_t\right)^{-\rho}, \tag{17}$$

then $\{\hat{k}_t\}_{t=0}^{\infty}$ is called an equilibrium path. Along an equilibrium path, the expectations of the representative agent on the externalities $\{e_t\}_{t=0}^{\infty}$ are realized.

Definition 2 *$\{k_t\}_{t=0}^{\infty}$ is an equilibrium path if $\{k_t\}_{t=0}^{\infty}$ satisfies (15), (16) and (17).*

Solving the equation (17) for e_t, we derive e_t that is given as a function of $(k_t,\ k_{t+1})$, namely $e_t = \hat{e}\left(k_t, k_{t+1}\right)$. Let us substitute $\hat{e}\left(k_t, k_{t+1}\right)$ into equations (13) and (14) and define the equilibrium prices as

$$p_t = p_t \left(k_t, k_{t+1}\right),$$

$$r_t = r_t \left(k_t, k_{t+1}\right).$$

Then the Euler equation (15) evaluated at $\{k_t\}_{t=0}^{\infty}$ is

$$-p \left(k_t, k_{t+1}\right) + \delta r \left(k_{t+1}, k_{t+2}\right) = 0. \tag{18}$$

We have the following lemma.

Lemma 1 *The partial derivatives of $T\left(k_t, k_{t+1}, e_t\right)$ with respect to k_t and k_{t+1} are given by*

$$\begin{aligned} T_1 \left(k_t, k_{t+1}, \hat{e}\left(k_t, k_{t+1}\right)\right) &= \alpha_1 \left[\alpha_1 + \alpha_2 \left(\frac{\alpha_1\beta_2}{\alpha_2\beta_1}\right)^{\frac{1+\rho}{\rho}} \left(\frac{\left(\frac{g}{k_{t+1}}\right)^{\rho}}{\beta_2 - b} - \frac{(\beta_1 + b)}{\beta_2 - b}\right)^{\rho}\right]^{-\frac{1+\rho}{\rho}} \\ T_2 \left(k_t, k_{t+1}, \hat{e}\left(k_t, k_{t+1}\right)\right) &= \frac{T_1(k_t,\ k_{t+1}, \hat{e}(k_t, k_{t+1}))}{\beta_1} \left(\frac{g}{k_{t+1}}\right)^{1+\rho} \end{aligned}$$

where

$$g = g\left(k_t, k_{t+1}\right) = \left\{K_{yt} \in [0, k_t] \mid \frac{\alpha_1\beta_2}{\alpha_2\beta_1} = \left(\frac{k_t - K_{yt}}{1 - L_{yt}(K_{yt}, k_{t+1})}\right)^{1+\rho} \left(\frac{L_{yt}(K_{yt}, k_{t+1})}{K_{yt}}\right)^{1+\rho}\right\}$$

and

$$L_{yt}\left(K_{yt},k_{t+1}\right)=\left(\frac{k_{t+1}^{-\rho}-(\beta_1+b)K_{yt}^{-\rho}}{\beta_2-b}\right)^{-\frac{1}{\rho}}.$$

3 Steady state

Definition 3 *A steady state is defined by $k_t=k_{t+1}=y_t=k^*$ and is given by the solution of $T_2\left(k^*,k^*,e^*\right)+\delta T_1\left(k^*,k^*,e^*\right)=0$ with $e^*=\hat{e}\left(k^*,k^*\right)$.*

In the rest of the paper we assume the following restriction on parameters' values that guarantees all the steady state values are positive.

Assumption 1 *The parameters δ, β_1, b and ρ satisfy*

$$(\delta\beta_1)^{\frac{-\rho}{1+\rho}}<\beta_1+b.$$

We obtain the steady state value.

Proposition 1 *In this model, there exists a unique stationary capital stock k^* such that:*

$$k^*=\left\{1+\left(\frac{\alpha_1\beta_2}{\alpha_2\beta_1}\right)^{\frac{-1}{1+\rho}}(\delta\beta_1)^{\frac{-1}{1+\rho}}\left[1-(\delta\beta_1)^{\frac{1}{1+\rho}}\right]\right\}^{-1}\left[\frac{1-\hat{\beta}_1(\delta\beta_1)^{\frac{-\rho}{1+\rho}}}{\beta_2}\right]^{\frac{1}{\rho}}. \tag{19}$$

To study local behavior of the equilibrium path around the steady state k^*, we linearize the Euler equation (15) at the steady state k^* and obtain the following characteristic equation

$$\delta T_{12}\lambda^2+\left[\delta T_{11}+T_{22}\right]\lambda+T_{21}=0,$$

or

$$\delta\lambda^2+\left[\delta\frac{T_{11}}{T_{12}}+\frac{T_{22}}{T_{12}}\right]\lambda+\frac{T_{21}}{T_{12}}=0. \tag{20}$$

As shown in Nishimura and Venditti [7], the expressions of the characteristic roots are as follows:

Proposition 2 *The characteristic roots of Equation (20) are*

$$\lambda_1 = \frac{1}{(\delta\beta_2)^{\frac{1}{1+\rho}}\left[\left(\frac{\beta_1}{\beta_2}\right)^{\frac{1}{1+\rho}}-\left(\frac{\alpha_1}{\alpha_2}\right)^{\frac{1}{1+\rho}}\right]}, \tag{21}$$

$$\lambda_2(b) = \frac{(\delta\beta_2)^{\frac{1}{1+\rho}}\left[\frac{\beta_1+b}{\beta_1}\left(\frac{\beta_1}{\beta_2}\right)^{\frac{1}{1+\rho}}-\frac{\beta_2-b}{\beta_2}\left(\frac{\alpha_1}{\alpha_2}\right)^{\frac{1}{1+\rho}}\right]}{\delta}.$$

The roots of the characteristic equation determine the local behavior of the equilibrium paths. The sign of λ_1 is determined by factor intensity differences from the private perspective.[5]

We now characterize the equilibrium paths in this model. In particular we can show that the local behavior of equilibrium path around the steady state changes according to the degree of external effect in the capital good sector.

Definition 4 *A steady state k^* is called locally indeterminate if there exists ε such that for any $k_0 \in (k^* - \varepsilon, k^* + \varepsilon)$, there are infinitely many equilibrium paths converging to the steady state.*

As there is one pre-determined variable, the capital stock, local indeterminacy occurs if the stable manifold has two dimension, i.e. if the two characteristic roots are within the unit circle. We will also present conditions for local determinacy (for saddle-point stability) in which there exists a unique equilibrium path. Such a configuration occurs if the stable manifold has one dimension, i.e. if one root is outside the unit circle while the other is inside.

When the investment good is capital intensive, local indeterminacy and flip bifurcation cannot occur.

Proposition 3 *Suppose that the capital good sector is capital intensive from the private perspective, i.e. $\alpha_2\beta_1 > \alpha_1\beta_2$. Then the characteristic roots λ_1 and $\lambda_2(b)$ are positive with $\lambda_1 > 1$.*

Next we present our results assuming that the capital good is labor intensive from the private perspective, i.e. $\alpha_2\beta_1 - \alpha_1\beta_2 < 0$. Equilibrium period-two cycles may occur in this case through a flip bifurcation. We will also get local indeterminacy of equilibria. By rewriting equation (21), the characteristic roots are

$$\begin{aligned} \lambda_1 &= -\frac{1}{(\delta\beta_2)^{\frac{1}{1+\rho}}\left[\left(\frac{\alpha_1}{\alpha_2}\right)^{\frac{1}{1+\rho}} - \left(\frac{\beta_1}{\beta_2}\right)^{\frac{1}{1+\rho}}\right]}, \qquad (22) \\ \lambda_2(b) &= -\frac{(\delta\beta_2)^{\frac{1}{1+\rho}}\left[\frac{\beta_2-b}{\beta_2}\left(\frac{\alpha_1}{\alpha_2}\right)^{\frac{1}{1+\rho}} - \frac{\beta_1+b}{\beta_1}\left(\frac{\beta_1}{\beta_2}\right)^{\frac{1}{1+\rho}}\right]}{\delta}. \end{aligned}$$

To get $\lambda_1 \in (-1, 0)$, we need however to suppose a slightly stronger condition than simply ensuring the capital good sector to be labor intensive from the private perspective. The capital intensity difference $\alpha_1\beta_2 - \alpha_2\beta_1$ needs to be large enough and the discount factor has to be close enough to 1.

Proposition 4 *Assume that $(\alpha_1\beta_2)^{\frac{1}{1+\rho}} - (\alpha_2\beta_1)^{\frac{1}{1+\rho}} > \alpha_2^{\frac{1}{1+\rho}}$ and $\delta \in (\delta_3, 1)$ with*

$$\delta_3 = \beta_2^{-1}\left[(\beta_1/\beta_2)^{\frac{1}{1+\rho}} - (\alpha_1/\alpha_2)^{\frac{1}{1+\rho}}\right]^{-1-\rho} < 1.$$

[5] If $\alpha_2\beta_1 - \alpha_1\beta_2 > 0$, the capital good sector is capital intensive from the private perspective.

Then there exist $\underline{b}(\delta) > 0$ *and* $\overline{b}(\delta) > \underline{b}(\delta)$ *such that the steady state is saddle point for* $b \in (0, \underline{b}(\delta))$, *undergoes a flip bifurcation when* $b = \underline{b}(\delta)$, *becomes locally indeterminate for* $b \in \left(\underline{b}(\delta), \overline{b}(\delta)\right)$ *and is again saddle-point stable for* $(\overline{b}(\delta), +\infty)$. *Generically, there exist period-two cycles in a left (right) neighborhood of* $\underline{b}(\delta)$ *that are locally indeterminate (saddle-point stable).*

Next we still assume that the capital good is labor intensive from the private perspective with $\alpha_2\beta_1 - \alpha_1\beta_2 < 0$, but make λ_1 an unstable root, i.e. $\lambda_1 < -1$. As a result local indeterminacy cannot occur but period-two cycles may still exist through a flip bifurcation. Two cases need to be considered: $(\alpha_1\beta_2)^{\frac{1}{1+\rho}} - (\alpha_2\beta_1)^{\frac{1}{1+\rho}} > \alpha_2^{\frac{1}{1+\rho}}$ and $\delta \in (0, \delta_3)$, as well as $(\alpha_1\beta_2)^{\frac{1}{1+\rho}} - (\alpha_2\beta_1)^{\frac{1}{1+\rho}} < \alpha_2^{\frac{1}{1+\rho}}$. The following result is proved along the same lines as Proposition 4.

Proposition 5 *Suppose that the capital goods sector is labor intensive from the private perspective and let*

$$\delta_4 = \beta_2^{\frac{1}{\rho}} \left[(\beta_1/\beta_2)^{\frac{1}{1+\rho}} - (\alpha_1/\alpha_2)^{\frac{1}{1+\rho}}\right]^{\frac{1+\rho}{\rho}} .$$

Assume also that one of the following sets of conditions hold:

i) $(\alpha_1\beta_2)^{\frac{1}{1+\rho}} - (\alpha_2\beta_1)^{\frac{1}{1+\rho}} > \alpha_2^{\frac{1}{1+\rho}}$ *and* $\delta \in (0, \delta^*)$ *with* $\delta^* = \min\{\delta_3, \delta_4\}$,

ii) $(\alpha_1\beta_2)^{\frac{1}{1+\rho}} - (\alpha_2\beta_1)^{\frac{1}{1+\rho}} < \alpha_2^{\frac{1}{1+\rho}}$, $\rho > 0$ *and* $\delta \in (0, \delta_4)$,

Then there exist $\underline{b}(\delta) > 0$ *and* $\overline{b}(\delta) > \underline{b}(\delta)$ *such that the steady state is totally unstable for* $b \in (0, \underline{b}(\delta))$, *undergoes a flip bifurcation when* $b = \underline{b}(\delta)$, *becomes saddle-point stable for* $b \in \left(\underline{b}(\delta), \overline{b}(\delta)\right)$ *and is again totally unstable for* $(\overline{b}(\delta), +\infty)$. *Generically, there exist period-two cycles in a left (right) neighborhood of* $\underline{b}(\delta)$ *that are locally saddle-point stable (unstable).*

Remark 2 Consider the production function from the social perspective as given in Definition 1 and recall from (5) that we can write it as follows

$$\tilde{y} = \left[(\beta_1 + b)\, k_y^{-\rho_y} + (\beta_2 - b)\right]^{-\frac{1}{\rho_y}} . \tag{23}$$

According to $b \gtrless \beta_2$, the following inequality holds: for any $\eta > 1$,

$$\begin{aligned} \left[(\beta_1 + b)\,(\eta k_y)^{-\rho} + (\beta_2 - b)\right]^{-\frac{1}{\rho}} &\gtrless \left[(\beta_1 + b)\,(\eta k_y)^{-\rho} + \eta^{-\rho}\,(\beta_2 - b)\right]^{-\frac{1}{\rho}} \\ &= \eta \left[(\beta_1 + b)\, k_y^{-\rho} + (\beta_2 - b)\right]^{-\frac{1}{\rho}} . \end{aligned}$$

If b is larger than β_2, the function $\tilde{y}$ exhibits increasing returns while if b is smaller than β_2 the function $\tilde{y}$ exhibits decreasing returns.

As we consider in Proposition 5 values of δ close to zero, the role of b on the local stability properties of the steady state is multiple. Indeed, starting from a low amount of externalities, an increase of b contributes to saddle-point stability and the existence of cycles through a flip bifurcation. But then if b is increased too much, total instability occurs since the returns to scale becomes increasing as shown in the previous Remark.

4 Concluding remarks

In this paper we have characterized the local dynamics of equilibrium paths depending on the size of external effects b. We have shown that when the consumption good is capital intensive, the effect of b on the local dynamics of equilibrium path depends on the value of the discount factor. If the discount factor is close enough to one and the capital intensity difference is large enough, local indeterminacy occurs for intermediary values of b while saddle-point stability is obtained when b is low enough or large enough. On the contrary, if the discount factor is low enough, local indeterminacy cannot occur. But the existence of equilibrium cycles and saddle-point stability require intermediary values of b while total instability is obtained when b is low enough or large enough.

5 Appendix

5.1 Proof of Lemma 1

We shall derive the first partial derivatives of $T(k_t, k_{t+1}, e_t)$ along an equilibrium path. The first order conditions derived from the Lagrangian are as below:

$$\alpha_1 \left(\frac{c_t}{K_{ct}}\right) - r_t = 0, \tag{A1.1}$$

$$\alpha_2 \left(\frac{c_t}{L_{ct}}\right) - w_t = 0, \tag{A1.2}$$

$$p_t \beta_1 \left(\frac{y_t}{K_{yt}}\right) - r_t = 0, \tag{A1.3}$$

$$p_t \beta_2 \left(\frac{y_t}{L_{yt}}\right) - w_t = 0. \tag{A1.4}$$

In the equilibrium the equation (2) is rewritten as

$$L_{yt} = \left(\frac{y_t^{-\rho} - (\beta_1 + b) K_{yt}^{-\rho}}{\beta_2 - b}\right)^{-\frac{1}{\rho}}. \tag{A1.5}$$

From the first order conditions (A1.1)-(A1.4),

$$\frac{\alpha_1 \beta_2}{\alpha_2 \beta_1} = \left(\frac{K_{ct}}{L_{ct}}\right)^{1+\rho} \left(\frac{L_{yt}}{K_{yt}}\right)^{1+\rho}.$$

Substituting $K_{ct} = k_t - K_{yt},\ L_{yt} = 1 - L_{ct}$ into the equation,

$$\frac{\alpha_1\beta_2}{\alpha_2\beta_1} = \left(\frac{k_t - K_{yt}}{1 - L_{yt}}\right)^{1+\rho}\left(\frac{L_{yt}}{K_{yt}}\right)^{1+\rho}. \tag{A1.6}$$

By solving equations (A1.5) and (A1.6) with respect to K_{yt} and substituting $y_t = k_{t+1}$, we have $K_{yt} = g(k_t,\ k_{t+1})$. From the equation (A1.1),

$$r_t = \alpha_1\left[\alpha_1 + \alpha_2\left(\frac{K_{ct}}{L_{ct}}\right)^{\rho}\right]^{-\frac{1+\rho}{\rho}}.$$

Using the equation (A1.6) we have

$$r_t = \alpha_1\left[\alpha_1 + \alpha_2\left(\frac{\alpha_1\beta_2}{\alpha_2\beta_1}\right)^{\frac{1+\rho}{\rho}}\left(\frac{g(k_t, k_{t+1})}{L_{ct}}\right)^{\rho}\right]^{-\frac{1+\rho}{\rho}}.$$

And then from (A1.5) r_t can be rewritten as the following equation by substituting $\left(\frac{g(k_t,k_{t+1})}{L_{yt}}\right)^{\rho} = \frac{\left(\frac{g(k_t,k_{t+1})}{y_t}\right)^{\rho}}{\beta_2-b} - \frac{(\beta_1+b)}{\beta_2-b}$ [6],

$$r_t = \alpha_1\left[\alpha_1 + \alpha_2\left(\frac{\alpha_1\beta_2}{\alpha_2\beta_1}\right)^{\frac{1+\rho}{\rho}}\left(\frac{\left(\frac{g(k_t,k_{t+1})}{y_t}\right)^{\rho}}{\beta_2 - b} - \frac{(\beta_1 + b)}{\beta_2 - b}\right)^{\rho}\right]^{-\frac{1+\rho}{\rho}}. \tag{A1.7}$$

Moreover from the equation (A1.3), we have

$$p_t = \frac{r_t}{\beta_1}\left(\frac{g(k_t, k_{t+1})}{y_t}\right)^{1+\rho}. \tag{A1.8}$$

Therefore we get T_1 and T_2 from the envelope theorem which gives

$$T_1 = r_t, \quad T_2 = -p_t.$$

■

5.2 Proof of Proposition 1

By definition k^* satisfies $T_2(k^*, k^*, e^*) + \delta T_1(k^*, k^*, e^*) = 0$ with $e^* = \hat{e}(k^*, k^*)$. In the steady state, $g^* = g(k^*, k^*)$ and $y^* = k^*$. Using Lemma 1, the Euler equation is

$$-\frac{r}{\beta_1}\left(\frac{g^*}{y^*}\right)^{1+\rho} + \delta r = 0.$$

[6]Substitute the equation (A1.5) into $\left(\frac{g}{L_{yt}}\right)^{\rho}$.

Thus,

$$g^* = (\delta\beta_1)^{\frac{1}{1+\rho}} k^*. \tag{A2.1}$$

As $y^* = k^*$ at the steady state, the equation (A1.5) becomes,

$$L_y^* = k^* \left(\frac{1 - (\beta_1 + b)(\delta\beta_1)^{\frac{-\rho}{1+\rho}}}{\beta_2 - b} \right)^{-\frac{1}{\rho}}. \tag{A2.2}$$

Using $K_c^* = k^* - K_y^*$ and $L_c^* = 1 - L_y^*$,

$$L_c^* = 1 - k^* \left(\frac{1 - (\beta_1 + b)(\delta\beta_1)^{\frac{-\rho}{1+\rho}}}{\beta_2 - b} \right)^{-\frac{1}{\rho}}, \tag{A2.3}$$

$$K_c^* = k^* \left(1 - (\delta\beta_1)^{\frac{1}{1+\rho}} \right). \tag{A2.4}$$

Then equation (A1.6) can be rewritten as follows;

$$\left(\frac{K_c^*}{L_c^*} \right)^{1+\rho} \left(\frac{L_y^*}{g^*} \right)^{1+\rho} = \frac{\alpha_1\beta_2}{\alpha_2\beta_1}. \tag{A2.5}$$

Substituting these input demand functions into the above equation and solving with respect to k^*, we can get

$$k^* = \left[1 + \left(\frac{\alpha_1\beta_2}{\alpha_2\beta_1} \right)^{\frac{-1}{1+\rho}} (\delta\beta_1)^{\frac{-1}{1+\rho}} \left(1 - (\delta\beta_1)^{\frac{1}{1+\rho}} \right) \right]^{-1} \left[\frac{1 - (\beta_1 + b)(\delta\beta_1)^{\frac{-\rho}{1+\rho}}}{\beta_2 - b} \right]^{\frac{1}{\rho}}.$$

Then k^* is well defined if and only if

$$(\delta\beta_1)^{\frac{-\rho}{1+\rho}} < \frac{1}{\beta_1 + b}.$$

■

5.3 Proof of Proposition 2

We give some lemmas in order to derive the characteristic roots.

Lemma 2 *At the steady state the following holds*

$$g_1 = \frac{1+\rho}{\Delta K_c},$$

$$g_2 = \left[\frac{(1+\rho) L_y^\rho + (1+\rho) \frac{L_y^{1+\rho}}{L_c}}{\Delta} \right] \frac{y^{-1-\rho}}{\beta_2 - b},$$

where

$$\Delta = \frac{1+\rho}{g} + \frac{1+\rho}{K_c} + \left((1+\rho) L_y^{\rho} + (1+\rho) \frac{L_y^{1+\rho}}{L_c} \right) \frac{\beta_1 + b}{\beta_2 - b} g^{-1-\rho}.$$

Proof. From equation (A2.5) we get

$$\frac{\alpha_1\beta_2}{\alpha_2\beta_1} = g^{-1-\rho} (k-g)^{1+\rho} \left(\frac{y^{-\rho} - (\beta_1 + b) g^{-\rho}}{\beta_2 - b} \right)^{-\frac{1+\rho}{\rho}} \left\{ 1 - \left(\frac{y^{-\rho} - (\beta_1 + b) g^{-\rho}}{\beta_2 - b} \right)^{-\frac{1}{\rho}} \right\}^{-1-\rho}.$$

Totally differentiating this equation, we have the following relationship,

$$\begin{aligned} &[(1+\rho) g^{-1} + (1+\rho)(k-g)^{-1} + (1+\rho) \left(\frac{y^{-\rho} - (\beta_1 + b) g^{-\rho}}{\beta_2 - b} \right)^{-1} \frac{\beta_1 + b}{\beta_2 - b} g^{-1-\rho} \\ &+ (1+\rho) \left\{ 1 - \left(\frac{y^{-\rho} - (\beta_1 + b) g^{-\rho}}{\beta_2 - b} \right)^{-\frac{1}{\rho}} \right\}^{-1} \left(\frac{y^{-\rho} - (\beta_1 + b) g^{-\rho}}{\beta_2 - b} \right)^{-\frac{1+\rho}{\rho}} \frac{\beta_1 + b}{\beta_2 - b} g^{-1-\rho}] dg \\ &= (1+\rho)(k-g)^{-1} dk + (1+\rho) \left(\frac{y^{-\rho} - (\beta_1 + b) g^{-\rho}}{\beta_2 - b} \right)^{-1-1} \frac{y^{-1-\rho}}{\beta_2 - b} dy \\ &+ (1+\rho) \left\{ 1 - \left(\frac{y^{-\rho} - (\beta_1 + b) g^{-\rho}}{\beta_2 - b} \right)^{-\frac{1}{\rho}} \right\}^{-1} \left(\frac{y^{-\rho} - (\beta_1 + b) g^{-\rho}}{\beta_2 - b} \right)^{-\frac{1+\rho}{\rho}} \frac{y^{-1-\rho}}{\beta_2 - b} dy. \end{aligned} \tag{A3.1.1}$$

Notice from equation (A1.5)

$$L_y^{-\rho} = \frac{y^{-\rho} - (\beta_1 + b) g^{-\rho}}{\beta_2 - b} \tag{A3.1.2}$$

and (A2.5)

$$\left(\frac{\alpha_1\beta_2}{\alpha_2\beta_1} \right)^{\frac{1}{1+\rho}} = \frac{K_c^*}{L_c^*} \left(\frac{g^*}{L_y^*} \right)^{-1}. \tag{A3.1.3}$$

Then substituting these equations and $dy_t = dk_{t+1}$ into (A3.1.1) gives

$$RHS = \frac{1+\rho}{K_c} dk_t + \left((1+\rho) L_y^{\rho} + (1+\rho) \frac{L_y^{1+\rho}}{L_c} \right) \frac{y^{1-\rho}}{\beta_2 - b} dk_{t+1},$$

$$LHS = \left[\frac{1+\rho}{g} + \frac{1+\rho}{K_c} + \left((1+\rho) L_y^{\rho} + (1+\rho) \frac{L_y^{1+\rho}}{L_c} \right) \frac{\beta_1 + b}{\beta_2 - b} g^{-1-\rho} \right] dg,$$

where we denote

$$\Delta \equiv \frac{1+\rho}{g} + \frac{1+\rho}{K_c} + \left((1+\rho) L_y^{\rho} + (1+\rho) \frac{L_y^{1+\rho}}{L_c} \right) \frac{\beta_1 + b}{\beta_2 - b} g^{-1-\rho},$$

and we derive

$$\Delta dg = \frac{1+\rho}{K_c} dk_t + \left[(1+\rho) L_y^{\rho} + (1+\rho) \frac{L_y^{1+\rho}}{L_c} \right] \frac{y^{1-\rho}}{\beta_2 - b} dk_{t+1}.$$

Therefore

$$dg = \frac{1+\rho}{\Delta K_c} dk_t + \frac{\left[(1+\rho) L_y^{\rho} + (1+\rho) \frac{L_y^{1+\rho}}{L_c} \right]}{\Delta} \frac{y^{1-\rho}}{\beta_2 - b} dk_{t+1}.$$

■

Lemma 3 *At the steady state the following holds*

$$g_1 y = (g - g_2 y) \frac{y}{g} \left(1 - \frac{K_c}{L_c} \frac{L_y}{K_y} \right)^{-1}$$

with g, K_c, L_y, L_c *respectively given by equations* $(A2.1) - (A2.4)$.

Proof. From equation (A1.5) we get

$$\left(L_y^{-\rho} + \frac{\beta_1 + b}{\beta_2 - b} K_y^{-\rho} \right) y^{-1} = \frac{y^{-\rho}}{\beta_2 - b} y^{-1}.$$

Substituting this equation into g_2,

$$g_2 = \frac{\left[(1+\rho) L_y^{\rho} + (1+\rho) \frac{L_y^{1+\rho}}{L_c} \right] \left(L_y^{-\rho} + \frac{\beta_1 + b}{\beta_2 - b} K_y^{-\rho} \right) y^{-1}}{\Delta}.$$

Using the expression of Δ we derive

$$g_2 y = g + \frac{(1+\rho)}{\Delta} \frac{L_y}{L_c} - \frac{(1+\rho)}{\Delta} \frac{g}{K_c}.$$

Then,

$$g - g_2 y = \frac{(1+\rho)}{\Delta K_c} g \left(1 - \frac{L_y}{L_c} \frac{K_c}{g} \right), \tag{A3.2.1}$$

$$g_1 y = \frac{1+\rho}{\Delta K_c} y. \tag{A3.2.2}$$

From equations (A3.2.1) and (A3.2.2), we finally get

$$g_1 y = (g - g_2 y) \frac{y}{g} \left(1 - \frac{K_c}{L_c} \frac{L_y}{K_y} \right)^{-1}.$$

■

Lemma 4 *Under Assumption* 1, *at the steady state,* $k_t = k_{t+1} = y_t = k^*$ *and the following holds*

$$\frac{V_{11}(k^*,\ k^*)}{V_{12}(k^*,\ k^*)} = -\frac{y}{g} \left(1 - \frac{K_c}{L_c} \frac{L_y}{K_y} \right)^{-1},$$

$$\frac{V_{22}\left(k^*,\ k^*\right)}{V_{12}\left(k^*,\ k^*\right)}=-\frac{g}{\beta_1 y}\left[\frac{\beta_1\left(\beta_2-b\right)}{\beta_2}\frac{K_c}{L_c}\frac{L_y}{g}+\left(\beta_2-b\right)\left(\frac{g}{L_y}\right)^{\rho}-\left(\frac{g}{y}\right)^{\rho}\right],$$

$$\frac{V_{21}\left(k^*,\ k^*\right)}{V_{12}\left(k^*,\ k^*\right)}=\frac{V_{22}\left(k^*,\ k^*\right)}{V_{12}\left(k^*,\ k^*\right)}\frac{V_{11}\left(k^*,\ k^*\right)}{V_{12}\left(k^*,\ k^*\right)},$$

where $g,\ K_c,\ L_y,\ L_c$ are given by equations $(A2.1)-(A2.4)$, respectively.

Proof. Let $V\left(k_t,\ k_{t+1}\right)$ denote $T_i\left(k_t,\ k_{t+1},\hat{e}\left(k_t,\ k_{t+1}\right)\right)$ for $i=1,\ 2$. By definition,

$$\begin{aligned}V_{11}^* &= \frac{\partial T_1}{\partial k_t}=\frac{\partial r}{\partial k_t},\\ V_{12}^* &= \frac{\partial T_1}{\partial k_{t+1}}=\frac{\partial r}{\partial k_{t+1}},\\ V_{21}^* &= \frac{\partial T_2}{\partial k_t}=-\frac{\partial p}{\partial k_t},\\ V_{22}^* &= \frac{\partial T_2}{\partial k_{t+1}}=-\frac{\partial p}{\partial k_{t+1}}.\end{aligned}$$

Computing the these equations, we have

$$V_{11}^*=\frac{\partial r}{\partial k_t}=-\left(1+\rho\right)\alpha_1^{-\frac{\rho}{1+\rho}}r^{\frac{1+2\rho}{1+\rho}}\frac{\alpha_2}{\beta_2-b}\left(\frac{\alpha_1\beta_2}{\alpha_2\beta_1}\right)^{\frac{\rho}{1+\rho}}\left(\frac{g}{y}\right)^{\rho}\frac{g_1}{g},$$

$$V_{12}^*=\frac{\partial r}{\partial k_{t+1}}=-\left(1+\rho\right)\alpha_1^{-\frac{\rho}{1+\rho}}r^{\frac{1+2\rho}{1+\rho}}\frac{\alpha_2}{\beta_2-b}\left(\frac{\alpha_1\beta_2}{\alpha_2\beta_1}\right)^{\frac{\rho}{1+\rho}}\left(\frac{g}{y}\right)^{\rho}\left(\frac{g_2 y-g}{yg}\right),$$

$$\frac{\partial p}{\partial k_t}=\frac{1}{\beta_1}\frac{\partial r}{\partial k_t}\left(\frac{g}{y}\right)^{1+\rho}+\left(1+\rho\right)\frac{r}{\beta_1}\left(\frac{g}{y}\right)^{1+\rho}\frac{g_1}{g},$$

$$\frac{\partial p}{\partial k_{t+1}}=\frac{1}{\beta_1}\frac{\partial r}{\partial k_{t+1}}\left(\frac{g}{y}\right)^{1+\rho}+\left(1+\rho\right)\frac{r}{\beta_1}\left(\frac{g}{y}\right)^{1+\rho}\left(\frac{g_2 y-g}{yg}\right).$$

From equation (A1.7),

$$\left(\frac{r}{\alpha_1}\right)^{\frac{\rho}{1+\rho}}=\left[\alpha_1+\alpha_2\left(\frac{\alpha_1\beta_2}{\alpha_2\beta_1}\right)^{\frac{\rho}{1+\rho}}\left(\frac{g}{L_c}\right)^{\rho}\right].$$

Substituting the above equation into V_{11}^*, and using (A3.1.2) and (A3.1.3) we obtain

$$V_{11}^*=-\left(1+\rho\right)r\left(\frac{g}{y}\right)^{\rho}\frac{g_1}{g}\left[\frac{\alpha_1\hat{\beta}_2}{\alpha_2}\left(\frac{\alpha_2\beta_1}{\alpha_1\beta_2}\right)\frac{K_c^*}{L_c^*}\frac{L_y^*}{g^*}+\hat{\beta}_2\left(\frac{y}{L_y}\right)^{\rho}\right]^{-1},$$

where

$$\mathcal{A}\equiv\frac{\alpha_1\hat{\beta}_2}{\alpha_2}\left(\frac{\alpha_2\beta_1}{\alpha_1\beta_2}\right)\frac{K_c^*}{L_c^*}\frac{L_y^*}{g^*}+\hat{\beta}_2\left(\frac{y}{L_y}\right)^{\rho}.$$

We can calculate $V_{21}^*,\ V_{12}^*$, and V_{22}^* as we did previously,

$$V_{21}^* = -(1+\rho)\frac{r}{\beta_1}\left(\frac{g}{y}\right)^{1+\rho}\frac{g_1}{g}\left[1-\left(\frac{g}{y}\right)^{\rho}\mathcal{A}^{-1}\right],$$

$$V_{12}^* = -(1+\rho)\,r\left(\frac{g}{y}\right)^{\rho}\left(\frac{g_2y-g}{yg}\right)\mathcal{A}^{-1},$$

$$V_{22}^* = -(1+\rho)\frac{r}{\beta_1}\left(\frac{g}{y}\right)^{1+\rho}\left(\frac{g_2y-g}{yg}\right)\left[1-\left(\frac{g}{y}\right)^{\rho}\mathcal{A}^{-1}\right].$$

Then we get

$$\frac{V_{11}^*}{V_{12}^*} = \frac{g_1y}{g_2y-g},$$

$$\frac{V_{22}^*}{V_{12}^*} = \frac{g}{\beta_1 y}\left[\mathcal{A}-\left(\frac{g}{y}\right)^{\rho}\right],$$

■

We shall now prove Proposition 2. From Lemma 4 the characteristic polynomial may be rewritten as

$$G(\lambda) = \left(\lambda+\frac{V_{11}^*}{V_{12}^*}\right)\left(\delta\lambda+\frac{V_{22}^*}{V_{12}^*}\right).$$

Then the characteristic roots are

$$\lambda_1 = -\frac{V_{11}^*}{V_{12}^*},\quad \lambda_2 = -\frac{V_{22}^*}{\delta V_{12}^*}. \tag{A3.3.1}$$

We can calculate $\frac{V_{11}^*}{V_{12}^*}$ and $\frac{V_{22}^*}{V_{12}^*}$ by substituting the following relationship

$$\frac{K_c}{L_c}\frac{L_y}{g} = \left(\frac{\alpha_1\beta_2}{\alpha_2\beta_1}\right)^{\frac{1}{1+\rho}},$$

$$\frac{g}{y} = (\delta\beta_1)^{\frac{1}{1+\rho}},$$

$$\left(\frac{g}{L_y}\right)^{\rho} = \frac{(\delta\beta_1)^{\frac{\rho}{1+\rho}}-\hat{\beta}_1}{\hat{\beta}_2},$$

$$g-g_2y = \frac{(1+\rho)}{\Delta K_c}g\left(1-\frac{L_y}{L_c}\frac{K_c}{g}\right),$$

and we obtain the first root by substituting all the above equations into the expressions given in Lemma 4

$$\lambda_1 = -\frac{1}{(\delta\beta_2)^{\frac{1}{1+\rho}}\left[\left(\frac{\alpha_1}{\alpha_2}\right)^{\frac{1}{1+\rho}}-\left(\frac{\beta_1}{\beta_2}\right)^{\frac{1}{1+\rho}}\right]}.$$

Moreover we can rewrite $\mathcal{A}$ by using these equations,

$$\mathcal{A} = \hat{\beta}_2 \frac{\alpha_1}{\alpha_2} \left(\frac{\alpha_2 \beta_1}{\alpha_1 \beta_2} \right)^{-\frac{\rho}{1+\rho}} + (\delta \beta_1)^{\frac{\rho}{1+\rho}} - \hat{\beta}_1.$$

From Lemma 4, we finally have the second characteristic root,

$$\lambda_2 = -\frac{(\delta\beta_2)^{\frac{1}{1+\rho}} \left[\frac{\beta_2 - b}{\beta_2} \left(\frac{\alpha_1}{\alpha_2} \right)^{\frac{1}{1+\rho}} - \frac{\beta_1 + b}{\beta_1} \left(\frac{\beta_1}{\beta_2} \right)^{\frac{1}{1+\rho}} \right]}{\delta}.$$

■

5.4 Proof of Proposition 3

Notice from (21) that $\lambda_1 > 0$. Denoting[7]

$$\delta_1 \equiv \beta_2^{-1} \left[(\beta_1/\beta_2)^{\frac{1}{1+\rho}} - (\alpha_1/\alpha_2)^{\frac{1}{1+\rho}} \right]^{-1-\rho} > 1$$

then we obtain $\lambda_1 = (\delta_1/\delta)^{\frac{1}{1+\rho}} > 1$ for $0 < \delta < 1$. Since $(\beta_1 + b)/\beta_1 > 1$ and $(\beta_2 - b)/\beta_2 < 1$, $\lambda_2(b)$ is always positive. ■

5.5 Proof of Proposition 4

If $(\alpha_1\beta_2)^{\frac{1}{1+\rho}} - (\alpha_2\beta_1)^{\frac{1}{1+\rho}} > \alpha_2^{\frac{1}{1+\rho}}$ and $\delta \in (\delta_3, 1)$, then $-1 < \lambda_1 < 0$. The size of $\lambda_2(b)$ is determined in the following way. Notice that $\lambda_2(b)$ is increasing in b. For $b = 0$, $\lambda_2(0) = 1/\delta\lambda_1 < -1$ by the above hypothesis and for $b = \beta_2$, $\lambda_2(\beta_2) = (\delta\beta_1)^{\frac{-\rho}{1+\rho}}$.

(i) If $-1 < \rho < 0$, $\lambda_2(\beta_2) < 1$. Therefore there exist $\underline{b}(\delta) \in (0, \beta_2)$ and $\overline{b}(\delta) > \beta_2$ such that $\lambda_2 < -1$ for $b \in (0, \underline{b}(\delta))$, $-1 < \lambda_2 < 1$ for any $b \in (\underline{b}(\delta), \overline{b}(\delta))$ and $\lambda_2 > 1$ for any $b > \overline{b}(\delta)$.

(ii) If $\rho = 0$, $\lambda_2(\beta_2) = 1$. Therefore $\lambda_2(b) < -1$ for $b \in (0, \beta_2 - 2\alpha_2)$, $-1 < \lambda_2(b) < 1$ for $b \in (\beta_2 - 2\alpha_2, \beta_2)$ and $\lambda_2(b) > 1$ for $b > \beta_2$.

(iii) If $\rho > 0$, $\lambda_2(\beta_2) > 1$. Therefore there exist $\underline{b}(\delta)$ and $\overline{b}(\delta)$ in $(0, \beta_2)$ such that $\lambda_2(b) < -1$ for $b \in (0, \underline{b}(\delta))$, $-1 < \lambda_2(b) < 1$ for $b \in (\underline{b}(\delta), \overline{b}(\delta))$, and $\lambda_2(b) > 1$ for $b > \overline{b}(\delta)$.

In each of these three cases, when $b = \underline{b}(\delta)$, $\lambda_2(b) = -1$ and $\lambda_2'(b)|_{b=\underline{b}(\delta)} > 0$. It follows that $b = \underline{b}(\delta)$ is a flip bifurcation value. The result follows from the flip bifurcation Theorem (see Ruelle [8]). ■

Received: April 2008. Revised: May 2008.

[7] Note that $\delta_1 = \alpha_2 \left[(\alpha_2\beta_1)^{\frac{1}{1+\rho}} - (\alpha_1\beta_2)^{\frac{1}{1+\rho}} \right]^{-1-\rho} > \frac{\alpha_2}{(\alpha_2\beta_1)} = \frac{1}{\beta_1} > 1.$

References

[1] J. Benhabib and R.E. Farmer, *Indeterminacy and Increasing Returns*, Journal of Economic Theory, **20** (1994), 19–41.

[2] J. Benhabib and R.E. Farmer, *Indeterminacy and Sunspots in Macroeconomics*, in Handbook of Macroeconomics, Edited by J.B. Taylor and M. Woodford, North-Holland, Amsterdam, Holland, (1999), pp. 387–448.

[3] J. Benhabib and K. Nishimura, *Indeterminacy and Sunspots with Constant Returns*, Journal of Economic Theory, **81** (1998), 58–96.

[4] J. Benhabib and K. Nishimura, *Indeterminacy Arising in Multisector Economies*, Japanese Economic Review, **50** (1999), 485–506.

[5] J. Benhabib, K. Nishimura, and A. Venditti, *Indeterminacy and Cycles in Two-Sector Discrete-Time Model*, Economic Theory, **20** (2002), 217–235.

[6] D. Cass and K. Shell, *Do Sunspots Matter ?*, Journal of Political Economy, **91** (1983), 193–227.

[7] K. Nishimura and A. Venditti, *Capital Depreciation, Factor Substitutability and Indeterminacy*, Journal of Difference Equations and Applications, **10** (2004), 1153–1169.

[8] D. Ruelle, *Elements of Differentiable Dynamics and Bifurcation Theory*, Academic Press, San Diego, 1989.

CUBO A Mathematical Journal
Vol.10, Nº03, (83–91). October 2008

Many-Ended Complete Minimal Surfaces Between Two Parallel Planes in $\mathbb{R}^3$

Francisco Brito*
Universidade Federal de Pernambuco – UFPE,
CCEN-Departamento de Matemática,
50.740.540, Recife-PE, Brazil
email: `brito@dmat.ufpe.br`

ABSTRACT

We use some special convergent Hadamard gap series to provide examples of complete minimal surfaces of many different conformal types between two parallel planes in three dimensional Euclidean space.

RESUMEN

Nosotros usamos algunas series convergentes especiales de Hadamard para dar ejemplo de superficies mínimas completas de varios diferentes tipos conforme entre dos planos paralelos en espacios euclideanos de dimensión tres.

Key words and phrases: *Complete minimal surfaces, Hadamard gap series.*

Math. Subj. Class.: *53A10.*

*I am grateful to Profs. Marcus V. M. Wanderley and Maria Elisa Oliveira for helpful hints and discussions.

1 Introduction

In [7] F. Xavier and L. P. M.Jorge established the existence of complete non-planar minimal surfaces between two parallel planes in $\mathbb{R}^3$. Their technique consisted of an artful use of Runge's Theorem to prove the existence of holomorphic functions on the unit disc $\mathbb{D}$ with the right properties they needed. Later their method was adapted by others to produce new surfaces as above with new features, like having cylindrical type as in [6] or being non-orientable as in [3].

Another way for rendering complete minimal surfaces between two parallel planes in $\mathbb{R}^3$ was developed in [1]. This method consisted mainly in proving the existence of bounded holomorphic functions h in $\mathbb{D}$, given by lacunary power series, and such that $\displaystyle\int_\gamma |h'(z)|^2|dz| = \infty$ for all divergent paths γ in the unit disc.

In this paper we intend to show the flexibility of the second method by producing examples of complete minimal surfaces between two parallel planes in $\mathbb{R}^3$ of the following conformal types:

1. A disc with finitely many points removed.
2. Any annulus, $0 < r < |z| < R$.
3. Any annulus as above with finitely many points removed.

This work is organized as follows: In §2 we give some definitions and prove the lemmas that will be needed in the other sections. In the three remaining sections we describe the examples of the types above.

Remark 1.1 *This work was written about fourteen years ago and circulated as a preprint for some time. Meanwhile N.Nadirashvili proved in [4] the existence of bounded complete minimal surfaces in* $\mathbb{R}^3$.

2 Some definitions and lemmas

Lacunary power series were defined in [1] with the restriction that they would have radius of convergence 1. This is just a mild technical point. Here we use any positive real number R as radius of convergence and make the necessary changes for having an analogue of Theorem 2 of [1].

Definition 2.1 *A convergent power series* $\displaystyle\sum_{k=0}^{\infty} a_k z^{n_k}$ *is lacunary if there exists a real number* $q > 1$ *such that* $\dfrac{n_{k+1}}{n_k} \geq q$ *for all* $k = 0, 1, \ldots$.

Lemma 2.2 *Let* $h(z) = \displaystyle\sum_{k=0}^{\infty} a_k z^{n_k}$ *be a lacunary power series of radius of convergence* $R > 0$*, and suppose that the following three conditions hold:*

a) $\displaystyle\sum_{k=0}^{\infty} |a_k| R^{n_k}$ *converges.*

b) $\displaystyle\lim_{k\to\infty} R^{n_k}|a_k| min\left\{\frac{n_{k+1}}{n_k}\,,\,\frac{n_k}{n_{k-1}}\right\} = \infty.$

c) $\displaystyle\sum_{k=0}^{\infty} R^{2n_k}|a_k|^2 n_k$ *diverges.*

Then h is bounded in $\mathbb{D}_R = \{z \in \mathbb{C}; |z| < R\}$, and for all divergent paths γ in $\mathbb{D}_R$, one has that $\displaystyle\int_\gamma |h'(z)|^2 |dz| = \infty.$

Proof. The change of variable $z = Rw$ together with (a) show that h is bounded. The same change of variable and Theorem 2 of [1] finish the proof. ■

Lemma 2.3 *If h satisfies the conditions of the above lemma in $\mathbb{D}_R$, $H(z) = h(z^{-1})$ is a bounded holomorphic function in $A_R = \{z \in \mathbb{C}; |z| > R^{-1}\}$, and for all divergent paths γ in A_R tending to a point of $|z| = R^{-1}$ one has that* $\displaystyle\int_\gamma |H^{'}(z)|^2 |dz| = \infty.$

Proof. The change of variable $z = w^{-1}$ and Lemma 2.2 prove this assertion. ■

Now, given $r, R \in \mathbb{R}$, $0 < r < R$, let $\Omega_{R,r}$ be the annulus $r < |z| < R$.

Lemma 2.4 *Suppose that $h_1(z) = \displaystyle\sum_{k=0}^{\infty} a_k z^{n_k}$ and $h_2(z) = \displaystyle\sum_{k=0}^{\infty} b_k z^{m_k}$ are lacunary power series that satisfy the conditions of Lemma 2.2, and have radii of convergence R and r^{-1} respectively, with $0 < r < R$. Suppose further that $h_1^{'}(z)$ and $h_2^{'}(z^{-1})$ do not vanish in $|z| = r$ and $|z| = R$ respectively. Then, for all divergent paths γ in $\Omega_{R,r}$ one has that* $\displaystyle\int_\gamma |h_1^{'}(z)|^2 |(h_2(z^{-1}))'|^2 |dz| = \infty.$

Proof. A divergent path in $\Omega_{R,r}$ either approach $|z| = r$ or $|z| = R$. Suppose that γ is a divergent path that approaches $|z| = r$. Since $h_1^{'}$ is holomorphic in a neighborhood of that circle and does not vanish at any point of it, it follows that there is a perhaps smaller neighborhood U of $|z| = r$ having compact closure $\overline{U}$, and such that $\displaystyle\inf_{z\in\gamma\cap\overline{U}} \{|h_1^{'}(z)|^2\} = A > 0$. Consequently, if $\tilde{\gamma}$ denotes the portion of γ inside $\overline{U}$ one has that

$$\int_\gamma |h_1^{'}(z)|^2 |(h_2(z^{-1}))^{'}|^2 |dz| \geq A^2 \int_{\tilde{\gamma}} |(h_2(z^{-1}))^{'}|^2 |dz| = \infty$$

by Lemma 2.3. The rest of the proof follows in a similar way. ■

In the next sections we will use mainly the Weierstrass representation for minimal surfaces in $\mathbb{R}^3$ (see [5]) and the three lemmas above.

3 Minimal immersions of a disc with finitely many points removed

Let $\Omega = \mathbb{D}_R - \{a_1, a_2, \ldots, a_n\}$, where $\mathbb{D}_R$ is the open disc of radius R centered at the origin and a_1,a_2,...,a_n are distinct points of $\mathbb{D}_R$.

Theorem 3.1 *There exist complete minimal immersions $\mathcal{M}$ of Ω between two parallel planes of $\mathbb{R}^3$. Furthermore the ends of $\mathcal{M}$ corresponding to the points a_1,a_2,...,a_n are all planar and have index one.*

Proof. To avoid notational inconveniences we will prove separately the cases $n = 1$ and $n > 1$. In the first case we take $\Omega = \mathbb{D}_R - \{a\}$, for some $a \in \mathbb{D}_R$. Using the Weierstrass representation, set

$$f(z) = (z-a)^{-2} \text{ and } g(z) = (z-a)^2 h'(z)$$

where h is any function as in Lemma 2.2. Clearly the data above defines a minimal immersion $\mathcal{M}$ of Ω in $\mathbb{R}^3$. Moreover, $\mathcal{M}$ is also complete for the metric is given by

$$\lambda(z)|dz| = \tfrac{1}{2}\{|z-a|^{-2} + |z-a|^2|h'(z)|^2\}|dz|,$$

and because $x_3(z) = \operatorname{Re}(h(z))$, it follows from the properties of h that the third coordinate of $\mathcal{M}$ is bounded.

The end corresponding to a is of course planar because fg is holomorphic at that point and have index one because f has a pole of order two at a and fg^2 vanishes at that same point. For information on the behavior of ends of complete minimal surfaces see [2]. ∎

Now we consider the case $n > 1$. In the Weierstrass representation for $\mathcal{M}$ set

$$f(z) = F(z)^{-1} \exp\left\{\sum_{j=1}^{n} A_j F_j(z)\right\} \quad \text{and} \quad g(z) = h'(z)F(z),$$

where h is a function as in Lemma 1.1, $F(z) = \prod_{j=1}^{n}(z-a_j)^2$ the functions F_j satisfy $(z-a_j)^2 F_j^{'}(z) = F(z)$, and the A_j are constants to be chosen so that

$$\int_\sigma f(z)dz = 0 \text{ for all closed curves } \sigma \text{ in } \Omega.$$

Since fg and fg^2 have holomorphic extensions to all of $|z| < R$, it follows that this will be enough to exclude the possibility of real periods appearing in the Weierstrass representation of $\mathcal{M}$. An easy computation shows that the choice $A_j = F_j^{''}(a_j)(F_j^{'}(a_j))^{-2}$, $j = 1, \ldots, n$ solves the problem.

Observe that the metric $\lambda(z)|dz|$ on $\mathcal{M}$ is given by

$$2\lambda(z)|dz| = \{|F(z)|^{-1} + |F(z)||h'(z)|^2\}\left|\exp\left\{\sum_{k=1}^{n} A_k F_k(z)\right\}\right||dz|,$$

and since there is a positive real C such that $\left|\exp\left\{\sum_{k=1}^{n} A_k F_k(z)\right\}\right| \geq C,\ z \in \Omega$, it follows from the properties of h and that f has poles of order 2 at the a_j that $\mathcal{M}$ is complete. Also, $x_3(z) = \mathrm{Re}\int h'(z)\exp\{\sum_{j=1}^{n} A_j F_j(z)\}dz$ is bounded in Ω. This can be seen in the following way: $\exp\left\{\sum_{k=1}^{n} A_k F_k(z)\right\}$ and its derivatives as well as h are all bounded holomorphic functions in Ω. As a matter of fact they are bounded in $\mathbb{D}_R$, so, by integration by parts, it follows that $\int h'(z)\exp\{\sum_{j=1}^{n} A_j F_j(z)\}dz$ is bounded in $\mathbb{D}_R$, so x_3 is also bounded.

By an argument similar to the one done in the case $n = 1$ we conclude that the ends corresponding to the points a_j are all planar and have index one. ■

4 Complete minimal annuli between two parallel planes in $\mathbb{R}^3$

All the examples of complete minimal annuli in [6] have the conformal type of an annulus of the form $R^{-1} < |z| < R$. Here we give examples of all possible annuli $0 < r < |z| < R < \infty$.

Theorem 4.1 *Given any annulus $\Omega_{R,r}$ there is a complete minimal immersion of it in $\mathbb{R}^3$ with one coordinate bounded.*

Proof. Take any two functions h_1 and h_2 as in Lemma 1.3, say $h_1(z) = \sum_{k=0}^{\infty} a_k z^{n_k}$ and $h_2(z) = \sum_{l=0}^{\infty} a_l z^{m_l}$, with radii of convergence R and r^{-1} respectively, and such that all the n_k and m_l are simultaneously either even or odd. Then in the Weierstrass representation we set in $\mathbb{D}_{R,r}$,

$$f(z) = 1 \quad \text{and} \quad g(z) = h_1'(z)H_2'(z),$$

where $H_2(z) = h_2(z^{-1})$. Because f is constant, and g is an even function where defined, it follows that for all closed curves γ in $\Omega_{R,r}$ one has that

$$\int_\gamma f dz = \int_\gamma fg dz = \int_\gamma fg^2 dz = 0.$$

Thus, the minimal surface so obtained is in fact well defined. Furthermore, since the metric $\lambda(z)|dz|$ is given by

$$2\lambda(z)|dz| = \left(1 + |h_1'(z)|^2|H_2'(z)|^2\right)|dz|$$

it follows from Lemma 1.3 that it is complete. It remains to prove only that one of the coordinates of that immersion is bounded.

Since $x_3(z) = Re\int g dz = Re\int h_1'(z)H_2'(z)dz$, it is enough to prove that $\int g dz$ is a bounded holomorphic function in $\Omega_{R,r}$. Consider $\rho \in \mathbb{R}$ such that $r < \rho < R$ and define the sets

$$A_1 = \Omega_{R,r} \cap \{z \in \mathbb{C} \text{ such that } |z| \leq \rho\} \text{ and } A_2 = \Omega_{R,r} \cap \{z \in \mathbb{C} \text{ such that } |z| \geq \rho\}.$$

We then observe that $h_1'(z)$, its derivatives and H_2 are bounded in A_1 and the same happens to H_2', its derivatives and h_1 in A_2. So, by integration by parts we can conclude that $\int g dz$ is bounded in both A_1 and A_2, hence in $\Omega_{R,r}$. ■

5 The case of a annulus with finitely many points removed

Let $A = \{a_1, \ldots, a_n\}$ be a set of distinct points of $\Omega_{R,r}$ such that $A \bigcap (-A) = \emptyset$, and set $\Omega = \Omega_{R,r} - A$.

Theorem 5.1 *There is a complete minimal immersion of Ω between two parallel planes of $\mathbb{R}^3$. Furthermore, the ends corresponding to $a_1, \ldots, a_n$ are planar.*

Proof. First suppose $n = 1$ and take $\Omega = \Omega_{R,r} - \{a\}$ with $a \in \Omega_{R,r}$. Consider holomorphic functions h_1 and h_2 as in Theorem 3.1, and define the surface $\mathcal{M}$ using the Weierstrass representation by taking

$$f(z) = (z-a)^{-2} \quad \text{and} \quad g(z) = (z-a)^2 h_1'(z)H_2'(z),$$

keeping the notation of Theorem 3.1. Then f is holomorphic in a neighborhood of $|z| \leq r$, and has residue zero at a, so $\int_\sigma f dz = 0$, for all closed curves σ inside Ω.

Because $fg = h_1'H_2'$ is an even holomorphic function in $\Omega_{R,r}$ it follows that $\int_\sigma fg dz = 0$ for all closed curves σ in Ω too.

Now we must make one more choice in order to have $\int_\sigma fg^2 dz = 0$ for all closed curves σ in Ω. The idea is to start with h_1 and h_2 having no low powers in their power series expansions. Since $f(z)g^2(z) = (z-a)^2(h_1'(z))^2(H_2'(z))^2$ is holomorphic in Ω, by expanding $(z-a)^2$, it is clear that the only term that may cause problems is $-2az(h_1'(z))^2(H_2'(z))^2$ because the other two terms are even functions. Hence if the functions h_1 and h_2 are chosen to satisfy Theorem 3.1 and have the form

$$h_1(z) = \sum_{k=1}^{\infty} a_k z^{1+2^{2^k}} \text{ and } h_2(z) = \sum_{k=1}^{\infty} b_k z^{1+2^{2^k}}$$

there is no term in z^{-1} in the Laurent expansion of $z(h_1'(z))^2(H_2'(z))^2$. Hence, for all closed curves σ in Ω, $\int_\sigma fg^2 dz = 0$ as wanted. Besides, as in Theorem 2.1, the end corresponding to the point a is planar and have index one.

In order to study the case $n > 1$ define the following holomorphic functions in Ω: $F(z) = \prod_{j=1}^{n}(z-a_j)^2$, $F_k(z) = (z-a_k)^{-2}F(z)$ and $G_k'(z) = zF_k(z)F_k(-z)$ for $k = 1, \ldots, n$, and finally $H(z) = \sum_{j=1}^{n} A_j G_j(z)$, where the constants A_j are to be determined so that, if the immersion $\mathcal{M}$ of Ω is defined in terms of the Weierstrass representation by setting

$$f(z) = (F(z))^{-1} \exp H(z) \text{ and } g(z) = F(z) \exp\left\{-\frac{1}{2}H(z)\right\} h_1'(z)H_2'(z),$$

then f has residue zero at all the points a_j. As before, the functions h_1 and h_2 are chosen satisfying the conditions of Lemma 1.3 and the exponents are chosen so that $\int_{|z|=\rho} f(z)g(z)^2 dz = 0$, for $r < \rho < R$.

It must be pointed out that once these constants A_j are determined the rest is done quite easily as follows: First we observe that H is an even holomorphic function in $\Omega_{R,r}$, and the same happens to h_1' and H_2', thus fg is an even holomorphic function in $\Omega_{R,r}$ and so $\int_\sigma fg dz = 0$ for all closed curves σ in $\Omega_{R,r}$ as wanted.

Furthermore,

$$2\lambda(z)|dz| = |F(z)|^{-1}|\exp H(z)| + |F(z)||h_1'(z)|^2|H_2'(z)|^2,$$

hence, repeating the reasoning in the proof of Theorem 2.1 we conclude that $\lambda(z)|dz|$ is complete and x_3 is bounded.

Now we determine the constants A_j. First, we observe that all the poles of f have order two and that for $j = 1, \ldots, n$, $(z-a_j)^2 f(z) = \dfrac{\exp H(z)}{F_j(z)}$, thus

$$\begin{aligned} \frac{d}{dz}\{(z-a_j)^2 f(z)\} &= \frac{\exp H(z)}{F_j^2(z)}\left[H'(z)F_j(z) - F_j'(z)\right] \\ &= \frac{\exp H(z)}{F_j^2(z)}\left[F_j(z)\left\{\sum_{k=1}^{n} A_k z F_k(z)F_k(-z)\right\} - F_j'(z)\right]. \end{aligned}$$

So, the residue of f at a_j is zero if and only if

$$F_j(a_j)\left\{\sum_{k=1}^{n} A_k a_j F_k(a_j)F_k(-a_j)\right\} - F_j^{'}(a_j) = 0.$$

Since $F_k(a_j) = 0$ for $k \neq j$, $F_j(a_j),\ F_j(-a_j) \neq 0$ and $a_j \neq -a_k$, for $1 \leq j,k \leq n$, it follows that $a_j F_j^2(a_j)F_j(-a_j)A_j - F_j^{'}(a_j) = 0$, for each j, $1 \leq j \leq n$, hence

$$A_j = \frac{F_j^{'}(a_j)}{a_j F_j^2(a_j)F_j(-a_j)}, \text{ for } j = 1, \dots, n.$$

To finish the proof it is enough to show that we can choose h_1 and h_2 in such a way that $\int\limits_{|z|=\rho} f(z)g(z)^2 dz = 0$, for $r < \rho < R$.

Since $f(z)g^2(z) = F(z)[h_1^{'}(z)H_2^{'}(z)]^2$, and F has degree $2n$, if we define

$$h_1(z) = \sum_{k=1}^{\infty} a_k z^{1+2n+2^{2^k}} \quad \text{and} \quad h_2(z) = \sum_{k=1}^{\infty} b_k z^{1+2n+2^{2^k}}$$

we are done. Also the observations about the ends in the other cases are valid here without change. ■

The assumption that $(-A) \cap A = \emptyset$ is not really needed. It is just a technical difficulty that can be easily overcome as follows:

Corollary 5.2 *If A is any finite subset of $\Omega_{R,r}$ there is a complete minimal immersion of $\Omega = \Omega_{R,r} - A$ between two parallel planes of $\mathbb{R}^3$. Furthermore, the ends corresponding to the points of A are planar.*

Proof. Induction on the number of elements of A shows that there exists a transformation

$$\pi : \Omega \longrightarrow \Omega^{'},$$

where $\pi(z) = z^{2^p}$ for some positive integer p and $\Omega^{'}$ satisfies the condition of Theorem 5.1.

It is clear that $(\pi\,,\Omega)$ is an unramified covering of $\Omega^{'}$, and by Theorem 5.1, there is a complete minimal immersion X of $\Omega^{'}$ between two parallel planes of $\mathbb{R}^3$. So $X\circ\pi$ is also a complete minimal immersion of Ω in $\mathbb{R}^3$ with the same properties as before. ■

Received: February 2008. Revised: June 2008.

References

[1] F.F. de Brito, *Power series with hadamard gaps and hyperbolic complete minimal surfaces*, Duke Math. Journal **68** (1992), 297–300.

[2] L.P.M. Jorge and W. Meeks, *The topology of complete minimal surfaces opf finite Gaussian curvature*, Topology **22** (1983), 203–221.

[3] F. Lopez, *A non orientable complete minimal surface in $\mathbb{R}^3$ between two parallel planes*, Proceedings of the Amer. Math. Soc **103** (1988), 913–917.

[4] N. Nadirashvili, *Hadamard and Calabi-Yau's conjectures on negatively curved and minimal surfaces*, Invent. Math. **126** (1996), 457–465.

[5] R. Osserman, *A survey of minimal surfaces*, Van Nostrand, New York, 1969.

[6] H. Rosenberg and E. Toubiana, *A cylindrical type minimal surface in a slab in $\mathbb{R}^3$*, Bull. Sc. Math. **111** (1987), 241–245.

[7] F. Xavier and L.P.M. Jorge, *A complete minimal surface in $\mathbb{R}^3$ between two parallel planes*, Ann. of Math. **112** (1980), 203–206.

CUBO A Mathematical Journal
Vol.10, Nº 03, (93–102). October 2008

Dualities Useful in Bond Percolation

Pedro Ferreira de Lima
Unversidade Regional do Cariri – URCAC
Av. Leão Sampaio S/N, CEP.63.010-970, Crajubar-CE, Brasil
email: `limapf@yahoo.com.br`

and

André Toom
Universidade Federal de Pernambuco – UFPE,
CCEN-Departamento de Estatística, CEP.50.740-540, Recife-PE, Brasil
email: `toom@de.ufpe.br`

ABSTRACT

We state four facts about dual pairs of graphs drawn in a plane. These facts pertain respectively to finite non-oriented, infinite non-oriented, finite oriented, and infinite oriented graphs. We do not include proofs of the former two facts (although we have them), but show that these facts are "evident" in some naive sense. Then we deduce the latter two facts from the former two ones. All of these facts can be used to obtain upper estimations for critical values of two-dimensional percolation models and we present three references and one example to illustrate this.

RESUMEN

Nosotros establecemos cuatro hechos acerca de pares duales de grafos en el plano. Estos hechos se relacionan respectivamente con grafos finito no-orientado, infinito no-orientado, finito orientado y infinito orientado. Nosotros no incluimos demostraciones de los dos anteriores hechos (aunque tenemos estas) pero demostramos que estos hechos son “evidentes” en algun sentido ingenuo. Entonces deducimos los dos ultimos hechos desde los dos anteriores. Todos estos hechos pueden ser usados para obtener estimativas por arriba para valores críticos de modelos de infiltración en dimensión dos y presentamos tres referencias y un ejemplo para ilustrar esto.

Key words and phrases: *Percolation, critical values, oriented planar graphs, duality.*

Math. Subj. Class.: *05C10, 60K35, 82B43, 94C15.*

1 Introduction

We state four facts about dual pairs of graphs drawn in a plane. These facts pertain respectively to finite non-oriented, infinite non-oriented, finite oriented, and infinite oriented graphs. All of these facts can be used to obtain upper estimations for critical values of two-dimensional percolation models and we present three references and one example to illustrate this. We call the former two facts the “main lemma” and the latter two facts “theorem”. We do not prove the main lemma although we have a proof in [L.2002]; instead we show that it is “evident” in some naive sense. Then we deduce our theorem from the main lemma.

We consider graphs with a finite or countable sets V of vertices and E of edges. Every edge connects two vertices, which are called its ends (and which may coincide). One and the same pair of vertices may be connected by several edges. A *finite path* is a finite sequence “ vertex-edge-vertex-edge-…-edge-vertex” in which every edge connects the vertices between which it is placed in this sequence and in which some vertices and/or edges may coincide. An *infinite path* is an infinite sequence “vertex-edge-vertex-edge-… ” with the same properties.. A path is called *self-avoiding* if all the vertices in its sequence are different. A *contour* is a finite path in which the initial and final vertices coincide. A contour is *self-avoiding* if all its vertices are different except, of course, the first and last vertices, which coincide. We say that a graph G is *connected* if every two vertices of G are connected by a path in G. All the graphs considered in this paper are assumed to be connected unless stated otherwise. We assume that every vertex is an end of only a finite number of edges. Therefore, if one of the sets V or E is finite, the other is finite also. If V and E are finite, we call G *finite*, otherwise G is *infinite*.

We consider *non-oriented* and *oriented* percolation models on graphs depending on the way in which we attribute certain states to their edges. In the *non-oriented model* each edge of a graph can be *open* or *closed*, independently of all the other edges. In the *oriented model* we distinguish two directions of each edge, and every edge can be open or closed in each direction independently of the state of the other direction of the same edge and states of all the other edges. Henceforth we shall write simply graphs instead of percolation models on graphs when it does not produce confusion. A path in a non-oriented graph is *open* if all its edges are open. A path in an oriented graph is *open* in a certain direction if all its edges are open in this direction. In a non-oriented graph a contour is open if all its edges are open. In an oriented graph a contour is open in a certain direction if all its edges are open in this direction.

Now let us speak about drawing graphs in a plane $\mathbb{R}^2$. A *curve* is a continuous function $f : [0,1] \to \mathbb{R}^2$. The points $f(0)$ and $f(1)$ are called the *ends* of this curve. A curve is called *self-avoiding* if

$$\forall x, y \in [0,1] \ : \ x \neq y \Rightarrow f(x) \neq f(y).$$

A curve is called *polygonal* if it is piecewise affine, that is there are

$$0 = t_0 < t_1 < \dots t_n = 1$$

such that $f(t)$ is an affine function of t in every segment $[t_{k-1}, t_k]$, $k = 1, \dots n$. The points $f(t_k)$, $0 < k < n$, are called *corners* and the sets

$$\{f(t) \mid t_{k-1} \leq t \leq t_k\}, \ k = 1, \dots, n$$

are called *pieces*. Henceforth we consider only polygonal curves. This approach is not unusual. We say that a graph G is *drawn in a plane* if the following conditions are satisfied:

1. Each vertex v of G is represented by a point $P(v)$ in the plane, so that different vertices are represented by different points. We denote the set of points representing all vertices of G by $P(V)$.
2. Each edge e of G is represented by a polygonal curve f_e, where:
   a) $f_e(0)$ and $f_e(1)$ represent the ends of this edge.
   b) The curve f_e is self-avoiding, except when e is a loop, in which case $f_e(0) = f_e(1)$.
3. If $e_i \neq e_j$ are two differents edges, the corresponding curves have no common points, except common ends, which represent common ends of e_i and e_j when they have such ones.
4. Each bounded subset of the plane intersects only a finite (or empty) set of curves representing edges.

A *closed curve* is a polygonal function f from a circle S^1 to $\mathbb{R}^2$. A *Jordan curve* is a closed self-avoiding curve, which means that the values of this function are different for different elements of S^1. According to the well-known Jordan theorem, any Jordan curve separates the remaining part of the plane into two open regions, one bounded, the other unbounded, so that it intersects any curve whose ends belong to different regions.

Notice that every finite path of a graph drawn in a plane is represented by a curve and every contour is represented by a closed curve, which are self-avoiding if and only if the original path, respectively contour is self-avoiding.

Given a graph G drawn in a plane, we call $P(G)$ the union of representations of its edges. Since $P(G)$ is closed, its complement is open. Connected components of this complement are called *faces* of G.

We say that two connected graphs G and G' drawn in the same plane are *dual* if they satisfy the following conditions. (Here and elsewhere we may write simply vertices and edges, while in fact we mean their representations in the plane.)

1. There is a 1-to-1 correspondence between the faces of G and vertices of G', called *duality*, such that each face of G contains its dual vertex of G'.
2. There is a 1-to-1 correspondence between the vertices of G and faces of G', called *duality*, such that each vertex of G is in the dual face of G'.
3. There is a 1-to-1 correspondence between the edges of G and edges of G' called *duality*, such that representation of each edge of each graph crosses the representation of its dual edge of the dual graph only in one point, which is not a corner or end of any of them. Representations of edges of G and G', which are not dual, have no common points.

This kind of duality is well-known and described in many textbooks. Observe that this kind of duality is a symmetric relationship, that is, if G' is dual of G, then G is dual of G'. Of course, if a graph is finite, its dual is finite too.

Now from duality of graphs drawn in a plane we go to dualities of percolation models on these graphs. The rules (1) and (2) are the central point of our definitions.

Rule for a dual pair of non-oriented graphs $(G,\ G')$:

Every edge of graph G' is open if and only if the dual edge of graph G is closed. (1)

Given a non-oriented graph G, we say that:

a) vertices α and β are *reachable* from each other if there is an open finite self-avoiding path connecting them.

b) vertex α and ∞ are *reachable* from each other if there is an open infinite self-avoiding path which starts at α.

Main Lemma. Let $(G,\ G')$ be a dual pair of non-oriented graphs, satisfying the rule (1). Then:

a) *If G is finite*: Two vertices α and β are not reachable from each other in G if and only if there is an open self-avoiding contour in G', whose representation in the plane leaves $P(\alpha)$ and $P(\beta)$ in different regions.

b) *If G is infinite*: A vertex α and ∞ are not reachable from each other in G if and only if there is an open self-avoiding contour in G', whose representation leaves the point $P(\alpha)$ in the bounded region.

The main lemma is "evident" in the same sense, in which all the basic topological facts are evident (Jordan's theorem, for example). We have a proof of it in [L.2002], but do not include it here. Let us notice that it is possible to turn the non-oriented percolation model into a graph by eliminating all closed edges. The resulting graph may be disconnected and the main lemma boils down to the following "evident" statements:

a) Two vertices of a finite graph drawn in a plane are not connected with a path in this graph if and only if there is a face containing a Jordan curve separating the representations of these vertices.

b) A vertex of an infinite graph drawn in a plane is not connected with infinity (i.e. there is no infinite self-avoiding path in this graph starting at this vertex) if and only if there is a face containing a Jordan curve surrounding the representation of this vertex.

Rule for a dual pair of oriented graphs $(G,\ G')$:

Given dual edges e and e', for each direction of e the corresponding direction of e' is the direction from right to left when we go along e in the given direction. Every edge of the graph G' is open in a certain direction if and only if the dual edge of the graph G is closed in the corresponding direction. (2)

Observe that in the case of oriented graphs the symmetry of duality becomes more complicated: given a direction of an edge e' of G', the corresponding direction of e is from left to right when we go along e' in this direction.

Given an oriented graph G, we say that:

a) vertex β is *reachable* from vertex α if there is a a self-avoiding path connecting α and β, open in the direction from α to β.

b) ∞ is *reachable* from a vertex α if there is an infinite self-avoiding path which begins at α and is open in the direction away from α.

Theorem. Let $(G,\ G')$ be a dual pair of oriented graphs, satisfying the rule (2). Then:

a) *If G is finite*: A vertex β is not reachable from another vertex α in G if and only if there is a self-avoiding contour in G', open in such a direction that its representation in the plane leaves $P(\alpha)$ on the left side and $P(\beta)$ on the right side.

b) *If G is infinite*: ∞ is not reachable from a vertex α in G if and only if there is an open self-avoiding contour in G', whose representation in the plane leaves the point $P(\alpha)$ in the finite area and surrounds it in the counter-clock direction.

Let us assume that the main lemma is proved and prove the theorem.

Proof in case a). Let $(G,\ G')$ be a dual pair of oriented finite graphs. Let us call a *good path* a self-avoiding path in the graph G connecting α and β, open in the direction from α for β. A *good contour* is a self-avoiding contour in the graph G', which is open in such a direction that its representation is a closed curve that leaves the point $P(\alpha)$ on the left side and $P(\beta)$ on the right side.

In one direction. Let us suppose that there is a good path in G and there is a good contour C in G' and obtain a contradiction. Let us denote H^* and C^* the representations of H and C in the plane respectively. By our assumption, $P(\alpha)$ and $P(\beta)$ are at different sides of C^*. From Jordan theorem, C^* and H^* have at least one common point. Let Q^* be the first point of intersection between C^* and H^* when we move along H^* starting at $P(\alpha)$. So Q^* belongs to representations of two dual edges, e of G and e' of G'. The edge e' belongs to the contour C and is open in the direction of C. Therefore the ends of e are on the opposite sides of C^* and the direction of e' in the direction of C corresponds to the direction of e from the left side to the right side of C. Since $P(\alpha)$ is on the left side when we move along C in the counter-clock direction, the edge e of the path H is open in the direction from left to right of C. So both e and e' are open in dual directions, which contradicts rule (2).

In the other direction. Let us assume that there is no good path in the graph G and prove that there is a good contour in the graph G'. Let us classify vertices of G into three *types* as follows:

$$\begin{cases} 1) \text{ A vertex } v \text{ of } G \text{ is \textit{type 1} if there is an open path from } \alpha \text{ to } v. \\ 2) \text{ A vertex } v \text{ is \textit{type 2} if there is a path from } v \text{ to } \beta \text{ without vertices type 1.} \\ 3) \text{ A vertex } v \text{ is \textit{type 3} if it is neither type 1, nor type 2.} \end{cases}$$

Notice that every vertex of G has exactly one type. Given a dual pair $(G,\ G')$, every face of G' is given the same type as the type of the corresponding vertex of G.

Let us introduce a dual pair $(\overline{G},\ \overline{G}')$ of non-oriented graphs, which have the same vertices and edges as G and G', and their representations in the plane. Let any edge of the graph $\overline{G}$ be open if and only if both ends of this edge are type 2 in G or both are not type 2 in G. After that, all the edges of $\overline{G}'$ are declared open or closed according to rule (1). Since α is type 1 in G and β is type 2 in G, every path in $\overline{G}$ connecting α and β has a closed edge. Therefore β is not reachable from α in $\overline{G}$. Hence, from the finite case of main lemma, there is an open self-avoiding contour C in the graph $\overline{G}'$, whose representation in the plane separates $P(\alpha)$ and $P(\beta)$. Let us denote $e'_1, e'_2, \ldots, e'_n$ the edges of that contour. All these edges are open, so all of their duals $e_1, e_2, \ldots, e_n$, in $\overline{G}$ are closed. Therefore every edge e_i connects a vertex not type 2 with a vertex type 2 of G. Let us denote these vertices $u_1, u_2, \ldots, u_n$ and $v_1, v_2, \ldots, v_n$ respectively. Let us prove that all the vertices $u_1, u_2, \ldots, u_n$ are type 1. Suppose that a vertex u_k is not type 1. We know that u_k is connected with a vertex v_k by an edge and that v_k is type 2. There is a path from v_k to β without vertices type 1, so there is a path from u_k to β without vertices type 1, so u_k is type 2. This is a contradiction. So every u_i is type 1, therefore every e_i is closed in G in the direction from u_i to v_i because otherwise v_i would be type 1. Therefore all e'_i should be open in G' in such a direction that the faces u_i are on the left side and the faces v_i are on the right side of them. Notice that no vertex type 2 can be inside the contour C. Therefore u_i are inside of our contour C and v_i are outside of C because v_i are type 2. Thus C is a self-avoiding contour, which separates $P(\alpha)$ from $P(\beta)$ and is open in such a direction that it leaves $P(\alpha)$ on the left side and $P(\beta)$ on the right side.

Proof in case b). Let $(G,\ G')$ be a dual pair of infinite oriented graphs. A *good path* is a self-avoiding infinite path in G, which begins in α and is open in this direction. A *good contour* is a self-avoiding contour in the graph G', which is open in that direction, whose representation in the plane goes around $P(\alpha)$ in the counter-clock direction.

In one direction. Let us suppose that there is a good path H in G and a good contour C in G' and obtain a contradiction. Let us denote H^* and C^* the representations of H and C in the plane respectively. Due to the condition 4 of definition of graph drawn in a plane, there is only a finite set of vertices of G inside C^*. Hence, since H is self-avoiding and infinite, it contains a vertex β, whose representation is in the exterior of C^*. So $P(\alpha)$ and $P(\beta)$ are in different regions of $\mathbb{R}^2 \setminus C^*$. Hence, from Jordan theorem, C^* and H^* have at least one common point. Let Q^* be the first point of intersection of C^* with H^* when we go along H^* starting at $P(\alpha)$. So Q^* belongs to representations of two dual edges, one of G and the other of G'. By the rule (2), these edges cannot be open at the same time in corresponding directions. But they are: the edge of H is open from inside to outside of the contour C^* and the edge of C is open in the counter-clock direction. This makes a contradiction.

In the other direction. Let us assume that there is no good path and prove that there is a good contour. Let us classify vertices of G into three *types* as follows (this classification is similar to that in the finite case, but not exactly the same):

1) A vertex v of G is *type 1* if there is an open path from α to v.
2) A vertex v is *type 2* if there is a path from v to ∞ without vertices type 1.
3) A vertex v is *type 3* if it is neither type 1, nor type 2.

Notice that every vertex of G has exactly one type. Given a dual pair $(G,\ G')$, every face of G' is given the same type as the type of the corresponding vertex of G.

As before, let us use a dual pair $(\overline{G},\ \overline{G}')$ of non-oriented graphs, which have the same vertices and edges and the same representations in the plane as G and G'. Let any edge of the graph $\overline{G}$ be open if and only if the ends of that edge are both type 2 or both of another type in G. After that, let the edges of $\overline{G}'$ be open or closed according to the rule (1). Since there is no good path in G, the set of vertices type 1 in G is finite. Therefore every self-avoiding path in G beginning at α has a finite number of vertices type 1. So for each infinite self-avoinding path in G, starting at α, there is a last vertex ω type 1, all the subsequent vertices being type 2. According to the definition of $\overline{G}$, the edge of $\overline{G}$, which connects ω with its successor in this path, is closed. So ∞ is not reachable from α in $\overline{G}$. Therefore, due to the infinite case of main lemma, in the dual graph $\overline{G}'$ there is an open self-avoiding contour C, whose representation surrounds $P(\alpha)$. Let us denote $e'_1,\ e'_2,\ \ldots,\ e'_n$ the edges of that contour. Since all these edges are open, all their duals $e_1, e_2, \ldots, e_n$ in $\overline{G}$ are closed. According to the definition of the graph $\overline{G}$, each edge e_i of $\overline{G}$ connects a vertex not type 2 with a vertex type 2 in G. Let us denote these vertices $u_1,\ u_2,\ \ldots,\ u_n$ and $v_1,\ v_2,\ \ldots,\ v_n$ respectively. Like in the finite case, we can prove that each u_i is type 1. So e_i has to be closed in G in the direction from u_i to v_i, because otherwise v_i would have type 1. So all e'_i should be open in G' in such a direction that the faces u_i are on the left side and the faces v_i are on the right side of them. As before, no vertex of type 2 can be inside the contour C^*. Therefore, every u_i is inside of C^* and v_i is outside of it because v_i is type 2. So C is a self-avoiding contour in G', whose representation surrounds $P(\alpha)$, which is open in a direction, which leaves $P(\alpha)$ on the left side, that is in the counter-clock direction. ■

Our proofs are over. Let us present examples of use of our main lemma and theorem to obtain upper estimations of critical values in percolation. As usual, we denote $\mathbb{L}^2$ the non-oriented graph in which the set of vertices is $\mathbb{Z}^2$ and two vertices are connected with an edge if the Euclidean distance between them is 1. An example of application of an assertion similar to the item b) of main lemma to the case when $G = \mathbb{L}^2$ is on pp. 15-19 of Grimmett's book [G.1999]. Also there is an example on pp. 6-13 of [T.2001]. An example of application of an assertion similar to the item a) of our theorem to a special class of cellular automata (a discrete analog of contact processes) is in [T.1968]. An example of application of the item b) of our theorem is in [T.2001] on pp. 16-20. It remains to present an example of use of item a) of the main lemma. Let us consider a finite rectangular part of the graph $\mathbb{L}^2$ with the width W and height H. It is shown on the figure 1 with $W = 8$ and $H = 6$.

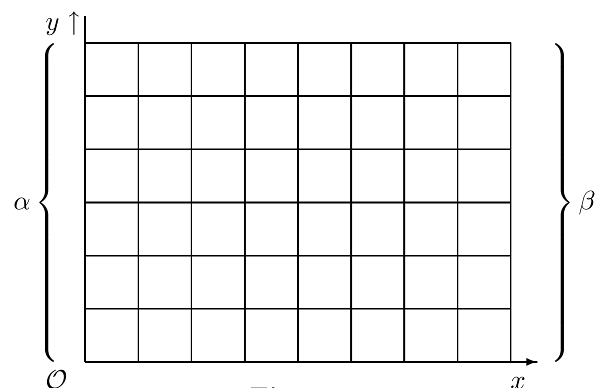

Figure 1.

Notice that we are using a rectangular coordinate system Oxy with the origin at the left-lower corner of the picture. Let us suppose that all the leftmost vertices with $x = 0$ are identified with α (say, one of them is α and the vertical edges connecting them are always open). All the rightmost vertices with $x = W$ are identified with β (say, one of them is β and the vertical edges connecting them are always open). All the other edges are open with probability ϵ and closed with probability $1 - \epsilon$ independently from each other. Let us estimate the probability that α and β are reachable from each other. It is easy to prove that, if $H \leq \text{const} \cdot W$ and ϵ is small enough, say $\epsilon < 1/3$, this probability tends to zero when $W \to \infty$. To estimate the same probability for large ϵ is not so easy; we shall do it using the item a) of main lemma. According to it, α and β are not reachable from each other if and only if there is a contour in the dual picture, separating α from β. This contour must contain the vertice dual of the unbounded face of the original graph. Cutting this contour at this vertex, we obtain a path connecting the upper and lower sides of the dual graph, which is similar to the original one. Let us denote $\delta = 1 - \epsilon$. Thus the probability that α and β are not reachable from each other does not exceed

$$W \cdot \sum_{k=H}^{\infty} (3\delta)^k \;=\; W \cdot \frac{(3\delta)^H}{1 - 3\delta}.$$

If $W \leq$ const $\cdot\, H$ and $\epsilon > 2/3$, this quantity tends to zero when $H \to \infty$. Thus the probability that α and β are reachable from each other behaves differently for small vs. large values of ϵ when $W \asymp H \to \infty$.

Now let us discuss some previous publications. Whitney [W.1932, W.1933] proved some statements about dual graphs. His theorem 4 on page 77 of [W.1933] is similar to the finite case of our main lemma, but it gives only a criterion whether a graph is connected or not and is not concerned whether two particular vertices are connected.

Our main lemma is often believed to be "well-known", but we cannot refer the reader to a source, where it is proved or even stated beyond a few special cases. We believe that this is not acceptable because mathematics is a general rigorous science and all important mathematical facts should be stated and proved in a general form.

Hammersley had some idea of duality when he wrote [H.1959], which provided the first upper estimation of a critical value in percolation. However, no general definition of duality was stated and no rule similar to our rule (1) was declared there. The most fundamental book on percolation is Grimmett's [G.1999]. On pages 16-18 and 283-287 he explains this kind of duality and its application to obtain an upper estimation of a critical value. However, he does all this only for the graph $\mathbb{L}^2$ and without a proof. Instead of going into topological details, [G.1999] refers the reader to the page 386 of [K.1982], the first page of Appendix "Some results for planar graphs", without specifying, which result of that Appendix is to be used. However, [K.1982] deals mostly with periodic, therefore, infinite case and mostly with site percolation and matching pairs.

Finally, about duality of oriented graphs. According to our knowledge, our theorem and the very definition of duality of oriented graphs have never been mentioned in print except [T.1968, T.2001], in both cases without a proof, and [L.2002], unpublished.

Although our theorem allows to obtain upper estimations in oriented percolation models, these estimations can be obtained by other means also, albeit not so easily. Let us present some examples. Shnirman [S.1968] proved non-ergodicity of Stavskaya process (a discrete-time contact process) without using duality. He considered the sequence of distributions for all natural t and proved by induction that all of them satisfy a certain infinite system of inequalities. His method was so complicated that it almost never was used again except [T.1972]. Durrett [D.1984] obtained an upper estimation of critical value in a few oriented percolation models, as a corollary of his study of a certain contact process. Liggett [L.1995] obtained upper estimations of critical values in some percolation models as by-products of his study of a certain growth model. For our example Liggett proves that the critical value does not exceed 2/3. Durrett's and Liggett's numerical estimations are better than that which we obtained in [T.2001], which was an educational text designed just to illustrate some ideas. However, the duality approach seems more general and can be amplified to obtain better estimations.

Acknowledgements. P. de Lima did this work at the department of statistics of the Federal University of Pernambuco as a part of his master program supported by CAPES. A. Toom's research was supported by CNPq, grant 300991/98-3.

Received: February 2008. Revised: June 2008.

References

[D.1984] R. Durrett, *Oriented Percolation in Two Dimensions*, *Annals of Probability*, **12** (1984), n. 4, pp. 999–1040.

[G.1999] J. Grimmett, *Percolation*, *Springer,* 1999.

[H.1959] J.M. Hammersley, *Bornes supérieures de la probabilité critique das un processus de filtration, Le calcul des probabilités et ses application*, Colloques Internationaux du Centre National de la Recherche Scientifique, LXXXVII (1959), Paris, pp. 17–37. (In French.)

[K.1982] H. Kesten, *Percolation theory for mathematicians*, *Birkhäuser*, 1982.

[L.1995] T.M. Liggett, *Survival of discrete time growth models, with applications to oriented percolation*, Annals of Applied Probability, **5** (1995), pp. 613–636.

[L.2002] P.F. de Lima, *Teoremas de Dualidade Usados na Percolação*, Master's Thesis, 2002. Unpublished. Deposited at the department of statistics of UFPE, Brazil. (In Portuguese.)

[S.1968] M. Shnirman, *On the problem of ergodicity of a Markov chain with infinite set of states*, Problemy Kibernetiki, **20** (1968), pp. 115–124. (In Russian.)

[T.1968] A. Toom, *A Family of Uniform Nets of Formal Neurons*, Soviet Math. Doklady, **9** (1968), pp. 1338–1341.

[T.1972] A. Toom, *On Invariant Measures in Non-Ergodic Random Media*, Probabilistical Methods of Investigation, issue 41. Ed. by A. Kolmogorov. Moscow University Press, 1972, pp. 43–51. (In Russian.)

[T.2001] A. Toom, *Contornos, conjuntos convexos e autômatas celulares*, Curso ministrado no 23º Colóquio Brasileiro de Matemática. IMPA, Rio de Janeiro, RJ, Brazil, 2001. (In Portugues.)

[W.1932] H. Whitney, *Non-Separable and Planar Graphs*, Trans. Amer. Math. Soc., **34** (1932) Cambridge, USA, pp. 339–362.

[W.1933] H. Whitney, *Planar Graphs*, Fund. Math., **21** (1933), Cambridge, USA, 1933, pp. 73–84.

CUBO A Mathematical Journal
Vol.10, Nº03, (103–114). October 2008

The Flip Crossed Products of the C^*-Algebras by Almost Commuting Isometries

Takahiro Sudo
Department of Mathematical Sciences,
Faculty of Science, University of the Ryukyus,
Nishihara, Okinawa 903-0213, Japan
email: sudo@math.u-ryukyu.ac.jp

ABSTRACT

We study the flip crossed products of the C^*-algebras by almost commuting isometries and obtain some results on their structure, K-theory, and continuity.

RESUMEN

Estudiamos el produto flip crossed de una C^*-algebra mediante isometrias casi commutando y obtenemos algunos resultados sobre su estructura, K-teoria, y continuidad.

Key words and phrases: *C^*-algebra, Continuous field, K-theory, Isometry.*

Math. Subj. Class.: *46L05, 46L80.*

Introduction

Recall that the soft torus A_ε of Exel [3] (for any $\varepsilon \in [0,2]$ the closed interval) is defined to be the universal C^*-algebra generated by almost commuting two unitaries $u_{\varepsilon,1}$ and $u_{\varepsilon,2}$ in the sense that $\|u_{\varepsilon,2}u_{\varepsilon,1} - u_{\varepsilon,1}u_{\varepsilon,2}\| \leq \varepsilon$. Its K-theory is computed in [3] by showing that it can be represented as a crossed product by $\mathbb{Z}$ and applying the Pimsner-Voiculescu six-term exact sequense for the crossed product. It is shown by Exel [4] that there exists a continuous field of C^*-algebras on $[0,2]$ with fibers the soft tori varying continuously. Furthermore, K-theory and continuity of the crossed products of A_ε by the flip (a $\mathbb{Z}_2$-action) are considered by Elliott, Exel and Loring [2].

On the other hand, we [8] began to study continuous fields of C^*-algebras by almost commuting isometries and obtained some similar results (but different in some senses) on their structure, K-theory and continuity as those by Exel. In this paper we consider those properties for the flip crossed products of the C^*-algebras generated by almost commuting isometries.

Refer to [1], [5], and [9] for some basics in C^*-algebras and K-theory.

1 The flip crossed products by isometries

The Toeplitz algebra is defined to be the universal C^*-algebra generated by a (non-unitary) isometry, and it is denoted by $\mathfrak{F}$, which is also the semigroup C^*-algebra $C^*(\mathbb{N})$ of the semigroup $\mathbb{N}$ of natural numbers. The C^*-algebra $C(\mathbb{T})$ of all continuous functions on the 1-torus $\mathbb{T}$ is the universal C^*-algebra generated by a unitary, which is also the group C^*-algebra $C^*(\mathbb{Z})$ of the group $\mathbb{Z}$ of integers. There is a canonical quotient map from $\mathfrak{F}$ to $C(\mathbb{T})$ by universality, whose kernel is isomorphic to the C^*-algebra $\mathbb{K}$ of all compact operators on a separable infinite dimensional Hilbert space (cf. [5]).

Definition 1.1 For $\varepsilon \in [0,2]$, the soft Toeplitz tensor product denoted by $\mathfrak{F} \otimes_\varepsilon \mathfrak{F}$ is defined to be the universal C^*-algebra generated by two isometries $s_{\varepsilon,1}$, $s_{\varepsilon,2}$ such that $\|s_{\varepsilon,2}s_{\varepsilon,1} - s_{\varepsilon,1}s_{\varepsilon,2}\| \leq \varepsilon$ (ε-commuting). Let $\pi : \mathfrak{F} \otimes_\varepsilon \mathfrak{F} \to A_\varepsilon$ be the canonical onto $*$-homomorphism sending the isometry generators to the unitary generators.

Remark. Refer to [8], in which super-softness is further defined and assumed, but it should be unnecessary from the universality argument (as given below). Instead, in fact, another norm estimate of the form $\|s_{\varepsilon,2}s^*_{\varepsilon,1} - s^*_{\varepsilon,1}s_{\varepsilon,2}\| \leq \varepsilon$ (ε-$*$-commuting) may be required, but we omit such an estimate in what follows. If not assuming the estimate, $\mathfrak{F} \otimes_\varepsilon \mathfrak{F}$ should be replaced with $C^*(\mathbb{N}^2)_\varepsilon$, where $C^*(\mathbb{N}^2)$ is the semigroup C^*-algebra of $\mathbb{N}^2$ (in what follows).

Definition 1.2 The flip on $\mathfrak{F} \otimes_\varepsilon \mathfrak{F}$ is the (non-unital) endomorphism σ defined by $\sigma(s_{\varepsilon,j}) = s^*_{\varepsilon,j}$ for $j = 1,2$. Since σ^2 is the identity on $\mathfrak{F} \otimes_\varepsilon \mathfrak{F}$, we denote by $(\mathfrak{F} \otimes_\varepsilon \mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2$ the crossed product of $\mathfrak{F} \otimes_\varepsilon \mathfrak{F}$ by the action σ of the order 2 cyclic group $\mathbb{Z}_2$, i.e., a flip crossed product.

Definition 1.3 For $\varepsilon \in [0,2]$, we define E_ε to be the universal C^*-algebra generated by an isometry t_1 and the elements $t_{n+1} = u^n t_1 (u^*)^n$ for $n \in \mathbb{N}$, where u is an isometry, such that $\|ut_1 - t_1 u\| \leq \varepsilon$. Let α_ε be the endomorphism of E_ε defined by $\alpha_\varepsilon(t_n) = t_{n+1} = ut_n u^*$ for $n \in \mathbb{N}$. Let $E_\varepsilon \rtimes_{\alpha_\varepsilon} \mathbb{N}$ be the semigroup crossed product of E_ε by the action α_ε of the additive semigroup $\mathbb{N}$ of natural numbers.

Remark. Note that $\mathfrak{F} \otimes_2 \mathfrak{F}$ (or $C^*(\mathbb{N}^2)_2$) is isomorphic to the unital full free product $\mathfrak{F} *_{\mathbb{C}} \mathfrak{F}$, which is also isomorphic to the full semigroup C^*-algebra $C^*(\mathbb{N} * \mathbb{N})$ of the free semigroup $\mathbb{N} * \mathbb{N}$. As in the above remark, another estimate $\|ut_1^* - t_1^* u\| \leq \varepsilon$ may be required accordingly.

It is shown in [8] that $\mathfrak{F} \otimes_\varepsilon \mathfrak{F} \cong E_\varepsilon \rtimes_{\alpha_\varepsilon} \mathbb{N}$, where the map φ from $\mathfrak{F} \otimes_\varepsilon \mathfrak{F}$ to $E_\varepsilon \rtimes_{\alpha_\varepsilon} \mathbb{N}$ is defined by $\varphi(s_{\varepsilon,1}) = t_1$ and $\varphi(s_{\varepsilon,2}) = u$, and its inverse ψ is given by $\psi(t_{n+1}) = s_{\varepsilon,2}^n s_{\varepsilon,1} (s_{\varepsilon,2}^*)^n$ for $n \in \mathbb{N}$ and $n = 0$ and $\psi(u) = s_{\varepsilon,2}$.

Proposition 1.4 *For $\varepsilon \in [0,2]$, we have the following isomorphism:*

$$(\mathfrak{F} \otimes_\varepsilon \mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2 \cong E_\varepsilon \rtimes_{\alpha_\varepsilon * \beta} (\mathbb{N} * \mathbb{Z}_2),$$

where $\mathbb{N} * \mathbb{Z}_2$ is the free product of $\mathbb{N}$ and $\mathbb{Z}_2$, and the action β on E_ε is given by $\beta(t_n) = t_n^*$ for $n \in \mathbb{N}$.

Proof. The crossed product $(\mathfrak{F} \otimes_\varepsilon \mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2$ is the universal C^*-algebra generated by isometries $s_{\varepsilon,1}$, $s_{\varepsilon,2}$ and a unitary ρ such that $\|s_{\varepsilon,2} s_{\varepsilon,1} - s_{\varepsilon,1} s_{\varepsilon,2}\| \leq \varepsilon$ and $\rho s_{\varepsilon,j} \rho^* = s_{\varepsilon,j}$ $(j = 1,2)$ with $\rho^2 = 1$, while $E_\varepsilon \rtimes_{\alpha_\varepsilon * \beta} (\mathbb{N} * \mathbb{Z}_2)$ is the C^*-algebra generated by isometries t_1, u and a unitary v such that $\|ut_1 - t_1 u\| \leq \varepsilon$ and $t_{n+1} = ut_n u^* = u^n t_1 (u^*)^n$ for $n \in \mathbb{N}$, and $vt_1 v^* = t_1^*$ and $vuv^* = u^*$ with $v^2 = 1$. The isomorphism between them is given by sending $s_{\varepsilon,1}, s_{\varepsilon,2}$, and ρ to t_1, u, and v respectively (cf. [2]). □

Theorem 1.5 *For $0 \leq \varepsilon < 2$, we obtain the K-theory isomorphisms:*

$$K_0((\mathfrak{F} \otimes_\varepsilon \mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2) \cong \mathbb{Z}^9, \quad K_1((\mathfrak{F} \otimes_\varepsilon \mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2) \cong 0.$$

Moreover, $K_j((\mathfrak{F} \otimes_\varepsilon \mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2) \cong K_j((\mathfrak{F} \otimes \mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2)$ for $j = 0, 1$.

Proof. Since $\mathfrak{F} \otimes_\varepsilon \mathfrak{F} \cong E_\varepsilon \rtimes_{\alpha_\varepsilon} \mathbb{N}$ and α_ε is a corner endomorphism on E_ε, note that $E_\varepsilon \rtimes_{\alpha_\varepsilon} \mathbb{N}$ is isomorphic to a corner of $(E_\varepsilon \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z}$, i.e., $p((E_\varepsilon \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z})p$ for a certain projection p, where $\rho_\varepsilon^\wedge$ is the dual action of the circle action on $E_\varepsilon \rtimes_{\alpha_\varepsilon} \mathbb{N}$ and id is the identity action on $\mathbb{K}$ (this is a variation of [6], and see also [7]). Hence, $(E_\varepsilon \rtimes_{\alpha_\varepsilon} \mathbb{N}) \rtimes_\sigma \mathbb{Z}_2$ is isomorphic to

$p((E_\varepsilon \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z})p \rtimes_\sigma \mathbb{Z}_2$. Therefore,

$$\begin{aligned} K_j((E_\varepsilon \rtimes_{\alpha_\varepsilon} \mathbb{N}) \rtimes_\sigma \mathbb{Z}_2) &\cong K_j(p((E_\varepsilon \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z})p \rtimes_\sigma \mathbb{Z}_2) \\ &\cong K_j^{\mathbb{Z}_2}(p((E_\varepsilon \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z})p) \\ &\cong K_j^{\mathbb{Z}_2}(p((E_\varepsilon \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z})p \otimes \mathbb{K}) \\ &\cong K_j^{\mathbb{Z}_2}(((E_\varepsilon \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z}) \otimes \mathbb{K}) \\ &\cong K_j((E_\varepsilon \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z} \rtimes \mathbb{Z}_2), \end{aligned}$$

where $K_j^{\mathbb{Z}_2}(\cdot)$ is the equivariant K-theory, and note that $p((E_\varepsilon \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z})p$ is stably isomorphic to $(E_\varepsilon \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z}$, and

$$\begin{aligned} (E_\varepsilon \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z} \rtimes \mathbb{Z}_2 &\cong (E_\varepsilon \otimes \mathbb{K}) \rtimes_{\sigma'_\varepsilon * \sigma \otimes \mathrm{id}} (\mathbb{Z}_2 * \mathbb{Z}_2) \\ &\cong (E_\varepsilon \rtimes_{\sigma'_\varepsilon * \sigma} (\mathbb{Z}_2 * \mathbb{Z}_2)) \otimes \mathbb{K} \end{aligned}$$

since $\mathbb{Z} \rtimes \mathbb{Z}_2 \cong \mathbb{Z}_2 * \mathbb{Z}_2$, where $\sigma'_\varepsilon(1) = \rho_\varepsilon^\wedge(1)\sigma(1)$ (cf. [2]). Set $F_\varepsilon = E_\varepsilon \rtimes_{\sigma'_\varepsilon * \sigma} (\mathbb{Z}_2 * \mathbb{Z}_2)$. There exists the following six-term exact sequence (A) (cf. [2]):

$$\begin{array}{ccccc} K_0(E_\varepsilon) & \longrightarrow & K_0(E_\varepsilon \rtimes_{\sigma'_\varepsilon} \mathbb{Z}_2) \oplus K_0(E_\varepsilon \rtimes_\sigma \mathbb{Z}_2) & \longrightarrow & K_0(F_\varepsilon) \\ \uparrow & & & & \downarrow \\ K_1(F_\varepsilon) & \longleftarrow & K_1(E_\varepsilon \rtimes_{\sigma'_\varepsilon} \mathbb{Z}_2) \oplus K_1(E_\varepsilon \rtimes_\sigma \mathbb{Z}_2) & \longleftarrow & K_1(E_\varepsilon). \end{array}$$

Consider the following exact sequence: $0 \to \mathfrak{I}_\varepsilon \to E_\varepsilon \to \pi(E_\varepsilon) = B'_\varepsilon \to 0$, where π is the canonical quotient map from E_ε to the quotient $\pi(E_\varepsilon) = B'_\varepsilon$, where B'_ε is the universal C^*-algebra generated by unitaries $u_{n+1} = w^n v(w^*)^n$ for $n \in \mathbb{N}$ and $n = 0$, where $\pi(t_{n+1}) = \pi(u)^n \pi(t_1)\pi(u^*)^n = u_{n+1}$ with $v = \pi(t_1)$ and $w = \pi(u)$. As shown in [8], K-theory groups of $\mathfrak{I}_\varepsilon$ are the same as those of $\mathbb{K}$. Since this quotient is invariant under the action $\beta = \sigma'_\varepsilon$ or σ, we have the following exact sequence:

$$(B): \quad 0 \to \mathfrak{I}_\varepsilon \rtimes_\beta \mathbb{Z}_2 \to E_\varepsilon \rtimes_\beta \mathbb{Z}_2 \to \pi(E_\varepsilon) \rtimes_\beta \mathbb{Z}_2 \to 0$$

and $\mathfrak{I}_\varepsilon \rtimes_\beta \mathbb{Z}_2 \cong \mathfrak{I}_\varepsilon \otimes C^*(\mathbb{Z}_2)$ and the group C^*-algebra $C^*(\mathbb{Z}_2)$ is isomorphic to $\mathbb{C}^2$ via the Fourier transform.

As shown in [2], it is deduced that $\pi(E_\varepsilon) \rtimes_\beta \mathbb{Z}_2$ is homotopy equivalent to the crossed product $C(\mathbb{T}) \rtimes_{\beta'} \mathbb{Z}_2$, where $\beta'(z) = z^{-1}$ for $z \in \mathbb{T}$. It follows that $K_j(\pi(E_\varepsilon) \rtimes_\beta \mathbb{Z}_2)$ is isomorphic to $K_j(C(\mathbb{T}) \rtimes_{\beta'} \mathbb{Z}_2)$. Since the points $\{\pm 1\}$ in $\mathbb{T}$ is fixed under the action β', we have

$$0 \to C_0(\mathbb{T} \setminus \{\pm 1\}) \rtimes_{\beta'} \mathbb{Z}_2 \to C(\mathbb{T}) \rtimes_{\beta'} \mathbb{Z}_2 \to \oplus^2 C^*(\mathbb{Z}_2) \to 0,$$

where $C_0(\mathbb{T} \setminus \{\pm 1\})$ is the C^*-algebra of all continuous functions on $\mathbb{T} \setminus \{\pm 1\}$ vanishing at infinity, and $C_0(\mathbb{T} \setminus \{\pm 1\}) \rtimes_{\beta'} \mathbb{Z}_2 \cong C_0(\mathbb{R}) \otimes (\mathbb{C}^2 \rtimes_{\beta'} \mathbb{Z}_2) \cong C_0(\mathbb{R}) \otimes M_2(\mathbb{C})$ and $C^*(\mathbb{Z}_2) \cong \mathbb{C}^2$. Hence the following six-term exact sequence is obtained:

$$\begin{array}{ccccc} 0 & \longrightarrow & K_0(C(\mathbb{T}) \rtimes_{\beta'} \mathbb{Z}_2) & \longrightarrow & \mathbb{Z}^4 \\ \uparrow & & & & \downarrow \\ 0 & \longleftarrow & K_1(C(\mathbb{T}) \rtimes_{\beta'} \mathbb{Z}_2) & \longleftarrow & \mathbb{Z}, \end{array}$$

where $K_j(C_0(\mathbb{R}) \otimes M_2(\mathbb{C})) \cong K_{j+1}(\mathbb{C})$ (mod 2) and $K_j(\oplus^2\mathbb{C}^2) \cong \oplus^4 K_j(\mathbb{C})$. It follows that $K_0(C(\mathbb{T}) \rtimes_{\beta'} \mathbb{Z}_2) \cong \mathbb{Z}^3$ and $K_1(C(\mathbb{T}) \rtimes_{\beta'} \mathbb{Z}_2) \cong 0$ (cf. [2]).

Therefore, for the above exact sequence (B), we obtain the diagram:

$$\begin{array}{ccccc} \mathbb{Z}^2 & \longrightarrow & K_0(E_\varepsilon \rtimes_\beta \mathbb{Z}_2) & \longrightarrow & \mathbb{Z}^3 \\ \uparrow & & & & \downarrow \\ 0 & \longleftarrow & K_1(E_\varepsilon \rtimes_\beta \mathbb{Z}_2) & \longleftarrow & 0, \end{array}$$

where $K_j(\mathbb{K} \otimes C^*(\mathbb{Z}_2)) \cong K_j(\mathbb{C}^2)$. Hence we obtain $K_0(E_\varepsilon \rtimes_\beta \mathbb{Z}_2) \cong \mathbb{Z}^5$ and $K_1(E_\varepsilon \rtimes_\beta \mathbb{Z}_2) \cong 0$. This implies that the diagram (A) is

$$\begin{array}{ccccc} \mathbb{Z} & \longrightarrow & \mathbb{Z}^5 \oplus \mathbb{Z}^5 & \longrightarrow & K_0(F_\varepsilon) \\ \uparrow & & & & \downarrow \\ K_1(F_0) & \longleftarrow & 0 \oplus 0 & \longleftarrow & 0 \end{array}$$

where it is shown in [8] that $K_0(E_\varepsilon) \cong \mathbb{Z}$ and $K_1(E_\varepsilon) \cong 0$. It follows that $K_0(F_\varepsilon) \cong \mathbb{Z}^9$ and $K_1(F_\varepsilon) \cong 0$. It follows from this and the first part shown above that $K_0((\mathfrak{F} \otimes_\varepsilon \mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2) \cong \mathbb{Z}^9$ and $K_1((\mathfrak{F} \otimes_\varepsilon \mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2) \cong 0$.

The second claim follows from the case $\varepsilon = 0$ and the same argument as above. Note that $\mathfrak{F} \otimes \mathfrak{F} \cong \mathfrak{F} \rtimes_{\mathrm{id}} \mathbb{N}$, where id is the trivial action. □

Corollary 1.6 *For $0 \leq \varepsilon < 2$, the natural onto $*$-homomorphism $\varphi_{\varepsilon,0}$ from $(\mathfrak{F} \otimes_\varepsilon \mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2$ to $(\mathfrak{F} \otimes \mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2$ sending $s_{\varepsilon,j}$ to $s_{0,j}$ $(j = 1, 2)$ induces the isomorphism between their K-groups.*

Proposition 1.7 *There exists a continuous field of C^*-algebras on the closed interval $[0, 2]$ such that its fibers are $(\mathfrak{F} \otimes_\varepsilon \mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2$ for $\varepsilon \in [0, 2]$, and for any $a \in (\mathfrak{F} \otimes_2 \mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2$, the sections $[0, 2] \ni \varepsilon \mapsto \varphi_\varepsilon(a) \in (\mathfrak{F} \otimes_\varepsilon \mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2$ are continuous, where $\varphi_\varepsilon : (\mathfrak{F} \otimes_2 \mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2 \to (\mathfrak{F} \otimes_\varepsilon \mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2$ is the natural onto $*$-homomorphism sending $s_{2,j}$ to $s_{\varepsilon,j}$ $(j = 0, 1)$.*

Proof. As shown before, $(\mathfrak{F} \otimes_\varepsilon \mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2 \cong (E_\varepsilon \rtimes_{\alpha_\varepsilon} \mathbb{N}) \rtimes_\sigma \mathbb{Z}_2$. Furthermore, this is isomorphic to $p((E_\varepsilon \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z})p \rtimes_\sigma \mathbb{Z}_2$. Hence it follows that

$$\begin{aligned} ((E_\varepsilon \rtimes_{\alpha_\varepsilon} \mathbb{N}) \rtimes_\sigma \mathbb{Z}_2) \otimes \mathbb{K} &\cong (p((E_\varepsilon \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z})p \rtimes_\sigma \mathbb{Z}_2) \otimes \mathbb{K} \\ &\cong (p((E_\varepsilon \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z})p \otimes \mathbb{K}) \rtimes_{\sigma \otimes \mathrm{id}} \mathbb{Z}_2 \\ &\cong (((E_\varepsilon \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z}) \otimes \mathbb{K}) \rtimes_{\sigma \otimes \mathrm{id}} \mathbb{Z}_2 \\ &\cong ((E_\varepsilon \otimes \mathbb{K} \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id} \otimes \mathrm{id}} \mathbb{Z}) \rtimes_{\sigma \otimes \mathrm{id}} \mathbb{Z}_2 \\ &\cong ((E_\varepsilon \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z}) \rtimes_\sigma \mathbb{Z}_2 \\ &\cong (E_\varepsilon \otimes \mathbb{K}) \rtimes_{\sigma'_\varepsilon * \sigma \otimes \mathrm{id}} (\mathbb{Z}_2 * \mathbb{Z}_2). \end{aligned}$$

It is deduced from [2] that there exists a continuous field of C^*-algebras on $[0, 2]$ such that its fibers are $(E_\varepsilon \otimes \mathbb{K}) \rtimes_{\sigma'_\varepsilon * \sigma \otimes \mathrm{id}} (\mathbb{Z}_2 * \mathbb{Z}_2)$ for $\varepsilon \in [0, 2]$, and for any $b \in (E_2 \otimes \mathbb{K}) \rtimes_{\sigma'_2 * \sigma \otimes \mathrm{id}} (\mathbb{Z}_2 * \mathbb{Z}_2)$, the sections $[0,2] \ni \varepsilon \mapsto \psi_\varepsilon(b) \in (E_\varepsilon \otimes \mathbb{K}) \rtimes_{\sigma'_\varepsilon * \sigma \otimes \mathrm{id}} (\mathbb{Z}_2 * \mathbb{Z}_2)$ are continuous, where ψ_ε is the unique onto $*$-homomorphism from $(E_2 \otimes \mathbb{K}) \rtimes_{\sigma'_2 * \sigma \otimes \mathrm{id}} (\mathbb{Z}_2 * \mathbb{Z}_2)$ to $(E_\varepsilon \otimes \mathbb{K}) \rtimes_{\sigma'_\varepsilon * \sigma \otimes \mathrm{id}} (\mathbb{Z}_2 * \mathbb{Z}_2)$. Cutting down this continuous field by cutting down the fibers from $((E_\varepsilon \rtimes_{\alpha_\varepsilon} \mathbb{N}) \rtimes_\sigma \mathbb{Z}_2) \otimes \mathbb{K}$ to $(E_\varepsilon \rtimes_{\alpha_\varepsilon} \mathbb{N}) \rtimes_\sigma \mathbb{Z}_2$ by minimal projections, we obtain the desired continuous field. □

2 The flip crossed products by n isometries

The n-fold tensor product $\otimes^n \mathfrak{F}$ of $\mathfrak{F}$ is the universal C^*-algebra generated by mutually commuting and $*$-commuting n isometries, while the universal C^*-algebra generated by mutually commuting n isometries is just the semigroup C^*-algebra $C^*(\mathbb{N}^n)$ of the semigroup $\mathbb{N}^n$. The C^*-algebra $C(\mathbb{T}^n)$ of all continuous functions on the n-torus $\mathbb{T}^n$ is the universal C^*-algebra generated by mutually commuting n unitaries, which is also the group C^*-algebra $C^*(\mathbb{Z}^n)$ of the group $\mathbb{Z}^n$. There is a canonical quotient map from $\otimes^n \mathfrak{F}$ to $C(\mathbb{T}^n) \cong \otimes^n C(\mathbb{T})$ by universality,

Definition 2.1 For $\varepsilon \in [0,2]$, the soft Toeplitz n-tensor product denoted by $\otimes^n_\varepsilon \mathfrak{F}$ is defined to be the universal C^*-algebra generated by n isometries $s_{\varepsilon,j}$ $(1 \le j \le n)$ such that $\|s_{\varepsilon,k}s_{\varepsilon,j} - s_{\varepsilon,j}s_{\varepsilon,k}\| \le \varepsilon$ $(1 \le j,k \le n)$.

Remark. Note that, in fact, the norm estimates of the form $\|s_{\varepsilon,k}s^*_{\varepsilon,j} - s^*_{\varepsilon,j}s_{\varepsilon,k}\| \le \varepsilon$ may be further required (and in what follows). If not assuming these estimates, $\otimes^n_\varepsilon \mathfrak{F}$ should be replaced with $C^*(\mathbb{N}^n)_\varepsilon$ in the same sense (and in what follows).

Definition 2.2 The flip on $\otimes^n_\varepsilon \mathfrak{F}$ is the (non-unital) endomorphism σ defined by $\sigma(s_{\varepsilon,j}) = s^*_{\varepsilon,j}$ for $1 \le j \le n$. Since σ^2 is the identity on $\otimes^n_\varepsilon \mathfrak{F}$, we denote by $(\otimes^n_\varepsilon \mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2$ the crossed product of $\otimes^n_\varepsilon \mathfrak{F}$ by the action σ of $\mathbb{Z}_2$.

Definition 2.3 For $\varepsilon \in [0,2]$, we define E^m_ε to be the universal C^*-algebra generated by n isometries $t^{(j)}_1$ $(1 \le j \le m)$ and the partial isometries $t^{(j)}_{n+1} = u^n t^{(j)}_1 (u^*)^n$ for $n \in \mathbb{N}$, where u is an isometry such that $\|ut^{(j)}_1 - t^{(j)}_1 u\| \le \varepsilon$ and $\|t^{(k)}_1 t^{(j)}_1 - t^{(j)}_1 t^{(k)}_1\| \le \varepsilon$ $(1 \le j,k \le m)$. Let α_ε be the endomorphism of E^m_ε defined by $\alpha_\varepsilon(t^{(j)}_n) = t^{(j)}_{n+1} = ut^{(j)}_n u^*$ for $n \in \mathbb{N}$. Let $E^m_\varepsilon \rtimes_{\alpha_\varepsilon} \mathbb{N}$ be the semigroup crossed product of E^m_ε by the action α_ε of $\mathbb{N}$.

Remark. Note that $\otimes^n_2 \mathfrak{F}$ (or $C^*(\mathbb{N}^n)_2$) is isomorphic to the unital full free product $*^n_{\mathbb{C}} \mathfrak{F}$, which is also isomorphic to the full semigroup C^*-algebra $C^*(*^n \mathbb{N})$ of the free semigroup $*^n \mathbb{N}$. As in the above remark, the additional estimates $\|u(t^{(j)}_1)^* - (t^{(j)}_1)^* u\| \le \varepsilon$ and $\|t^{(k)}_1 (t^{(j)}_1)^* - (t^{(j)}_1)^* t^{(k)}_1\| \le \varepsilon$ may be required accordingly.

It is shown as in [8] that $\otimes^{m+1}_\varepsilon \mathfrak{F} \cong E^m_\varepsilon \rtimes_{\alpha_\varepsilon} \mathbb{N}$ as in the case in Section 1.

Proposition 2.4 *For* $\varepsilon \in [0,2]$*, we have*

$$(\otimes^{m+1}_\varepsilon \mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2 \cong E^m_\varepsilon \rtimes_{\alpha_\varepsilon * \beta} (\mathbb{N} * \mathbb{Z}_2),$$

where the action β *on* E_ε^m *is given by* $\beta(t_n^{(j)}) = (t_n^{(j)})^*$ *for* $n \in \mathbb{N}$ *and* $1 \le j \le m$.

Proof. This is shown as in the proof of Proposition 1.4 similarly. □

Theorem 2.5 *For* $0 \le \varepsilon < 2$, *we obtain* (*inductively*)

$$K_0((\otimes_\varepsilon^{m+1}\mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2) \cong \mathbb{Z}^{2^{m+2}+3}, \quad K_1((\otimes_\varepsilon^{m+1}\mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2) \cong 0.$$

Moreover, $K_j((\otimes_\varepsilon^{m+1}\mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2) \cong K_j((\otimes^{m+1}\mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2)$ *for* $j = 0, 1$.

Proof. Since $\otimes_\varepsilon^{m+1}\mathfrak{F} \cong E_\varepsilon^m \rtimes_{\alpha_\varepsilon} \mathbb{N}$, note that $E_\varepsilon^m \rtimes_{\alpha_\varepsilon} \mathbb{N}$ is isomorphic to a corner of $(E_\varepsilon^m \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z}$, i.e., $p((E_\varepsilon^m \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z})p$ for a certain projection p, where $\rho_\varepsilon^\wedge$ is the dual action of the circle action on $E_\varepsilon^m \rtimes_{\alpha_\varepsilon} \mathbb{N}$ and id is the identity action on $\mathbb{K}$ (this is a variation of [6], and see also [7]). Hence, $(E_\varepsilon^m \rtimes_{\alpha_\varepsilon} \mathbb{N}) \rtimes_\sigma \mathbb{Z}_2$ is isomorphic to $p((E_\varepsilon^m \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z})p \rtimes_\sigma \mathbb{Z}_2$. Therefore,

$$\begin{aligned} K_j((E_\varepsilon^m \rtimes_{\alpha_\varepsilon} \mathbb{N}) \rtimes_\sigma \mathbb{Z}_2) &\cong K_j(p((E_\varepsilon^m \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z})p \rtimes_\sigma \mathbb{Z}_2) \\ &\cong K_j^{\mathbb{Z}_2}(p((E_\varepsilon^m \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z})p) \\ &\cong K_j^{\mathbb{Z}_2}(p((E_\varepsilon^m \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z})p \otimes \mathbb{K}) \\ &\cong K_j^{\mathbb{Z}_2}(((E_\varepsilon^m \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z}) \otimes \mathbb{K}) \\ &\cong K_j((E_\varepsilon^m \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z} \rtimes \mathbb{Z}_2), \end{aligned}$$

where $p((E_\varepsilon^m \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z})p$ is stably isomorphic to $(E_\varepsilon^m \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z}$, and

$$\begin{aligned} (E_\varepsilon^m \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z} \rtimes \mathbb{Z}_2 &\cong (E_\varepsilon^m \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge * \sigma \otimes \mathrm{id}} (\mathbb{Z}_2 * \mathbb{Z}_2) \\ &\cong (E_\varepsilon^m \rtimes_{\rho_\varepsilon^\wedge * \sigma} (\mathbb{Z}_2 * \mathbb{Z}_2)) \otimes \mathbb{K} \end{aligned}$$

since $\mathbb{Z} \rtimes \mathbb{Z}_2 \cong \mathbb{Z}_2 * \mathbb{Z}_2$ (cf. [2]). Set $F_\varepsilon^m = E_\varepsilon^m \rtimes_{\rho_\varepsilon^\wedge * \sigma} (\mathbb{Z}_2 * \mathbb{Z}_2)$. There exists the following six-term exact sequence $(A)_m$ (cf. [2]):

$$\begin{array}{ccccc} K_0(E_\varepsilon^m) & \longrightarrow & K_0(E_\varepsilon^m \rtimes_{\rho_\varepsilon^\wedge} \mathbb{Z}_2) \oplus K_0(E_\varepsilon^m \rtimes_\sigma \mathbb{Z}_2) & \longrightarrow & K_0(F_\varepsilon^m) \\ \uparrow & & & & \downarrow \\ K_1(F_\varepsilon^m) & \longleftarrow & K_1(E_\varepsilon^m \rtimes_{\rho_\varepsilon^\wedge} \mathbb{Z}_2) \oplus K_1(E_\varepsilon^m \rtimes_\sigma \mathbb{Z}_2) & \longleftarrow & K_1(E_\varepsilon^m). \end{array}$$

We now have the following exact sequence:

$$0 \to \mathfrak{I}_\varepsilon^m \rtimes \mathbb{Z}_2 \to E_\varepsilon^m \rtimes \mathbb{Z}_2 \to \pi(E_\varepsilon^m) \rtimes \mathbb{Z}_2 \to 0,$$

where the map π is sending isometries of E_ε^m to unitaries with the same norm estimates by universality, and $\mathfrak{I}_\varepsilon^m$ is the kernel of π, and the action of $\mathbb{Z}_2$ is given by $\rho_\varepsilon^\wedge$ or σ. Furthermore, it follows that $\mathfrak{I}_\varepsilon^m \rtimes \mathbb{Z}_2 \cong \mathfrak{I}_\varepsilon^m \otimes C^*(\mathbb{Z}_2)$ and the K-theory of $\mathfrak{I}_\varepsilon^m$ is the same as that of $\mathbb{K}$.

It is deduced that $\pi(E_\varepsilon^m) \rtimes \mathbb{Z}_2$ is homotopy equivalent to $C(\mathbb{T}^m) \rtimes_\sigma \mathbb{Z}_2$, where $\beta(z_j) = (z_j^{-1})$ for $(z_j) \in \mathbb{T}^m$. Since the points $(\pm 1, \cdots, \pm 1) \in \mathbb{T}^m$ are fixed under α, we have

$$0 \to C_0(\mathbb{T}^m \setminus (\pm 1, \cdots, \pm 1)) \rtimes \mathbb{Z}_2 \to C(\mathbb{T}^m) \rtimes \mathbb{Z}_2 \to \oplus^{2^m} C^*(\mathbb{Z}_2) \to 0,$$

where $C_0(X)$ is the C^*-algebra of all continuous functions on a locally compact Hausdorff space X vanishing at infinity (in what follows). Set $X_{m+1} = \mathbb{T}^m \setminus (\pm 1, \cdots, \pm 1)$. By considering invariant subspaces in X_{m+1} under β, we obtain a finite composition series $\{\mathfrak{L}_j\}_{j=1}^m$ of $C_0(X_{m+1}) \rtimes \mathbb{Z}_2$ such that $\mathfrak{L}_0 = \{0\}$, $\mathfrak{L}_j = C_0(X_j) \times \mathbb{Z}_2$, and

$$\mathfrak{L}_j/\mathfrak{L}_{j-1} \cong \oplus^{{}_mC_{m-j+1}} C_0((\mathbb{T} \setminus \{\pm 1\})^{m-j+1}) \rtimes \mathbb{Z}_2,$$

where ${}_mC_{m-j+1}$ mean the combinations. Furthermore,

$$C_0((\mathbb{T} \setminus \{\pm 1\})^{m-j+1}) \rtimes \mathbb{Z}_2 \cong C_0(\mathbb{R}^{m-j+1}) \otimes (C(\Pi^{m-j+1}\{\pm i\}) \rtimes \mathbb{Z}_2)$$

and $C(\Pi^{m-j+1}\{\pm i\}) \rtimes \mathbb{Z}_2 \cong \oplus^{m-j+1}(\mathbb{C}^2 \rtimes \mathbb{Z}_2) \cong \oplus^{m-j+1} M_2(\mathbb{C})$, where $\mathbb{T} \setminus \{\pm 1\}$ is homeomorphic to $i\mathbb{R} \cup (-i)\mathbb{R}$ so that the above isomorphisms are deduced from considering orbits under β in this identification. Set $C(m,j) = {}_mC_{m-j+1}(m-j+1)$. Thus, the following six-term exact sequences are obtained:

$$\begin{array}{ccccc} K_0(\mathfrak{L}_{j-1}) & \longrightarrow & K_0(\mathfrak{L}_j) & \longrightarrow & K_{m-j+1}(\oplus^{C(m,j)}\mathbb{C}) \\ \uparrow & & & & \downarrow \\ K_{m-j+2}(\oplus^{C(m,j)}\mathbb{C}) & \longleftarrow & K_1(\mathfrak{L}_j) & \longleftarrow & K_1(\mathfrak{L}_{j-1}). \end{array}$$

Now consider the case $m = 2$. Then

$$0 \to C_0(\mathbb{T}^2 \setminus (\pm 1, \pm 1)) \rtimes \mathbb{Z}_2 \to C(\mathbb{T}^2) \rtimes \mathbb{Z}_2 \to \oplus^{2^2} C^*(\mathbb{Z}_2) \to 0.$$

Furthermore, $0 \to C_0(X_1) \rtimes \mathbb{Z}_2 \to C_0(X_2) \rtimes \mathbb{Z}_2 \to C_0(X_2 \setminus X_1) \rtimes \mathbb{Z}_2 \to 0$, where $X_2 = \mathbb{T}^2 \setminus (\pm 1, \pm 1)$, $X_1 = (\mathbb{T} \setminus \{\pm 1\})^2$, and $C_0(X_2 \setminus X_1) \rtimes \mathbb{Z}_2$ is isomorphic to $\oplus^2 C_0(\mathbb{T} \setminus \{\pm 1\}) \rtimes \mathbb{Z}_2$. We have the following six-term exact sequence:

$$\begin{array}{ccccc} \mathbb{Z}^2 & \longrightarrow & K_0(C_0(X_2) \rtimes \mathbb{Z}_2) & \longrightarrow & 0 \\ \uparrow & & & & \downarrow \\ \mathbb{Z}^2 & \longleftarrow & K_1(C_0(X_2) \rtimes \mathbb{Z}_2) & \longleftarrow & 0, \end{array}$$

which implies $K_0(C_0(X_2) \rtimes \mathbb{Z}_2) \cong 0$ and $K_1(C_0(X_2) \rtimes \mathbb{Z}_2) \cong 0$. Thus,

$$\begin{array}{ccccc} 0 & \longrightarrow & K_0(C(\mathbb{T}^2) \rtimes \mathbb{Z}_2) & \longrightarrow & \mathbb{Z}^{2^3} \\ \uparrow & & & & \downarrow \\ 0 & \longleftarrow & K_1(C(\mathbb{T}^2) \rtimes \mathbb{Z}_2) & \longleftarrow & 0, \end{array}$$

which implies $K_0(C(\mathbb{T}^2) \rtimes \mathbb{Z}_2) \cong \mathbb{Z}^{2^3}$ and $K_1(C(\mathbb{T}^2) \rtimes \mathbb{Z}_2) \cong 0$. Therefore,

$$\begin{array}{ccccc} \mathbb{Z}^2 & \longrightarrow & K_0(E_\varepsilon^2 \rtimes \mathbb{Z}_2) & \longrightarrow & \mathbb{Z}^{2^3} \\ \uparrow & & & & \downarrow \\ 0 & \longleftarrow & K_1(E_\varepsilon^2 \rtimes \mathbb{Z}_2) & \longleftarrow & 0. \end{array}$$

It follows that $K_0(E_\varepsilon^2 \rtimes \mathbb{Z}_2) \cong \mathbb{Z}^{2^3+2}$ and $K_1(E_\varepsilon^2 \rtimes \mathbb{Z}_2) \cong 0$. Therefore,

$$\begin{array}{ccccc} \mathbb{Z} & \longrightarrow & \mathbb{Z}^{2^3+2} \oplus \mathbb{Z}^{2^3+2} & \longrightarrow & K_0(F_\varepsilon^2) \\ \uparrow & & & & \downarrow \\ K_1(F_\varepsilon^2) & \longleftarrow & 0 \oplus 0 & \longleftarrow & 0. \end{array}$$

Hence, it follows that $K_0(F_\varepsilon^2) \cong \mathbb{Z}^{2^4+3}$ and $K_1(F_\varepsilon^2) \cong 0$.

Next consider the case $m = 3$. Then

$$0 \to C_0(\mathbb{T}^3 \setminus (\pm 1, \pm 1, \pm 1)) \rtimes \mathbb{Z}_2 \to C(\mathbb{T}^3) \rtimes \mathbb{Z}_2 \to \oplus^{2^3} C^*(\mathbb{Z}_2) \to 0.$$

Furthermore, $0 \to C_0(X_2) \rtimes \mathbb{Z}_2 \to C_0(X_3) \rtimes \mathbb{Z}_2 \to C_0(X_3 \setminus X_2) \rtimes \mathbb{Z}_2 \to 0$, where $X_3 = \mathbb{T}^3 \setminus (\pm 1, \pm 1, \pm 1)$, and

$$0 \to C_0(X_1) \rtimes \mathbb{Z}_2 \to C_0(X_2) \rtimes \mathbb{Z}_2 \to C_0(X_2 \setminus X_1) \rtimes \mathbb{Z}_2 \to 0,$$

where $X_1 = (\mathbb{T} \setminus \{\pm 1\})^3$. We have the following six-term exact sequence:

$$\begin{array}{ccccc} 0 & \longrightarrow & K_0(C_0(X_2) \rtimes \mathbb{Z}_2) & \longrightarrow & \mathbb{Z}^6 \\ \uparrow & & & & \downarrow \\ 0 & \longleftarrow & K_1(C_0(X_2) \rtimes \mathbb{Z}_2) & \longleftarrow & \mathbb{Z}^3, \end{array}$$

which implies $K_0(C_0(X_2) \rtimes \mathbb{Z}_2) \cong \mathbb{Z}^3$ and $K_1(C_0(X_2) \rtimes \mathbb{Z}_2) \cong 0$. Furthermore,

$$\begin{array}{ccccc} \mathbb{Z}^3 & \longrightarrow & K_0(C_0(X_3) \rtimes \mathbb{Z}_2) & \longrightarrow & 0 \\ \uparrow & & & & \downarrow \\ \mathbb{Z}^3 & \longleftarrow & K_1(C_0(X_3) \rtimes \mathbb{Z}_2) & \longleftarrow & 0, \end{array}$$

which implies $K_0(C_0(X_3) \rtimes \mathbb{Z}_2) \cong 0$ and $K_1(C_0(X_3) \rtimes \mathbb{Z}_2) \cong 0$. Thus,

$$\begin{array}{ccccc} 0 & \longrightarrow & K_0(C(\mathbb{T}^3) \rtimes \mathbb{Z}_2) & \longrightarrow & \mathbb{Z}^{2^4} \\ \uparrow & & & & \downarrow \\ 0 & \longleftarrow & K_1(C(\mathbb{T}^3) \rtimes \mathbb{Z}_2) & \longleftarrow & 0, \end{array}$$

which implies $K_0(C(\mathbb{T}^3) \rtimes \mathbb{Z}_2) \cong \mathbb{Z}^{2^4}$ and $K_1(C(\mathbb{T}^2) \rtimes \mathbb{Z}_2) \cong 0$. Therefore,

$$\begin{array}{ccccc} \mathbb{Z}^2 & \longrightarrow & K_0(E_\varepsilon^3 \rtimes \mathbb{Z}_2) & \longrightarrow & \mathbb{Z}^{2^4} \\ \uparrow & & & & \downarrow \\ 0 & \longleftarrow & K_1(E_\varepsilon^3 \rtimes \mathbb{Z}_2) & \longleftarrow & 0. \end{array}$$

It follows that $K_0(E_\varepsilon^3 \rtimes \mathbb{Z}_2) \cong \mathbb{Z}^{2^4+2}$ and $K_1(E_\varepsilon^3 \rtimes \mathbb{Z}_2) \cong 0$. Therefore,

$$\begin{array}{ccccc} \mathbb{Z} & \longrightarrow & \mathbb{Z}^{2^4+2} \oplus \mathbb{Z}^{2^4+2} & \longrightarrow & K_0(F_\varepsilon^3) \\ \uparrow & & & & \downarrow \\ K_1(F_\varepsilon^3) & \longleftarrow & 0 \oplus 0 & \longleftarrow & 0. \end{array}$$

Hence, it follows that $K_0(F_\varepsilon^3) \cong \mathbb{Z}^{2^5+3}$ and $K_1(F_\varepsilon^3) \cong 0$.

Next consider the case $m = 4$. Then

$$0 \to C_0(\mathbb{T}^4 \setminus (\pm 1, \pm 1, \pm 1, \pm 1)) \rtimes \mathbb{Z}_2 \to C(\mathbb{T}^4) \rtimes \mathbb{Z}_2 \to \oplus^{2^4} C^*(\mathbb{Z}_2) \to 0.$$

Furthermore, $0 \to C_0(X_3) \rtimes \mathbb{Z}_2 \to C_0(X_4) \rtimes \mathbb{Z}_2 \to C_0(X_4 \setminus X_3) \rtimes \mathbb{Z}_2 \to 0$, where $X_4 = \mathbb{T}^4 \setminus (\pm 1, \pm 1, \pm 1, \pm 1)$, and

$$0 \to C_0(X_1) \rtimes \mathbb{Z}_2 \to C_0(X_2) \rtimes \mathbb{Z}_2 \to C_0(X_2 \setminus X_1) \rtimes \mathbb{Z}_2 \to 0,$$

where $X_1 = (\mathbb{T} \setminus \{\pm 1\})^4$. We have the following six-term exact sequence:

$$\begin{array}{ccccc} \mathbb{Z}^4 & \longrightarrow & K_0(C_0(X_2) \rtimes \mathbb{Z}_2) & \longrightarrow & 0 \\ \uparrow & & & & \downarrow \\ \mathbb{Z}^{12} & \longleftarrow & K_1(C_0(X_2) \rtimes \mathbb{Z}_2) & \longleftarrow & 0, \end{array}$$

which implies $K_0(C_0(X_2) \rtimes \mathbb{Z}_2) \cong 0$ and $K_1(C_0(X_2) \rtimes \mathbb{Z}_2) \cong \mathbb{Z}^8$. Furthermore,

$$\begin{array}{ccccc} 0 & \longrightarrow & K_0(C_0(X_3) \rtimes \mathbb{Z}_2) & \longrightarrow & \mathbb{Z}^{12} \\ \uparrow & & & & \downarrow \\ 0 & \longleftarrow & K_1(C_0(X_3) \rtimes \mathbb{Z}_2) & \longleftarrow & \mathbb{Z}^8, \end{array}$$

which implies $K_0(C_0(X_3) \rtimes \mathbb{Z}_2) \cong \mathbb{Z}^4$ and $K_1(C_0(X_3) \rtimes \mathbb{Z}_2) \cong 0$. Furthermore,

$$\begin{array}{ccccc} \mathbb{Z}^4 & \longrightarrow & K_0(C_0(X_4) \rtimes \mathbb{Z}_2) & \longrightarrow & 0 \\ \uparrow & & & & \downarrow \\ \mathbb{Z}^4 & \longleftarrow & K_1(C_0(X_4) \rtimes \mathbb{Z}_2) & \longleftarrow & 0, \end{array}$$

which implies $K_0(C_0(X_4) \rtimes \mathbb{Z}_2) \cong 0$ and $K_1(C_0(X_4) \rtimes \mathbb{Z}_2) \cong 0$. Thus,

$$\begin{array}{ccccc} 0 & \longrightarrow & K_0(C(\mathbb{T}^4) \rtimes \mathbb{Z}_2) & \longrightarrow & \mathbb{Z}^{2^5} \\ \uparrow & & & & \downarrow \\ 0 & \longleftarrow & K_1(C(\mathbb{T}^4) \rtimes \mathbb{Z}_2) & \longleftarrow & 0, \end{array}$$

which implies $K_0(C(\mathbb{T}^4) \rtimes \mathbb{Z}_2) \cong \mathbb{Z}^{2^5}$ and $K_1(C(\mathbb{T}^4) \rtimes \mathbb{Z}_2) \cong 0$. Therefore,

$$\begin{array}{ccccc} \mathbb{Z}^2 & \longrightarrow & K_0(E_\varepsilon^4 \rtimes \mathbb{Z}_2) & \longrightarrow & \mathbb{Z}^{2^5} \\ \uparrow & & & & \downarrow \\ 0 & \longleftarrow & K_1(E_\varepsilon^4 \rtimes \mathbb{Z}_2) & \longleftarrow & 0. \end{array}$$

It follows that $K_0(E_\varepsilon^4 \rtimes \mathbb{Z}_2) \cong \mathbb{Z}^{2^5+2}$ and $K_1(E_\varepsilon^4 \rtimes \mathbb{Z}_2) \cong 0$. Therefore,

$$\begin{array}{ccccc} \mathbb{Z} & \longrightarrow & \mathbb{Z}^{2^5+2} \oplus \mathbb{Z}^{2^5+2} & \longrightarrow & K_0(F_\varepsilon^4) \\ \uparrow & & & & \downarrow \\ K_1(F_\varepsilon^4) & \longleftarrow & 0 \oplus 0 & \longleftarrow & 0. \end{array}$$

Hence, it follows that $K_0(F_\varepsilon^4) \cong \mathbb{Z}^{2^6+3}$ and $K_1(F_\varepsilon^4) \cong 0$.

The case for m general can be treated by the step by step argument as shown above. The argument for K-theory is inductive in a sense that it involves essentially suspensions and direct sums inductively. The second claim follows from considering the case $\varepsilon = 0$ and the same argument as above. □

Corollary 2.6 *For $0 \le \varepsilon < 2$, the natural onto $*$-homomorphism $\varphi_{\varepsilon,0}$ from $(\otimes_\varepsilon^{m+1}\mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2$ to $(\otimes^{m+1}\mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2$ sending $s_{\varepsilon,j}$ to $s_{0,j}$ $(1 \le j \le m+1)$ induces the isomorphism between their K-groups.*

Proposition 2.7 *There exists a continuous field of C^*-algebras on the closed interval $[0,2]$ such that fibers are $(\otimes_\varepsilon^{m+1}\mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2$ for $\varepsilon \in [0,2]$, and for any $a \in (\otimes_2^{m+1}\mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2$, the sections $[0,2] \ni \varepsilon \mapsto \varphi_\varepsilon(a) \in (\otimes_\varepsilon^{m+1}\mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2$ are continuous, where $\varphi_\varepsilon : (\otimes_2^{m+1}\mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2 \to (\otimes_\varepsilon^{m+1}\mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2$ is the natural onto $*$-homomorphism sending $s_{2,j}$ to $s_{\varepsilon,j}$ $(1 \le j \le m+1)$.*

Proof. As shown before, $(\otimes_\varepsilon^{m+1}\mathfrak{F}) \rtimes_\sigma \mathbb{Z}_2 \cong (E_\varepsilon^m \rtimes_{\alpha_\varepsilon} \mathbb{N}) \rtimes_\sigma \mathbb{Z}_2$. Furthermore, this is isomorphic to $p((E_\varepsilon^m \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z})p \rtimes_\sigma \mathbb{Z}_2$. Hence it follows that

$$\begin{aligned}((E_\varepsilon^m \rtimes_{\alpha_\varepsilon} \mathbb{N}) \rtimes_\sigma \mathbb{Z}_2) \otimes \mathbb{K} &\cong (p((E_\varepsilon^m \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z})p \rtimes_\sigma \mathbb{Z}_2) \otimes \mathbb{K} \\ &\cong (p((E_\varepsilon^m \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z})p \otimes \mathbb{K}) \rtimes_{\sigma \otimes \mathrm{id}} \mathbb{Z}_2 \\ &\cong (((E_\varepsilon^m \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z}) \otimes \mathbb{K}) \rtimes_{\sigma \otimes \mathrm{id}} \mathbb{Z}_2 \\ &\cong ((E_\varepsilon^m \otimes \mathbb{K} \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id} \otimes \mathrm{id}} \mathbb{Z}) \rtimes_{\sigma \otimes \mathrm{id}} \mathbb{Z}_2 \\ &\cong ((E_\varepsilon^m \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge \otimes \mathrm{id}} \mathbb{Z}) \rtimes_\sigma \mathbb{Z}_2 \\ &\cong (E_\varepsilon^m \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge * \sigma \otimes \mathrm{id}} (\mathbb{Z}_2 * \mathbb{Z}_2).\end{aligned}$$

It is deduced from [2] that there exists a continuous field of C^*-algebras on $[0,2]$ such that fibers are $(E_\varepsilon^m \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge * \sigma \otimes \mathrm{id}} (\mathbb{Z}_2 * \mathbb{Z}_2)$ for $\varepsilon \in [0,2]$, and for any $b \in (E_2^m \otimes \mathbb{K}) \rtimes_{\rho_2^\wedge * \sigma \otimes \mathrm{id}} (\mathbb{Z}_2 * \mathbb{Z}_2)$, the sections $[0,2] \ni \varepsilon \mapsto \psi_\varepsilon(b) \in (E_\varepsilon^m \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge * \sigma \otimes \mathrm{id}} (\mathbb{Z}_2 * \mathbb{Z}_2)$ are continuous, where ψ_ε is the unique onto $*$-homomorphism from $(E_2^m \otimes \mathbb{K}) \rtimes_{\rho_2^\wedge * \sigma \otimes \mathrm{id}} (\mathbb{Z}_2 * \mathbb{Z}_2)$ to $(E_\varepsilon^m \otimes \mathbb{K}) \rtimes_{\rho_\varepsilon^\wedge * \sigma \otimes \mathrm{id}} (\mathbb{Z}_2 * \mathbb{Z}_2)$. Cutting down this continuous field by cutting down the fibers from $((E_\varepsilon^m \rtimes_{\alpha_\varepsilon} \mathbb{N}) \rtimes_\sigma \mathbb{Z}_2) \otimes \mathbb{K}$ to $(E_\varepsilon^m \rtimes_{\alpha_\varepsilon} \mathbb{N}) \rtimes_\sigma \mathbb{Z}_2$ by minimal projections, we obtain the desired continuous field. □

Received: May 2008. Revised: June 2008.

References

[1] B. Blackadar, *K-theory for Operator Algebras*, Second Edition, Cambridge, (1998).

[2] G.A. Elliott, R. Exel and T.A. Loring, *The soft torus III: The flip*, J. Operator Theory 26 (1991), 333–344.

[3] R. Exel, *The soft torus and applications to almost commuting matrices*, Pacific J. Math. **160** no. 2, (1993), 207–217.

[4] R. Exel, *The soft torus: A variational analysis of commutator norms*, J. Funct. Anal. **126** (1994), 259–273.

[5] G.J. Murphy, *C^*-algebras and Operator theory*, Academic Press, (1990).

[6] W. Paschke, *The crossed product of a C^*-algebra by an endomorphism*, Proc. Amer. Math. Soc. **80**, no. 1, (1980), 113–118.

[7] M. Rørdam, *Classification of certain infinite simple C^*-algebras*, J. Funct. Anal. **131**, (1995), 415–458.

[8] T. Sudo, *Continuous fields of C^*-algebras by almost commuting isometries*, Int. J. Pure Appl. Math. **44** (3) (2008), 425–438.

[9] N.E. Wegge-Olsen, *K-theory and C^*-algebras*, Oxford Univ. Press (1993).

CUBO A Mathematical Journal
Vol.10, Nº03, (115–132). October 2008

Limit Cycles of Liénard-Type Dynamical Systems

Valery A. Gaiko*
Department of Mathematics,
Belarusian State University of Informatics and Radioelectronics,
L. Beda Str. 6-4, Minsk 220040 – Belarus
email: `valery.gaiko@yahoo.com`

ABSTRACT

In this paper, using geometric properties of the field rotation parameters, we present a solution of *Smale's Thirteenth Problem* on the maximum number of limit cycles for Liénard's polynomial system, generalize the obtained results for special classes of polynomial systems, and complete the global qualitative analysis of a piecewise linear dynamical system approximating a Liénard-type polynomial system with an arbitrary number of finite singularities.

RESUMEN

En este artículo, usando propriedades geometricas del campo de rotación de parametros, nosotros presentamos una solución del problema trece de Smale sobre el número máximo de ciclos límite para el sistema polinomial de Liénard, generalizamos los resultados obtenidos para clases especiales de sistemas polinomiales, y completamos el analisis cualitativo global de un sistema dinamico linear por pedazos aproximando un sistema polinomial de tipo Liérnard con un número arbitrario finito de singularidades.

*This work was supported by the Netherlands Organization for Scientific Research (NWO).

Key words and phrases: *Planar polynomial dynamical system, Liénard's polynomial system, generalized Liénard's cubic system, piecewise linear Liénard-type dynamical system, Hilbert's sixteenth problem, Smale's thirteenth problem, field rotation parameter, bifurcation, limit cycle.*

Math. Subj. Class.: *34C05, 34C07, 34C23, 37G05, 37G10, 37G15.*

1 Introduction

We consider planar dynamical systems

$$\dot{x} = P_n(x, y), \quad \dot{y} = Q_n(x, y), \tag{1.1}$$

where $P_n(x, y)$ and $Q_n(x, y)$ are polynomials with real coefficients in the real variables x, y and not greater than n degree. First of all, we consider a special case of (1.1): a classical Liénard's polynomial system of the form

$$\dot{x} = y, \quad \dot{y} = -x + \mu_1\, y + \mu_2\, y^2 + \mu_3\, y^3 + \ldots + \mu_{2k}\, y^{2k} + \mu_{2k+1}\, y^{2k+1}. \tag{1.2}$$

The main problem of qualitative theory of such systems is *Hilbert's Sixteenth Problem* on the maximum number and relative position of their limit cycles, i. e., closed isolated trajectories of (1.1). This problem was formulated as one of the fundamental problems for mathematicians of the XX century, however it has not been solved even in the simplest (quadratic, cubic, etc.) cases of the polynomial systems. In this paper, we suggest a new geometric approach to solving the problem in the case of Liénard's system (1.2). In this special case, it is considered as *Smale's Thirteenth Problem* becoming one of the main problems for mathematicians of the XXI century [16], [20].

In Section 2 of this paper, applying a canonical system with field rotation parameters and using geometric properties of the spirals filling the interior and exterior domains of limit cycles, we present a solution of *Smale's Thirteenth Problem* for Liénard's polynomial system (1.2). In Section 3, by means of the same geometric approach, we generalize the obtained result and present a solution of *Hilbert's sixteenth problem* on the maximum number of limit cycles surrounding a singular point for an arbitrary polynomial system. In Section 4, we consider generalized Liénard's cubic system with three finite singularities, for which the developed geometric approach can complete its global qualitative analysis: in particular, it easily solves the problem on the maximum number of limit cycles in their different distribution. In this section, we give also an alternative proof of the main theorem for the generalized Liénard's system applying the Wintner–Perko termination principle for multiple limit cycles. In Section 5, by means of the same principle, we complete the global qualitative analysis of a piecewise linear dynamical system approximating a Liénard-type polynomial system with an arbitrary number of finite singularities.

2 Liénard's polynomial system

System (1.2) and more general Liénard's systems have been studied in numerous works (see, for example, [2], [16], [17], [19], [20]). It is easy to see that (1.2) has the only finite singularity: an anti-saddle at the origin. At infinity, system (1.2) for $k \geqslant 1$ has two singular points: a node at the "ends" of the y-axis and a saddle at the "ends" of the x-axis. For studying the infinite singularities, the methods applied in [2] for Rayleigh's and van der Pol's equations and also Erugin's two-isocline method developed in [10] can be used. Following [10], we will study limit cycle bifurcations of (1.2) by means of a canonical system containing only the field rotation parameters of (1.2). It is valid the following theorem.

Theorem 2.1. *Liénard's polynomial system* (1.2) *with limit cycles can be reduced to the canonical form*

$$\dot{x} = y \equiv P, \quad \dot{y} = -x + \mu_1\, y + y^2 + \mu_3\, y^3 + \ldots + y^{2k} + \mu_{2k+1}\, y^{2k+1} \equiv Q, \tag{2.1}$$

where $\mu_1,\, \mu_3, \ldots,\, \mu_{2k+1}$ *are field rotation parameters of* (2.1).

Proof. Vanish all odd parameters of (1.2),

$$\dot{x} = y, \quad \dot{y} = -x + \mu_2\, y^2 + \mu_4\, y^4 + \ldots + \mu_{2k}\, y^{2k}, \tag{2.2}$$

and consider the corresponding equation

$$\frac{dy}{dx} = \frac{-x + \mu_2\, y^2 + \mu_4\, y^4 + \ldots + \mu_{2k}\, y^{2k}}{y} \equiv F(x, y). \tag{2.3}$$

Since $F(x, -y) = -F(x, y)$, the direction field of (2.3) (and the vector field of (2.2) as well) is symmetric with respect to the x-axis. It follows that for arbitrary values of the parameters μ_2, $\mu_4, \ldots,\, \mu_{2k}$ system (2.2) has a center at the origin and cannot have a limit cycle surrounding this point. Therefore, without loss of generality, all even parameters of system (1.2) can be supposed to be equal, for example, to one: $\mu_2 = \mu_4 = \ldots = \mu_{2k} = 1$ (they could be also supposed to be equal to zero).

To prove that the rest (odd) parameters rotate the vector field of (2.1), let us calculate the following determinants:

$$\begin{gathered}
\Delta_{\mu_1} = PQ'_{\mu_1} - QP'_{\mu_1} = y^2 \geq 0, \\
\Delta_{\mu_3} = PQ'_{\mu_3} - QP'_{\mu_3} = y^2 \geq 0, \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
\Delta_{\mu_{2k+1}} = PQ'_{\mu_{2k+1}} - QP'_{\mu_{2k+1}} = y^2 \geq 0.
\end{gathered}$$

By definition of a field rotation parameter [4], for increasing each of the parameters $\mu_1,\, \mu_3, \ldots,$ μ_{2k+1}, under the fixed others, the vector field of system (2.1) is rotated in positive direction (counterclockwise) in the whole phase plane; and, conversely, for decreasing each of these parameters, the vector field of (2.1) is rotated in negative direction (clockwise).

Thus, for studying limit cycle bifurcations of (1.2), it is sufficient to consider canonical system (2.1) containing only its odd parameters, $\mu_1,\, \mu_3, \ldots,\, \mu_{2k+1}$, which rotate the vector field of (2.1). The theorem is proved. □

By means of canonical system (2.1), let us study global limit cycle bifurcations of (1.2) and prove the following theorem.

Theorem 2.2. *Liénard's polynomial system* (1.2) *has at most k limit cycles.*

Proof. According to Theorem 2.1, for the study of limit cycle bifurcations of system (1.2), it is sufficient to consider canonical system (2.1) containing only the field rotation parameters of (1.2): $\mu_1,\, \mu_3, \ldots,\, \mu_{2k+1}$.

Vanish all these parameters:

$$\dot{x} = y, \quad \dot{y} = -x + y^2 + y^4 + \ldots + y^{2k}. \tag{2.4}$$

System (2.4) is symmetric with respect to the x-axis and has a center at the origin. Let us input successively the field rotation parameters into this system beginning with the parameters at the highest degrees of y and alternating with their signs. So, begin with the parameter μ_{2k+1} and let, for definiteness, $\mu_{2k+1} > 0$:

$$\dot{x} = y, \quad \dot{y} = -x + y^2 + y^4 + \ldots + y^{2k} + \mu_{2k+1}\, y^{2k+1}. \tag{2.5}$$

In this case, the vector field of (2.5) is rotated in positive direction (counterclockwise) turning the origin into a nonrough unstable focus.

Fix μ_{2k+1} and input the parameter $\mu_{2k-1} < 0$ into (2.5):

$$\dot{x} = y, \quad \dot{y} = -x + y^2 + y^4 + \ldots + \mu_{2k-1}\, y^{2k-1} + y^{2k} + \mu_{2k+1}\, y^{2k+1}. \tag{2.6}$$

Then the vector field of (2.6) is rotated in opposite direction (clockwise) and the focus immediately changes the character of its stability (since its degree of nonroughness decreases and the sign of the field rotation parameter at the lower degree of y changes) generating a stable limit cycle. Under further decreasing μ_{2k-1}, this limit cycle will expand infinitely, not disappearing at infinity (because of the parameter μ_{2k+1} at the higher degree of y).

Denote the limit cycle by Γ_1, the domain outside the cycle by D_1, the domain inside the cycle by D_2 and consider logical possibilities of the appearance of other (semi-stable) limit cycles from a "trajectory concentration" surrounding the origin. It is clear that, under decreasing the parameter μ_{2k-1}, a semi-stable limit cycle cannot appear in the domain D_2, since the focus spirals filling this domain will untwist and the distance between their coils will increase because of the vector field rotation.

By contradiction, we can also prove that a semi-stable limit cycle cannot appear in the domain D_1. Suppose it appears in this domain for some values of the parameters $\mu^*_{2k+1} > 0$ and $\mu^*_{2k-1} < 0$. Return to initial system (2.4) and change the inputting order for the field rotation parameters. Input first the parameter $\mu_{2k-1} < 0$:

$$\dot{x} = y, \quad \dot{y} = -x + y^2 + y^4 + \ldots + \mu_{2k-1}\, y^{2k-1} + y^{2k}. \tag{2.7}$$

Fix it under $\mu_{2k-1} = \mu^*_{2k-1}$. The vector field of (2.7) is rotated clockwise and the origin turns into a nonrough stable focus. Inputting the parameter $\mu_{2k+1} > 0$ into (2.7), we get again system (2.6), the vector field of which is rotated counterclockwise. Under this rotation, a stable limit cycle Γ_1 will immediately appear from infinity, more precisely, from a separatrix cycle of the Poincaré circle form containing infinite singularities of the saddle and node types [2]. This cycle will contract, the outside spirals winding onto the cycle will untwist and the distance between their coils will increase under increasing μ_{2k+1} to the value μ^*_{2k+1}. It follows that there are no values of $\mu^*_{2k-1} < 0$ and $\mu^*_{2k+1} > 0$, for which a semi-stable limit cycle could appear in the domain D_1.

This contradiction proves the uniqueness of a limit cycle surrounding the origin in system (2.6) for any values of the parameters μ_{2k-1} and μ_{2k+1} of different signs. Obviously, if these parameters have the same sign, system (2.6) has no limit cycles surrounding the origin at all.

Let system (2.6) have the unique limit cycle Γ_1. Fix the parameters $\mu_{2k+1} > 0$, $\mu_{2k-1} < 0$ and input the third parameter, $\mu_{2k-3} > 0$, into this system:

$$\dot{x} = y, \quad \dot{y} = -x + y^2 + \ldots + \mu_{2k-3}\, y^{2k-3} + y^{2k-2} + \ldots + \mu_{2k+1}\, y^{2k+1}. \tag{2.8}$$

The vector field of (2.8) is rotated counterclockwise, the focus at the origin changes the character of its stability and the second (unstable) limit cycle, Γ_2, immediately appears from this point. Under further increasing μ_{2k-3}, the limit cycle Γ_2 will join with Γ_1 forming a semi-stable limit cycle, Γ_{12}, which will disappear in a "trajectory concentration" surrounding the origin. Can another semi-stable limit cycle appear around the origin in addition to Γ_{12}? It is clear that such a limit cycle cannot appear either in the domain D_1 bounded on the inside by the cycle Γ_1 or in the domain D_3 bounded by the origin and Γ_2 because of the increasing distance between the spiral coils filling these domains under increasing the parameter μ_{2k-3}.

To prove impossibility of the appearance of a semi-stable limit cycle in the domain D_2 bounded by the cycles Γ_1 and Γ_2 (before their joining), suppose the contrary, i. e., for some set of values of the parameters, $\mu^*_{2k+1} > 0$, $\mu^*_{2k-1} < 0$, and $\mu^*_{2k-3} > 0$, such a semi-stable cycle exists. Return to system (2.4) again and input first the parameters $\mu_{2k-3} > 0$ and $\mu_{2k+1} > 0$:

$$\dot{x} = y, \quad \dot{y} = -x + y^2 + \ldots + \mu_{2k-3}\, y^{2k-3} + y^{2k-2} + y^{2k} + \mu_{2k+1}\, y^{2k+1}. \tag{2.9}$$

Both parameters act in a similar way: they rotate the vector field of (2.9) counterclockwise turning the origin into a nonrough unstable focus.

Fix these parameters under $\mu_{2k-3} = \mu^*_{2k-3}$, $\mu_{2k+1} = \mu^*_{2k+1}$ and input the parameter $\mu_{2k-1} < 0$ into (2.9) getting again system (2.8). Since, by our assumption, this system has two limit cycles for $\mu_{2k-1} > \mu^*_{2k-1}$, there exists some value of the parameter, μ^{12}_{2k-1} ($\mu^*_{2k-1} < \mu^{12}_{2k-1} < 0$), for which a semi-stable limit cycle, Γ_{12}, appears in system (2.8) and then splits into a stable cycle, Γ_1, and an unstable cycle, Γ_2, under further decreasing μ_{2k-1}. The formed domain D_2 bounded by the limit cycles Γ_1, Γ_2 and filled by the spirals will enlarge since, on the properties of a field rotation parameter, the interior unstable limit cycle Γ_2 will contract and the exterior stable limit cycle Γ_1 will expand under decreasing μ_{2k-1}. The distance between the spirals of the domain D_2 will naturally increase, what will prevent the appearance of a semi-stable limit cycle in this domain for $\mu_{2k-1} < \mu^{12}_{2k-1}$.

Thus, there are no such values of the parameters, $\mu^*_{2k+1} > 0$, $\mu^*_{2k-1} < 0$, and $\mu^*_{2k-3} > 0$, for which system (2.8) would have an additional semi-stable limit cycle. Obviously, there are no other values of the parameters μ_{2k+1}, μ_{2k-1}, and μ_{2k-3} for which system (2.8) would have more than two limit cycles surrounding the origin. Therefore, two is the maximum number of limit cycles for system (2.8). This result agrees with [19], where it was proved for the first time that the maximum number of limit cycles for Liénard's system of the form

$$\dot{x} = y, \quad \dot{y} = -x + \mu_1\, y + \mu_3\, y^3 + \mu_5\, y^5 \tag{2.10}$$

was equal to two.

Suppose that system (2.8) has two limit cycles, Γ_1 and Γ_2 (this is always possible if $\mu_{2k+1} \gg -\mu_{2k-1} \gg \mu_{2k-3} > 0$), fix the parameters μ_{2k+1}, μ_{2k-1}, μ_{2k-3} and consider a more general system than (2.8) (and (2.10)) inputting the fourth parameter, $\mu_{2k-5} < 0$, into (2.8):

$$\dot{x} = y, \quad \dot{y} = -x + y^2 + \ldots + \mu_{2k-5}\, y^{2k-5} + y^{2k-4} + \ldots + \mu_{2k+1}\, y^{2k+1}. \tag{2.11}$$

Under decreasing μ_{2k-5}, the vector field of (2.11) will be rotated clockwise and the focus at the origin will immediately change the character of its stability generating the third (stable) limit cycle, Γ_3. Under further decreasing μ_{2k-5}, Γ_3 will join with Γ_2 forming a semi-stable limit cycle, Γ_{23}, which will disappear in a "trajectory concentration" surrounding the origin; the cycle Γ_1 will expand infinitely tending to the Poincaré circle at infinity.

Let system (2.11) have three limit cycles: Γ_1, Γ_2, Γ_3. Could an additional semi-stable limit cycle appear under decreasing μ_{2k-5}, after splitting of which system (2.11) would have five limit cycles around the origin? It is clear that such a limit cycle cannot appear either in the domain D_2 bounded by the cycles Γ_1 and Γ_2 or in the domain D_4 bounded by the origin and Γ_3 because of the increasing distance between the spiral coils filling these domains under decreasing μ_{2k-5}. Consider two other domains: D_1 bounded on the inside by the cycle Γ_1 and D_3 bounded by the cycles Γ_2 and Γ_3. As before, we will prove impossibility of the appearance of a semi-stable limit cycle in these domains by contradiction.

Suppose that for some set of values of the parameters $\mu^*_{2k+1} > 0$, $\mu^*_{2k-1} < 0$, $\mu^*_{2k-3} > 0$, and $\mu^*_{2k-5} < 0$, such a semi-stable cycle exists. Return to system (2.4) again, input first the parameters $\mu_{2k-5} < 0$, $\mu_{2k-1} < 0$ and then the parameter $\mu_{2k+1} > 0$:

$$\dot{x} = y, \quad \dot{y} = -x + y^2 + \ldots + \mu_{2k-5} y^{2k-5} + \ldots + \mu_{2k-1} y^{2k-1} + y^{2k} + \mu_{2k+1} y^{2k+1}. \tag{2.12}$$

Fix the parameters μ_{2k-5}, μ_{2k-1} under the values μ^*_{2k-5}, μ^*_{2k-1}, respectively. Under increasing μ_{2k+1}, the node at infinity will change the character of its stability, the separatrix behavior of the infinite saddle will be also changed and a stable limit cycle, Γ_1, will immediately appear from the Poincaré circle at infinity [2]. Fix μ_{2k+1} under the value μ^*_{2k+1} and input the parameter $\mu_{2k-3} > 0$ into (2.12) getting system (2.11).

Since, by our assumption, (2.11) has three limit cycles for $\mu_{2k-3} < \mu^*_{2k-3}$, there exists some value of the parameter μ^{23}_{2k-3} $(0 < \mu^{23}_{2k-3} < \mu^*_{2k-3})$ for which a semi-stable limit cycle, Γ_{23}, appears in this system and then splits into an unstable cycle, Γ_2, and a stable cycle, Γ_3, under further increasing μ_{2k-3}. The formed domain D_3 bounded by the limit cycles Γ_2, Γ_3 and also the domain D_1 bounded on the inside by the limit cycle Γ_1 will enlarge and the spirals filling these domains will untwist excluding a possibility of the appearance of a semi-stable limit cycle there.

All other combinations of the parameters μ_{2k+1}, μ_{2k-1}, μ_{2k-3}, and μ_{2k-5} are considered in a similar way. It follows that system (2.11) has at most three limit cycles. If we continue the procedure of successive inputting the odd parameters, $\mu_{2k-7}, \ldots, \mu_3, \mu_1$, into system (2.4), it is possible first to obtain k limit cycles ($\mu_{2k+1} \gg -\mu_{2k-1} \gg \mu_{2k-3} \gg -\mu_{2k-5} \gg \mu_{2k-7} \gg \ldots$) and then to conclude that canonical system (2.1) (i. e., Liénard's polynomial system (1.2) as well) has at most k limit cycles. The theorem is proved. □

3 An arbitrary polynomial system

Let us consider an arbitrary polynomial system

$$\dot{x} = P_n(x, y, \mu_1, \ldots, \mu_k), \quad \dot{y} = Q_n(x, y, \mu_1, \ldots, \mu_k), \tag{3.1}$$

where P_n and Q_n are polynomials in the real variables x, y and not greater than n degree containing k field rotation parameters, $\mu_1, \ldots, \mu_k$, and having an anti-saddle at the origin. Generalizing the main result of the previous section, we will prove the following theorem.

Theorem 3.1. *Polynomial system* (3.1) *containing k field rotation parameters and having a singular point of the center type at the origin for the zero values of these parameters can have at most $k-1$ limit cycles surrounding the origin.*

Proof. Vanish all parameters of (3.1) and suppose that the obtained system

$$\dot{x} = P_n(x, y, 0, \ldots, 0), \quad \dot{y} = Q_n(x, y, 0, \ldots, 0) \tag{3.2}$$

has a singular point of the center type at the origin. Let us input successively the field rotation parameters, $\mu_1, \ldots, \mu_k$, into this system.

Suppose, for example, that $\mu_1 > 0$ and that the vector field of the system

$$\dot{x} = P_n(x, y, \mu_1, 0, \ldots, 0), \quad \dot{y} = Q_n(x, y, \mu_1, 0, \ldots, 0) \tag{3.3}$$

is rotated counterclockwise turning the origin into a stable focus under increasing μ_1.

Fix μ_1 and input the parameter μ_2 into (3.3) changing it so that the field of the system

$$\dot{x} = P_n(x, y, \mu_1, \mu_2, 0, \ldots, 0), \quad \dot{y} = Q_n(x, y, \mu_1, \mu_2, 0, \ldots, 0) \tag{3.4}$$

would be rotated in opposite direction (clockwise). Let be so for $\mu_2 < 0$. Then, for some value of this parameter, a limit cycle will appear in system (3.4). There are three logical possibilities for such a bifurcation: 1) the limit cycle appears from the focus at the origin; 2) it can also appear from some separatrix cycle surrounding the origin; 3) the limit cycle appears from a so-called "trajectory concentration". In the last case, the limit cycle is semi-stable and, under further decreasing μ_2, it splits into two limit cycles (stable and unstable), one of which then disappears at (or tends to) the origin and the other disappears on (or tends to) some separatrix cycle surrounding this point. But since the stability character of both a singular point and a separatrix cycle is quite easily controlled [10], this logical possibility can be excluded. Let us choose one of the two other possibilities: for example, the first one, the so-called Andronov–Hopf bifurcation. Suppose that, for some value of μ_2, the focus at the origin becomes non-rough, changes the character of its stability and generates a stable limit cycle, Γ_1.

Under further decreasing μ_2, three new logical possibilities can arise: 1) the limit cycle Γ_1 disappears on some separatrix cycle surrounding the origin; 2) a separatrix cycle can be formed earlier than Γ_1 disappears on it, then it generates one more (unstable) limit cycle, Γ_2, which joins with Γ_1 forming a semi-stable limit cycle, Γ_{12}, disappearing in a "trajectory concentration" under further decreasing μ_2; 3) in the domain D_1 outside the cycle Γ_1 or in the domain D_2 inside Γ_1, a semi-stable limit cycle appears from a "trajectory concentration" and then splits into two limit cycles (logically, the appearance of such semi-stable limit cycles can be repeated).

Let us consider the third case. It is clear that, under decreasing μ_2, a semi-stable limit cycle cannot appear in the domain D_2, since the focus spirals filling this domain will untwist and the distance between their coils will increase because of the vector field rotation. By contradiction, we can prove that a semi-stable limit cycle cannot appear in the domain D_1. Suppose it appears in this domain for some values of the parameters $\mu_1^* > 0$ and $\mu_2^* < 0$. Return to initial system (3.2) and change the inputting order for the field rotation parameters. Input first the parameter $\mu_2 < 0$:

$$\dot{x} = P_n(x, y, \mu_2, 0, \ldots, 0), \quad \dot{y} = Q_n(x, y, \mu_2, 0, \ldots, 0). \tag{3.5}$$

Fix it under $\mu_2 = \mu_2^*$. The vector field of (3.5) is rotated clockwise and the origin turns into a unstable focus. Inputting the parameter $\mu_1 > 0$ into (3.5), we get again system (3.4), the vector field of which is rotated counterclockwise. Under this rotation, a stable limit cycle, Γ_1, will appear from some separatrix cycle. The limit cycle Γ_1 will contract, the outside spirals winding onto this cycle will untwist and the distance between their coils will increase under increasing μ_1 to the value μ_1^*. It follows that there are no values of $\mu_2^* < 0$ and $\mu_1^* > 0$, for which a semi-stable limit cycle could appear in the domain D_1.

The second logical possibility can be excluded by controlling the stability character of the separatrix cycle [10]. Thus, only the first possibility is valid, i. e., system (3.4) has at most one limit cycle.

Let system (3.4) have the unique limit cycle Γ_1. Fix the parameters $\mu_1 > 0$, $\mu_2 < 0$ and input the third parameter, $\mu_3 > 0$, into this system supposing that μ_3 rotates its vector field counterclockwise:

$$\dot{x} = P_n(x, y, \mu_1, \mu_2, \mu_3, 0, \ldots, 0), \quad \dot{y} = Q_n(x, y, \mu_1, \mu_2, \mu_3, 0, \ldots, 0). \tag{3.6}$$

Here we can have two basic possibilities: 1) the limit cycle Γ_1 disappears at the origin; 2) the second (unstable) limit cycle, Γ_2, appears from the origin and, under further increasing the parameter μ_3, the cycle Γ_2 joins with Γ_1 forming a semi-stable limit cycle, Γ_{12}, which disappears in a "trajectory concentration" surrounding the origin. Besides, we can also suggest that: 3) in the domain D_2 bounded by the origin and Γ_1, a semi-stable limit cycle, Γ_{23}, appears from a "trajectory concentration", splits into an unstable cycle, Γ_2, and a stable cycle, Γ_3, and then the cycles Γ_1, Γ_2 disappear through a semi-stable limit cycle, Γ_{12}, and the cycle Γ_3 disappears through the Andronov–Hopf bifurcation; 4) a semi-stable limit cycle, Γ_{34}, appears in the domain D_2 bounded by the cycles Γ_1, Γ_2 and, for some set of values of the parameters, μ_1^*, μ_2^*, μ_3^*, system (3.6) has at least four limit cycles.

Let us consider the last, fourth, case. It is clear that a semi-stable limit cycle cannot appear either in the domain D_1 bounded on the inside by the cycle Γ_1 or in the domain D_3 bounded by the origin and Γ_2 because of the increasing distance between the spiral coils filling these domains under increasing the parameter μ_3. To prove impossibility of the appearance of a semi-stable limit cycle in the domain D_2, suppose the contrary, i. e., for some set of values of the parameters, $\mu_1^* > 0$, $\mu_2^* < 0$, and $\mu_3^* > 0$, such a semi-stable cycle exists. Return to system (3.2) again and input first the parameters $\mu_3 > 0$, $\mu_1 > 0$:

$$\dot{x} = P_n(x, y, \mu_1, \mu_3, 0, \ldots, 0), \quad \dot{y} = Q_n(x, y, \mu_1, \mu_3, 0, \ldots, 0). \tag{3.7}$$

Fix these parameters under $\mu_3 = \mu_3^*$, $\mu_1 = \mu_1^*$ and input the parameter $\mu_2 < 0$ into (3.7) getting again system (3.6). Since, by our assumption, this system has two limit cycles for $\mu_2 > \mu_2^*$, there exists some value of the parameter, μ_2^{12} ($\mu_2^* < \mu_2^{12} < 0$), for which a semi-stable limit cycle, Γ_{12}, appears in system (3.6) and then splits into a stable cycle, Γ_1, and an unstable cycle, Γ_2, under further decreasing μ_2. The formed domain D_2 bounded by the limit cycles Γ_1, Γ_2 and filled by the spirals will enlarge, since, on the properties of a field rotation parameter, the interior unstable limit cycle Γ_2 will contract and the exterior stable limit cycle Γ_1 will expand under decreasing μ_2. The distance between the spirals of the domain D_2 will naturally increase, what will prevent the appearance of a semi-stable limit cycle in this domain for $\mu_2 < \mu_2^{12}$.

Thus, there are no such values of the parameters, $\mu_1^* > 0$, $\mu_2^* < 0$, $\mu_3^* > 0$, for which system (3.6) would have an additional semi-stable limit cycle. Therefore, the fourth case cannot be realized. The third case is considered absolutely similarly. It follows from the first two cases that system (3.6) can have at most two limit cycles.

Suppose that system (3.6) has two limit cycles, Γ_1 and Γ_2, fix the parameters $\mu_1 > 0$, $\mu_2 < 0$, $\mu_3 > 0$ and input the fourth parameter, $\mu_4 < 0$, into this system supposing that μ_4 rotates its vector field clockwise:

$$\dot{x} = P_n(x, y, \mu_1, \ldots, \mu_4, 0, \ldots, 0), \quad \dot{y} = Q_n(x, y, \mu_1, \ldots, \mu_4, 0, \ldots, 0). \tag{3.8}$$

The most interesting logical possibility here is that when the third (stable) limit cycle, Γ_3, appears from the origin and then, under preservation of the cycles Γ_1 and Γ_2, in the domain D_3 bounded on the inside by the cycle Γ_3 and on the outside by the cycle Γ_2, a semi-stable limit cycle, Γ_{45}, appears and then splits into a stable cycle, Γ_4, and an unstable cycle, Γ_5, i. e., when system (3.8) for some set of values of the parameters, μ_1^*, μ_2^*, μ_3^*, μ_4^*, has at least five limit cycles. Logically, such a semi-stable limit cycle could also appear in the domain D_1 bounded on the inside by the cycle Γ_1, since, under decreasing μ_4, the spirals of the trajectories of (3.8) will twist and the distance between their coils will decrease. On the other hand, in the domain D_2 bounded on the inside by the cycle Γ_2 and on the outside by the cycle Γ_1 and also in the domain D_4 bounded by the origin and Γ_3, a semi-stable limit cycle cannot appear, since, under decreasing μ_4, the spirals will untwist and the distance between their coils will increase. To prove impossibility of the appearance of a semi-stable limit cycle in the domains D_3 and D_1, suppose the contrary, i. e., for some set of values of the parameters, $\mu_1^* > 0$, $\mu_2^* < 0$, $\mu_3^* > 0$, and $\mu_4^* < 0$, such a semi-stable cycle exists. Return to system (3.2) again, input first the parameters $\mu_4 < 0$, $\mu_2 < 0$ and then the parameter $\mu_1 > 0$:

$$\dot{x} = P_n(x, y, \mu_1, \mu_2, \mu_4, 0, \ldots, 0), \quad \dot{y} = Q_n(x, y, \mu_1, \mu_2, \mu_4, 0, \ldots, 0). \tag{3.9}$$

Fix the parameters μ_4, μ_2 under the values μ_4^*, μ_2^*, respectively. Under increasing μ_1, a separatrix cycle is formed around the origin generating a stable limit cycle, Γ_1. Fix μ_1 under the value μ_1^* and input the parameter $\mu_3 > 0$ into (3.9) getting system (3.8).

Since, by our assumption, system (3.8) has three limit cycles for $\mu_3 < \mu_3^*$, there exists some value of the parameter μ_3^{23} $(0 < \mu_3^{23} < \mu_3^*)$ for which a semi-stable limit cycle, Γ_{23}, appears in this system and then splits into an unstable cycle, Γ_2, and a stable cycle, Γ_3, under further increasing μ_3. The formed domain D_3 bounded by the limit cycles Γ_2, Γ_3 and also the domain D_1 bounded on the inside by the limit cycle Γ_1 will enlarge and the spirals filling these domains will untwist excluding a possibility of the appearance of a semi-stable limit cycle there.

All other combinations of the parameters μ_1, μ_2, μ_3, and μ_4 are considered in a similar way. It follows that system (3.8) has at most three limit cycles. If we continue the procedure of successive inputting the field rotation parameters, $\mu_5, \mu_6, \ldots, \mu_k$, into system (3.2), it is possible to conclude that system (3.1) can have at most $k - 1$ limit cycles surrounding the origin. The theorem is proved. □

4 Generalized liénard's cubic system

In [13], we considered generalized Liénard's cubic system of the form:

$$\dot{x} = y, \quad \dot{y} = -x + (\lambda - \mu)\, y + (3/2)\, x^2 + \mu\, xy - (1/2)\, x^3 + \alpha\, x^2 y\,. \tag{4.1}$$

This system has three finite singularities: a saddle $(1,0)$ and two antisaddles — $(0,0)$ and $(2,0)$. At infinity system (4.1) can have either the only nilpotent singular point of fourth order with two closed elliptic and four hyperbolic domains or two singular points: one of them is a hyperbolic saddle and the other is a triple nilpotent singular point with two elliptic and two hyperbolic domains. We studied global bifurcations of limit and separatrix cycles of (4.1), found possible distributions of its limit cycles and carried out a classification of its separatrix cycles. We proved also the following theorems.

Theorem 4.1. *The foci of system* (4.1) *can be at most of second order.*

Theorem 4.2. *System* (4.1) *has at least three limit cycles.*

Using the results obtained in [13] and applying the approach developed in this paper, we can easily prove a much stronger theorem.

Theorem 4.3. *System* (4.1) *has at most three limit cycles with the following their distributions*: $((1,1),1)$, $((1,2),0)$, $((2,1),0)$, $((1,0),2)$, $((0,1),2)$, *where the first two numbers denote the numbers of limit cycles surrounding each of two anti-saddles and the third one denotes the number of limit cycles surrounding simultaneously all three finite singularities.*

Theorem 4.3 agrees, for example, with the earlier results by Iliev and Perko [15], but it does not agree with a quite recent result by Dumortier and Li [5] published in the same journal. The authors of both papers use very similar methods: small perturbations of a Hamiltonian system. In [15], the zeros of the Melnikov functions are studied and, in particular, it is proved that at most two limit cycles can bifurcate from either the interior or exterior period annulus of the Hamiltonian under small parameter perturbations giving a generalized Liénard system. In [5], zeros of the Abelian integrals are studied and it is "proved" that at most four limit cycles can bifurcate from the exterior period annulus. Thus, Dumortier and Li "obtain" a configuration of four big limit cycles surrounding three finite singularities together with the fifth small limit cycle which surrounds one of the anti-saddles.

The result by Dumortier and Li [5] also does not agree with the Wintner–Perko termination principle for multiple limit cycles [10], [18]. Applying the method as developed in [3], [7]–[13], we can show that system (4.1) cannot have either a multiplicity-three limit cycle or more than three limit cycles in any configuration. That will be another proof of Theorem 4.3 (the same approach can be applied to proving Theorems 2.2 and 3.1 as well). But first let us formulate the Wintner–Perko termination principle [18] for the polynomial system

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{\mu}), \tag{4.2$\boldsymbol{\mu}$}$$

where $\boldsymbol{x} \in \mathbf{R}^2$; $\boldsymbol{\mu} \in \mathbf{R}^n$; $\boldsymbol{f} \in \mathbf{R}^2$ ($\boldsymbol{f}$ is a polynomial vector function).

Theorem 4.4 (Wintner–Perko termination principle). *Any one-parameter family of multiplicity-m limit cycles of relatively prime polynomial system* $(4.2_{\boldsymbol{\mu}})$ *can be extended in a unique way to a maximal one-parameter family of multiplicity-m limit cycles of* $(4.2_{\boldsymbol{\mu}})$ *which is either open or cyclic.*

If it is open, then it terminates either as the parameter or the limit cycles become unbounded; or, the family terminates either at a singular point of $(4.2_{\boldsymbol{\mu}})$, *which is typically a fine focus of multiplicity* m, *or on a* (*compound*) *separatrix cycle of* $(4.2_{\boldsymbol{\mu}})$, *which is also typically of multiplicity* m.

The proof of this principle for general polynomial system $(4.2_{\boldsymbol{\mu}})$ with a vector parameter $\boldsymbol{\mu} \in \mathbf{R}^n$ parallels the proof of the planar termination principle for the system

$$\dot{x} = P(x, y, \lambda), \quad \dot{y} = Q(x, y, \lambda) \tag{4.2_λ}$$

with a single parameter $\lambda \in \mathbf{R}$ (see [10], [18]), since there is no loss of generality in assuming that system $(4.2_{\boldsymbol{\mu}})$ is parameterized by a single parameter λ; i. e., we can assume that there exists an analytic mapping $\boldsymbol{\mu}(\lambda)$ of $\mathbf{R}$ into $\mathbf{R}^n$ such that $(4.2_{\boldsymbol{\mu}})$ can be written as $(4.2\,_{\boldsymbol{\mu}(\lambda)})$ or even (4.2_λ) and then we can repeat everything, what had been done for system (4.2_λ) in [18]. In particular, if λ is a field rotation parameter of (4.2_λ), the following Perko's theorem on monotonic families of limit cycles is valid.

Theorem 4.5. *If* L_0 *is a nonsingular multiple limit cycle of* (4.2_0), *then* L_0 *belongs to a one-parameter family of limit cycles of* (4.2_λ); *furthermore*:

1) *if the multiplicity of* L_0 *is odd, then the family either expands or contracts monotonically as* λ *increases through* λ_0;

2) *if the multiplicity of* L_0 *is even, then* L_0 *bifurcates into a stable and an unstable limit cycle as* λ *varies from* λ_0 *in one sense and* L_0 *disappears as* λ *varies from* λ_0 *in the opposite sense; i. e., there is a fold bifurcation at* λ_0.

Proof of Theorem 4.3. The proof is carried out by contradiction. Suppose that system (4.1) with three field rotation parameters, λ, μ, and α, has three limit cycles around, for example, the origin (the case when limit cycles surround another focus is considered in a similar way). Then we get into some domain in the space of these parameters which is bounded by two fold bifurcation surfaces forming a cusp bifurcation surface of multiplicity-three limit cycles.

The corresponding maximal one-parameter family of multiplicity-three limit cycles cannot be cyclic, otherwise there will be at least one point corresponding to the limit cycle of multiplicity four (or even higher) in the parameter space. Extending the bifurcation curve of multiplicity-four limit cycles through this point and parameterizing the corresponding maximal one-parameter family of multiplicity-four limit cycles by a field rotation parameter, according to Theorem 4.5, we will obtain a monotonic curve which, by the Wintner–Perko termination principle (Theorem 4.4), terminates either at the origin or on some separatrix cycle surrounding the origin. Since we know absolutely precisely at least the cyclicity of the singular point (Theorem 4.1) which is equal to two, we have got a contradiction with the termination principle stating that the multiplicity of limit cycles cannot be higher than the multiplicity (cyclicity) of the singular point in which they terminate.

If the maximal one-parameter family of multiplicity-three limit cycles is not cyclic, on the same principle (Theorem 4.4), this again contradicts to Theorem 4.1 not admitting the multiplicity of limit cycles higher than two. Moreover, it also follows from the termination principle that either the ordinary separatrix loop or the eight-loop cannot have the multiplicity (cyclicity) higher than two (in that way, it can be proved that the cyclicity of three other separatrix cycles [13] is at most two). Therefore, according to the same principle, there are no more than two limit cycles in the exterior domain surrounding all three finite singularities of (4.1). Thus, system (4.1) cannot have either a multiplicity-three limit cycle or more than three limit cycles in any configuration. The theorem is proved. □

5 A piecewise linear dynamical system

Consider a Liénard-type dynamical system

$$\dot{x} = y - \varphi(x), \;\; \dot{y} = \beta - \alpha x - y, \;\; \alpha > 0, \;\; \beta > 0, \tag{5.1}$$

where $\varphi(x)$ is a piecewise linear function containing k dropping sections and approximating an arbitrary polynomial of degree $2k+1$. The line $\beta - \alpha x - y = 0$ and the curve $y = \varphi(x)$ can be considered as the isoclines of zero and infinity, respectively, for the corresponding equation. Such systems and equations may occur, for example, when tunnel diode circuits and some other problems are studied (see [1], [2], [6], [14]).

Suppose that the ascending sections of system (5.1) have an inclination $k_1 > 0$ and the descending (dropping) sections have an inclination $k_2 < 0$. Then the phase plane of (5.1) can be divided onto $2k+1$ parts in every of which (5.1) is a linear system: the ascending sections are in $k+1$ strip regions $(I, III, V, \ldots, 2K+1)$ and the descending sections are in other k such regions $(II, IV, VI, \ldots, 2K)$. The parameters k_1, k_2, and also α can be considered as rotation parameters for the sewed vector field of (5.1) (see [2], [10]).

System (5.1) can have an odd number of simple singular points: $1, 3, 5, \ldots, 2k+1$. If (5.1) has the only singular point, this point will be always an antisaddle (center, focus or node). A focus (node) will be always stable in odd regions and unstable in even regions if $k_2 > 1$. If system (5.1) has $2k+1$ singularities, then k of them are saddles (they are in even regions) and $k+1$ others are antisaddles (foci or nodes) which are always stable (they are in odd regions). The pieces of the straight lines $\beta = x_{2i-1}\alpha + y_{2i-1}$ and $\beta = x_{2i}\alpha + y_{2i}$ $(i = 1, 2, \ldots, k)$, where (x_{2i-1}, y_{2i-1}) and

(x_{2i}, y_{2i}) are the coordinates of the upper and lower corner points of the curve $\varphi(x)$, respectively, form a discriminant curve separating the domains in the plane (α, β), where $\alpha \le k_2$, with different numbers of singular points. The points of the discriminant curve correspond to the sewed singularities of saddle-focus or saddle-node type ($\alpha < k_2$) and its corner points correspond to the unstable equilibrium segments ($\alpha = k_2$) which coincide with the dropping sections of the curve $y = \varphi(x)$.

In the case when $k_2 < 1$, closed trajectories cannot exist and only bifurcations of singular points are possible in system (5.1). Therefore, we will consider further only the case when $k_2 > 1$ and $(k_1 - 1)^2 < 4k_2$ giving various bifurcations and, first of all, the bifurcations of limit cycles. Studying all such bifurcations (local and global), we will give a proof of the following theorem.

Theorem 5.1. *System* (5.1) *with k dropping sections and $2k + 1$ singular points can have at most $k + 2$ limit cycles, $k + 1$ of which surround the foci one by one and the last, $(k + 2)$-th, limit cycle surrounds all of the singular points of* (5.1).

Proof of Theorem 5.1. To prove the theorem, we will study both local and global bifurcations of limit cycles. The limit cycle of system (5.1) will be called *small* if it belongs to at most two adjoining regions; the cycle will be called *big* if it belongs to at least three adjoining regions.

5.1 Local bifurcations

Following [1], we will study first stability of the singular points on the line of sewing. Suppose that the straight line $\beta - \alpha x - y = 0$ passes through the corner point (x_1, y_1) of the curve $y = \varphi(x)$ on the boundary of regions I, II and that $\alpha > (k_2 + 1)^2/4$. Then the region I (II) will be filled by the pieces of trajectories of the stable (unstable) focus.

Introduce positive coordinates S_0 (lower (x_1, y_1)) and S_1 (upper (x_1, y_1)) on the line of sewing of regions I and II; S_2 (lower (x_2, y_2)) and S_3 (upper (x_2, y_2)) on the line of sewing of regions II and III, etc. The maps $S_0 \to S_1$ along the trajectories of region I and $S_1 \to S_0$ along the trajectories of region II are written as follows:

$$S_1 = S_0 e^{\pi\sigma_1/\omega_1}, \quad \bar{S}_0 = S_1 e^{\pi\sigma_2/\omega_2}, \tag{5.2}$$

where σ_i, ω_i $(i = 1, 2)$ are the real and imaginary parts of the roots of the characteristic equation for a singular point of regions I, II, respectively.

The singular point (x_1, y_1) will be a sewed center ($\bar{S}_0 = S_0$) iff $\sigma_1/\omega_1 + \sigma_2/\omega_2 = 0$, i. e., when $\alpha = \alpha^* \equiv (1 - k_1/k_2)/(k_2 - k_1 + 2)$. The sewed focus (x_1, y_1) will be stable ($\bar{S}_0 < S_0$) when $\alpha > \alpha^*$ and unstable ($\bar{S}_0 > S_0$) when $\alpha < \alpha^*$.

Consider the return map $S_0 \to \bar{S}_0$ along the trajectories of regions I and II. For region I, we will have

$$
\begin{aligned}
S_0 &= \frac{\delta_0}{\sin\omega_1\tau_1}(\omega_1\cos\omega_1\tau_1 - \sigma_1\sin\omega_1\tau_1 - \omega_1 e^{-\sigma_1\tau_1}) \equiv \delta_0\zeta(\tau_1),\\
S_1 &= \frac{\delta_0}{\sin\omega_1\tau_1}(\omega_1\cos\omega_1\tau_1 + \sigma_1\sin\omega_1\tau_1 - \omega_1 e^{\sigma_1\tau_1}) \equiv \delta_0\chi(\tau_1),
\end{aligned}
\tag{5.3}
$$

where δ_0 is the distance from the boundary of regions I, II to the singular point; ζ and χ are monotonic functions. The return map along the trajectories of region II has a similar form.

Calculation of the first derivative for the return map gives

$$
\frac{\mathrm{d}\bar{S}_0}{\mathrm{d}S_0} = \frac{S_0}{\bar{S}_0}\, e^{2(\sigma_1\tau_1+\sigma_2\tau_2)}, \tag{5.4}
$$

where τ_i $(i = 1, 2)$ is motion time along the trajectories of regions I, II, respectively; $\sigma_i = (1+k_i)/2$ $(i = 1, 2)$.

Studying the return map $S_0 \to \bar{S}_0$ by means of (5.4), we prove that at most one limit cycle can exist in regions I and II (see also [1]). The same result can be obtained for regions III and $IV, \ldots, 2K - 1$ and $2K$.

Consider now the map $\bar{S}_0 = f(S_0)$ sewed of two pieces: $\bar{S}_0 = \xi(S_0)$ along the trajectories in regions I, $II, \ldots, 2K$ and $\bar{S}_0 = \psi(S_0)$ along the trajectories in all regions, I, $II, \ldots, 2K$, $2K + 1$. The map $S_0 \to S_1$ in region I is given by (5.3). The maps $S_1 \to S_3$, $S_3 \to S_5, \ldots, S_{2k-1} \to S_{2k-2}$ $(S_{2k-1} \to S_{2k+1}$, $S_{2k+1} \to S_{2k}$, $S_{2k} \to S_{2k-2})$, $S_{2k-2} \to S_{2k-4}, \ldots, S_2 \to S_0$ have similar forms.

The derivatives for the functions $\xi(S_0)$, $\psi(S_0)$ are given by the following expressions, respectively:

$$
\frac{\mathrm{d}\bar{S}_0}{\mathrm{d}S_0} = \frac{S_0}{\bar{S}_0}\, e^{2(\sigma_1(\tau_1+\tau_3^{+}+\tau_3^{-}+\ldots+\tau_{2k-1})+\sigma_2(\tau_2^{+}+\tau_2^{-}+\ldots+\tau_{2k-2}^{+}+\tau_{2k-2}^{-}))}, \tag{5.5}
$$

$$
\frac{\mathrm{d}\bar{S}_0}{\mathrm{d}S_0} = \frac{S_0}{\bar{S}_0}\, e^{2(\sigma_1(\tau_1+\tau_3^{+}+\tau_3^{-}+\ldots+\tau_{2k+1})+\sigma_2(\tau_2^{+}+\tau_2^{-}+\ldots+\tau_{2k}^{+}+\tau_{2k}^{-}))}, \tag{5.6}
$$

where τ_1, τ_{2k-1}, τ_{2k+1} are motion times in regions I, $2K - 1$, $2K + 1$ and τ_{2i}^{+} (τ_{2i}^{-}), τ_{2i+1}^{+} (τ_{2i+1}^{-}), $i = 1, 2, \ldots, k$, are motion times in the upper (lower) parts of regions $II, III, \ldots, 2K$, respectively.

Studying the return map $\bar{S}_0 = f(S_0)$ by means of (5.5) and (5.6), we prove that at most two limit cycles can be generated by the boundary of the domain filled by closed trajectories of (5.1) and that these two limit cycles can be only outside the boundary.

Suppose that a part of the straight line $\beta - \alpha x - y = 0$ coincides with a dropping section of (5.1), for example, with the first one $(\alpha = k_2)$. The dropping section of (5.1) will be an unstable equilibrium segment and regions I, II (because of the condition $(k_1 - 1)^2 < 4k_2$) will be filled by trajectories of the stable foci. It is easy to obtain an explicit expression for the map of the half-line S_0 into itself:

$$
\bar{S}_0 = S_0\, e^{2\pi\sigma_1/\omega_1} + \delta(k_2 - 1)(1 + e^{\pi\sigma_1/\omega_1}), \tag{5.7}
$$

where δ is the width of regions II.

This map has the only stable fixed point, and we can show that two stable foci surrounded by unstable limit cycles (one by one) are generated from the ends of the equilibrium segment under the rotation of the line $\beta - \alpha x - y = 0$ (see also [1]).

The simplest type of separatrix cycles of (5.1) is a so-called eight-loop formed by two ordinary saddle loops. In the case of $2k+1$ simple singular points, a separatrix cycle can contain $k+1$ saddle loops, the first and the last of which are ordinary loops with one rough saddle on each and the $k-1$ others are separatrix digons with two rough saddles on each. Such a separatrix cycle will be called *nondegenerate.* In the cases when the straight line $\beta - \alpha x - y = 0$ passes through the corner points of the curve $y = \varphi(x)$, we will have *degenerate* separatrix cycles of lips-type containing one or two sewed saddle-nodes. It is clear that the bifurcations of separatrix cycles do not depend on the parameter β (see [1]). The separatrix cycles can be formed or destroyed only under a variation of the parameter α. The character of their stability will be determined by the sign of the saddle quantities which are always positive in our case, when the saddles are inside or on the boundary of even regions II, $IV, \ldots, 2K$ and $k_2 > 1$ (the corresponding theorems are valid for the piecewise linear dynamical systems as well [2]). It follows that the separatrix cycles of (5.1) are always unstable (inside and outside) and, under a variation of α, a nondegenerate separatrix cycle can generate at most $k+1$ small unstable limit cycles inside its loops (digons) or the only big unstable limit cycle outside it.

5.2 Global bifurcations

Now we are able to consider the global bifurcations of limit cycles. Suppose again that the zero isocline $\beta - \alpha x - y = 0$ passes through the corner point (x_1, y_1) of the infinite isocline $y = \varphi(x)$ and that $\alpha > \alpha^*$. In this case, the only singular point in the phase plane is a sewed stable focus and all trajectories of (5.1) tend to it when $t \to +\infty$. For decreasing α ($k_2 < \alpha < \alpha^*$), the sewed focus becomes unstable and a stable limit cycle is generated from the boundary curve of the domain filled by closed trajectories (immediately after passing the value α^* by the parameter α).

For $\alpha = k_2$, the first dropping section of (5.1) will coincide with a part of the straight line $\beta - \alpha x - y = 0$ and an unstable equilibrium segment will appear inside the stable limit cycle. If we rotate the line $\beta - \alpha x - y = 0$ around an interior point of the segment (changing both of the parameters, α and β), two unstable limit cycles surrounding stable foci (one by one) will be generated from the ends (x_1, y_1) and (x_2, y_2) of the equilibrium segment. Under the further rotation of the line $\beta - \alpha x - y = 0$, it will pass first through the next corner point, (x_4, y_4), and then, successively, through the points $(x_6, y_6), \ldots, (x_{2k}, y_{2k})$. Every time, the corner point becomes a sewed saddle-node generating an unstable limit cycle surrounding a stable focus. So, we will get a piecewise linear system with $2k+1$ singular points having at least $k+1$ small unstable limit cycles surrounding the stable foci (one by one) inside a big stable limit cycle, $k+2$, surrounding all of the singular points.

Under the further rotation of the zero isocline, all $k+1$ small limit cycles simultaneously disappear in a separatrix cycle consisting of $k+1$ loops (digons), this separatrix cycle generates a big (unstable) limit cycle which combines with another big (stable) limit cycle of (5.1) forming a semi-stable (double) limit cycle which finally disappears in a so-called trajectory condensation.

Let us prove that system (5.1) cannot have more than $k+2$ limit cycles. The proof is carried out by contradiction by means of the Wintner–Perko termination principle [2], [10], [18]. Since a small limit cycle is always unique in the corresponding strip regions, suppose that system (5.1) with three field rotation parameters, k_1, k_2, and α, has three big limit cycles. Then we get into some domain in the space of these parameters which is bounded by two fold bifurcation surfaces forming a cusp bifurcation surface of multiplicity-three limit cycles [10], [18].

The corresponding maximal one-parameter family of multiplicity-three limit cycles cannot be cyclic, otherwise there will be at least one point corresponding to the limit cycle of multiplicity four (or even higher) in the parameter space. Extending the bifurcation curve of multiplicity-four limit cycles through this point and parameterizing the corresponding maximal one-parameter family of multiplicity-four limit cycles by a field rotation parameter, for example, by the parameter α, we will obtain a monotonic curve which, by the Wintner–Perko termination principle, terminates either at the boundary curve of the domain filled by closed trajectories of (5.1) or on some degenerate separatrix cycle of (5.1) [10], [18].

Since we know at least the cyclicity of the boundary curve which is equal to two, we have got a contradiction with the termination principle stating that the multiplicity of limit cycles cannot be higher than the multiplicity (cyclicity) of the end bifurcation points in which they terminate [10], [18].

If the maximal one-parameter family of multiplicity-three limit cycles is not cyclic, using the same principle, this again contradicts with the cyclicity result for the boundary curve not admitting the multiplicity of limit cycles to be higher than two. Moreover, it also follows from the termination principle that the degenerate separatrix cycles of (5.1) cannot have the multiplicity (cyclicity) higher than two. Therefore, according to the same principle, there are no more than two big limit cycles in the exterior domain outside the boundary curve of (5.1).

The same results can be obtained by means of the new geometric methods developed in [12]. The phase portraits and bifurcation diagrams for system (5.1) will be similar to that which were constructed in [1], [2]. Thus, system (5.1) with $2k+1$ singular points cannot have more than $k+2$ limit cycles, i. e., $k+2$ is the maximum number of limit cycles of such system and the obtained distribution ($k+1$ small limit cycles plus a big limit cycle) is the only possibility for their distribution. The theorem is proved. □

Received: May 2008. Revised: June 2008.

References

[1] A.N. Bautin, *Qualitative investigation of a piecewise linear system*, J. Appl. Math. Mech., **38** (1974), 691–700.

[2] N.N. Bautin and E.A. Leontovich, *Methods and Examples of the Qualitative Analysis of Dynamical Systems in a Plane*, Nauka, Moscow, 1990 (in Russian).

[3] F. Botelho and V.A. Gaiko, *Global analysis of planar neural networks*, Nonlinear Anal., **64** (2006), 1002–1011.

[4] G.F.D. Duff, *Limit-cycles and rotated vector fields*, Ann. Math., **67** (1953), 15–31.

[5] F. Dumortier and C. Li, *Perturbation from an elliptic Hamiltonian of degree four: (IV) figure eight-loop*, J. Diff. Equat., **188** (2003), 512–554.

[6] A.F. Filippov, *Differential Equations with Discontinuous Right-Hand Sides*, Kluwer, Boston, 1988.

[7] V.A. Gaiko, *Qualitative theory of two-dimensional polynomial dynamical systems: problems, approaches, conjectures*, Nonlinear Anal., **30** (1997), 1385–1394.

[8] V.A. Gaiko, *Hilbert's sixteenth problem and global bifurcations of limit cycles*, Nonlinear Anal., **47** (2001), 4455–4466.

[9] V.A. Gaiko, *Global bifurcation families of multiple limit cycles in polynomial dynamical systems*, Nonlin. Phenom. Compl. Syst., **6** (2003), 734–745.

[10] V.A. Gaiko, *Global Bifurcation Theory and Hilbert's Sixteenth Problem*, Kluwer, Boston, 2003.

[11] V.A. Gaiko, *Wintner–Perko termination principle, parameters rotating a field, and limit-cycle problem*, J. Math. Sci., **126** (2005), 1259–1266.

[12] V.A. Gaiko, *Limit cycles of quadratic systems*, Nonlinear Anal., **69** (2008), 2150–2157.

[13] V.A. Gaiko and W.T. van Horssen, *Global bifurcations of limit and separatrix cycles in a generalized Liénard system*, Nonlin. Anal., **59** (2004), 189–198.

[14] F. Giannakopoulos and K. Pliete, *Planar systems of piecewise linear differential equations with a line of discontinuity*, Nonlinearity, **14** (2001), 1–22.

[15] I.D. Iliev and L.M. Perko, *Higher order bifurcations of limit cycles*, J. Diff. Equat., **154** (1999), 339–363.

[16] Yu. Ilyashenko, *Centennial history of Hilbert's 16th problem*, Bul. Amer. Math. Soc., **39** (2002), 301–354.

[17] A. Lins, W. de Melo, and C.C. Pugh, *On Liénard's equation*, In: Lecture Notes in Mathematics, Vol. 597, Springer, Berlin, 1977, pp. 335–357.

[18] L. Perko, *Differential Equations and Dynamical Systems*, Springer, New York, 2002.

[19] G.S. Rychkov, *The maximal number of limit cycles of the system* $\dot{y}=-x, \dot{x}=y-\sum_{i=0}^{2} a_i x^{2i+1}$ *is equal to two*, Diff. Equat., **11** (1975), 301–302.

[20] S. Smale, *Mathematical problems for the next century*, Math. Intellig., **20** (1998), 7–15.

CUBO A Mathematical Journal
Vol.10, N$^{\underline{o}}$03, (133–136). October 2008

The Fibonacci Zeta-Function is Hypertranscendental

Jörn Steuding
Department of Mathematics, Würzburg University,
Am Hubland, 97218 Würzburg, Germany
email: `steuding@mathematik.uni-wuerzburg.de`

ABSTRACT

Applying a theorem of Reich on Dirichlet series satisfying difference-differential equations, we show that the Fibonacci zeta-function satisfies no algebraic differential equation.

RESUMEN

Aplicando el Teorema de Reich sobre series de Dirichlet satisfaziendo ecuaciones diferenciales-diferencias, nosotros mostramos que la función zeta de Fibonacci satisfaze una ecuación diferencial no algebraica.

Key words and phrases: *Hypertranscendence, Fibonacci zeta-function.*

Math. Subj. Class.: *11B39, 11M41, 34M15.*

1 Introduction

The Fibonacci numbers are recursively defined by

$$F_0 = 0,\ F_1 = 1 \qquad \text{and} \qquad F_{n+2} = F_{n+1} + F_n \quad \text{for} \quad n \in \mathbb{N}.$$

In number theory one can often obtain arithmetic information by studying a generating function of a given number theoretical object. In the case of Fibonacci numbers this is usually the corresponding Lambert series; however, in the recent past also the generating Dirichlet series was studied; this function is more interesting with respect to its analytic properties. Let s be a complex variable. For $\operatorname{Re} s > 0$ the Fibonacci zeta-function is defined by

$$\zeta_{\mathsf{F}}(s) = \sum_{n\in\mathbb{N}} F_n^{-s},$$

and by analytic continuation throughout the complex plane except for simple poles at $s = -2k + \pi i(2n+k)/\log\varphi$ for $n \in \mathbb{Z}, k \in \mathbb{N}_0$, where φ is the golden ratio; this was first proved by Navas [6] (and relies mainly on Binet's formula). In [1], Elsner et al. obtained several results on the algebraic independence of the values taken by ζ_{F} on the positive integers, e.g. $\zeta_{\mathsf{F}}(2), \zeta_{\mathsf{F}}(4), \zeta_{\mathsf{F}}(6)$ are algebraically independent.

In this note we show that the Fibonacci zeta-function $\zeta_{\mathsf{F}}(s)$ is hypertranscendental, i.e., it satisfies no non-trivial algebraic differential equation (that is no finite collection of derivatives of ζ_{F}) is algebraically dependent over the field of rational functions). Actually, we shall prove a slightly stronger statement by applying Reich's theorem on Dirichlet series satisfying holomorphic difference–differential equations. In order to state this result denote by $\mathcal{D}$ the set of all ordinary Dirichlet series $f(s) = \sum_{n=1}^{\infty} a_n n^{-s}$ satisfying the following two assumptions:

- the abscissa of absoulte convergence is finite: $\sigma_a(f) < \infty$,
- the set of all divisors of indices n with $a_n \neq 0$ contains infinitely many prime numbers.

Furthermore, we introduce the following abbreviation: for a non-negative integer ν we write

$$\underline{f}^{[\nu]}(s) = (f(s), f'(s), \ldots, f^{(\nu)}(s)).$$

Reich [9] proved the following theorem: *Assume that* $f \in \mathcal{D}$. *Let* $h_0 < h_1 < \ldots < h_m$ *be any real numbers,* $\nu_0, \nu_1, \ldots, \nu_m$ *be any non-negative integers, and let* $\sigma_0 > \sigma_a(f) - h_0$. *Put* $k := \sum_{j=0}^{m} (\nu_j + 1)$. *If* $\Phi : \mathbb{C}^k \to \mathbb{C}$ *is continuous and the difference-differential equation*

$$\Phi(\underline{f}^{[\nu_0]}(s+h_0), \underline{f}^{[\nu_1]}(s+h_1), \ldots, \underline{f}^{[\nu_m]}(s+h_m)) = 0$$

holds for all s *with* $\operatorname{Re} s > \sigma_0$*, then* Φ *vanishes identically.* To apply this result to the Fibonacci zeta-function it suffices to show that the set of all Fibonacci numbers F_n is not generated by finitely many primes. However, this follows immediately from Lucas' theorem

$$\gcd(F_m, F_n) = F_{\gcd(m,n)}$$

(see [5]), since the right-hand side is equal to $F_1 = 1$ for any pair m, n of relatively coprime integers. Thus we obtain:

Theorem 1. *Given any real numbers $h_0 < h_1 < \ldots < h_m$, any non-negative integers $\nu_0, \nu_1, \ldots, \nu_m$, and any $\sigma_0 > -h_0$, if $\Phi : \mathbb{C}^k \to \mathbb{C}$ is continuous and the difference-differential equation*

$$\Phi(\underline{\zeta_{\mathsf{F}}}^{[\nu_0]}(s+h_0), \underline{\zeta_{\mathsf{F}}}^{[\nu_1]}(s+h_1), \ldots, \underline{\zeta_{\mathsf{F}}}^{[\nu_m]}(s+h_m)) = 0$$

holds for all s with $\operatorname{Re} s > \sigma_0$, then Φ vanishes identically.

Notice that the proof does not use the meromorphic continuation of $\zeta_{\mathsf{F}}(s)$ to $\mathbb{C}$, obtained by Navas. The statement of the theorem can easily be extended to other Dirichlet series built from linear recursive sequences. Here we only need that such sequences are divisible by infinitely many prime numbers which is true except for *degenerate* cases when the characteristic polynomial has two roots whose quotient is a root of unity; since roots are counted with multiplicities, this also includes the case of repeated roots. This was first shown by Pólya [8] and has been rediscovered by several mathematicians (see [2, 10] for some history).

We conclude with a few historical remarks on hypertranscendence and an interesting question. In 1887, Hölder [4] proved that the Gamma-function is hypertranscendental. In his challenging lecture at the International Congress for Mathematicians in Paris 1900, Hilbert [3] asked in problem 18 for a description of classes of functions definable by differential equations. In this context Hilbert stated that the Riemann zeta-function $\zeta(s)$ is hypertranscendental; the first published proof was written down by Stadigh in his dissertation (cf. Ostrowski [7]). The idea is to deduce the hypertranscendence of $\zeta(s)$ from Hölder's theorem and the fact that the Gamma-function appears in the functional equation for zeta. Besides, Hilbert [3] asked for a proof of the hypertranscendence for the more general series $\sum_{n=1}^{\infty} x^n n^{-s}$. This problem was solved by Ostrowski [7] as a particular case of a more general theorem which also applies to the case when there is no functional equation at hand; his argument relies on a comparison of the differential independence with the linear independence of its frequencies. Reich's theorem [9], which we have used to prove Theorem 1, may be regarded as the most general and powerful extension of this method. A different way for proving hypertranscendence was found by Voronin. In [11], he developped a new technique to study the joint value distribution of Dirichlet L-functions to pairwise inequivalent characters and their derivatives; in [12], he extended the method in order to prove his famous universality theorem for the Riemann zeta-function: *Let $0 < r < \frac{1}{4}$ and suppose that $g(s)$ is a non-vanishing continuous function on the disk $|s| \leq r$ which is analytic in the interior. Then, for any $\epsilon > 0$,*

$$\liminf_{T\to\infty} \frac{1}{T} \operatorname{meas} \left\{ \tau \in [0,T] \,:\, \max_{|s|\leq r} \left| \zeta\left(s + \tfrac{3}{4} + i\tau\right) - g(s) \right| < \epsilon \right\} > 0.$$

Voronin's results imply the hypertranscendence of these Dirichlet series. It is natural to ask whether the Fibonacci zeta-function shares this or some other universality property: *is it true or false that any (suitable) analytic function $g(s)$ can be uniformly approximated by certain shifts of $\zeta_{\mathsf{F}}(s)$?*

Acknowledgements. The author wants to thank the anonymous referee for her or his remarks and references to clarify to which extent the statement of the theorem can be generalized.

Received: April 2008. Revised: June 2008.

References

[1] C. Elsner, S. Shimomura and I. Shiokawa, *Algebraic relations for reciprocal sums of Fibonacci numbers*, Acta Arith., **130** (2007), 37–60.

[2] G. Everest, A. van der Poorten, I. Shparlinski and T. Ward, *Recurrence sequences*, AMS Mathematical Surveys and Monographs, **104** (2003).

[3] D. Hilbert, *Mathematische Probleme*, Archiv f. Math. u. Physik, **1** (1901), 44–63, 213–317.

[4] O. Hölder, *Über die Eigenschaft der Gammafunktion keiner algebraischen Differentialgleichung zu genügen*, Math. Ann., **28** (1887), 1–13.

[5] T. Koshy, *Fibonacci and Lucas Numbers with applications*, John Wiley & Sons, New York 2001.

[6] L. Navas, *Analytic continuation of the Fibonacci Dirichlet series*, Fibonacci Q., **39** (2001), 409–418.

[7] A. Ostrowski, *Über Dirichletsche Reihen und algebraische Differentialgleichungen*, Math. Z., **8** (1920), 241–298.

[8] G. Pólya, *Arithmetische Eigenschaften der Reihenentwicklungen rationaler Funktionen*, J. Reine Angew. Math., **151** (1921), 1–31.

[9] A. Reich, *Über Dirichletsche Reihen und holomorphe Differentialgleichungen*, Analysis, **4** (1984), 27–44.

[10] H. Roskam, *Prime divisors of linear recurrences and Artin's primitive root conjecture for number fields*, J. Théo. Nombres Bordeaux, **13** (2001), 303–314.

[11] S.M. Voronin, *On the functional independence of Dirichlet L-functions*, Acta Arith., **27** (1975), 493–503 (in Russian).

[12] S.M. Voronin, *Theorem on the 'universality' of the Riemann zeta-function*, Izv. Akad. Nauk SSSR, Ser. Matem., **39** (1975) (in Russian); engl. translation in Math. USSR Izv., **9** (1975), 443–445.

CUBO A Mathematical Journal
Vol.10, N^{o}03, (137–144). October 2008

On Some Bitopological γ-Separation Axioms

M. Lellis Thivagar
Department of Mathematics, Arul Anandar College,
Karumathur, Madurai (Dt.)-625514, Tamilnadu, India
email: mlthivagar@yahoo.co.in

S. Athisaya Ponmani
Department of Mathematics, Jayaraj Annapackiam,
College for Women, Periyakulam, Theni (Dt.)-625601, Tamilnadu, India
email: athisayaponmani@yahoo.co.in

R. Raja Rajeswari
Department of Mathematics, Sri Parasakthi College,
Courtalam, Tirunelveli (Dt.)-627802, Tamilnadu, India
email: raji_arul2000@yahoo.co.in

and

Erdal Ekici
Department of Mathematics, Canakkale Onsekiz Mart University,
Terzioglu Campus, 17020 Canakkale, Turkey
email: e.mail:eekici@comu.edu.tr

ABSTRACT

The aim of this paper is to introduce the notions of (i,j)-γ-T_1, $(i,j)\gamma$-R_1, (i,j)-γ-T_2 and (i,j)-γ-US spaces, (i,j)-γ-open mappings and (i,j)-γ-irresolute mappings.

RESUMEN

El objetivo de este artículo es introducir las nociones de espacios (i,j)-γ-T_1, $(i,j)\gamma$-R_1, (i,j)-γ-T_2 y (i,j)-γ-US, aplicaciones (i,j)-abiertas y (i,j)-γ-irresolutas.

Key words and phrases: *(i,j)-γ-open set, (i,j)-γ-T_1 space, (i,j)-γ-R_1space, (i,j)-γ-US space, (i,j)-γ-open mapping, and (i,j)-γ-irresolute mapping.*

Math. Subj. Class.: *54C55.*

1 Introduction

In 1982, Mashhour et al. [11] introduced the notion of preopen sets, also called locally dense sets by Corson and Michael [4]. The class of preopen sets properly contains the class of open sets. As the intersection of two preopen sets may fail to be preopen, the class of preopen sets does not always form a topology. In a submaximal space i.e. a topological space X in which every dense subset is open, collection of all preopen sets form a topology. Indeed, many notions in Topology can be defined in terms of preopen sets (see [3], [5], [8], [12] and [13]). In 1987, Andrijevic [2] offered a new class of open sets called γ-open sets by utilizing preopen sets. Recently, Abd El Monsef et al. [1] have applied preopen sets in connection with the topological applications of rough set measures in information systems. Moreover, it has been shown in [6] that the notion preopen sets is important with respect to the digital topology. Many researchers also used the notion of preopen sets in fuzzy topological spaces which Professor El-Naschie has recently shown in [7] the importance of the notion of fuzzy topology which may be relevent to quantam particle physics in connection with string theory and ϵ^{∞}theory.

In a bitopological space (X, τ_1, τ_2), the γ-open set is generalized in the form of (i,j)-γ-open set, i, $j = 1, 2$ and $i \neq j$ [14] and these sets are used to define the separation axiom (i,j)-γ-T_0 [14].

In this paper we define $(i,j)\gamma$-T_1, (i,j)-γ-R_1, (i,j)-γ-T_2 and (i,j)-γ-US spaces and show that (i,j)-γ-US axiom is stronger than (i,j)-γ-T_1 axiom and is weaker than (i,j)-γ-T_2 axiom.

We recall some definitions and concepts which are useful in the following sections.

2 Preliminaries

In a topological space (X, τ), the interior and the closure of a subset A are denoted by $int(A)$ and $cl(A)$, respectively.

Definition 1 *A subset A of X is called pre-open set [11] if $A \subset int(cl(A))$.*

Definition 2 *A subset A of a topological space (X, τ) is called a γ-set [2] if $A \cap S \in PO(X)$ for each $S \in PO(X)$.*

In the above definition, $PO(X)$ is the family of all pre-open sets in X. The family of all γ-sets in X is denoted by $\gamma O(X)$.

In the following sections by a space X, we mean a bitopological space (X, τ_1, τ_2).

Definition 3 *A subset A of X is called (i, j)-pre-open [9] if $A \subset \tau_i$-int$(\tau_j$-cl$(A))$.*

Definition 4 *A subset A of X is called (i, j)-γ-open [14], if $A \cap B$ is (i, j)-pre-open for every (i, j)-pre-open set B in X.*

We denote the family of (i, j)-γ-open sets in X by (i, j)-$\gamma O(X)$.

Theorem 5 *[14] The family of all (i, j)-γ-open sets in X forms a topology on X.*

Definition 6 *A subset $A \subset X$ is called (i, j)-γ-closed [14] if its complement, A^c in X is (i, j)-γ-open.*

Definition 7 *For any $A \subset X$*

(i) (i, j)-γ-closure of A [14] is the intersection of all the (i, j)-γ-closed sets containing A and is written as (i, j)-γ-cl(A).

(ii) (i, j)-γ-kernal of A [14] is the intersection of all the (i, j)-γ-open sets containing A and is written as (i, j)-γ-ker(A).

Definition 8 *A space X is called (i, j)-γ-T_0[14] if for $x,\ y \in X,\ x \neq y$, there exists $U \in (i, j)$-$\gamma O(X)$ such that U contains only one of x and y but not the other where $i,\ j = 1, 2,\ i \neq j$.*

Definition 9 *A map $f{:}(X, \tau_1, \tau_2) \rightarrow (Y, \sigma_1, \sigma_2)$ is called pairwise γ-continuous (briefly p.γ-continuous)[14] if the inverse image of each σ_i-open set of Y is (i, j)-γ-open in X for $i,\ j = 1, 2$ and $i \neq j$.*

In the following section the (i, j)-γ-open sets are used to define some separation axioms.

3 Some separation axioms

In this section we define the (i,j)-γ-T_1, (i,j)-γ-R_1, (i,j)-γ-T_2 and (i,j)-γ-US spaces and study some characterizations.

Definition 10 *A space X is called (i,j)-γ-T_1 if for x, y in X, $x \neq y$, there exist U, $V \in (i,j)$-$\gamma O(X)$ such that $x \in U$, $y \notin V$ and $y \in V$, $x \notin V$.*

Definition 11 *A space X is said to be (i,j)-γ-R_1 if for x, y in X, $x \neq y$ with (i,j)-γ-$cl(\{x\}) \neq (i,j)$-$\gamma cl(\{y\})$, there exist disjoint (i,j)-γ-open sets U, V such that (i,j)-γ-$cl(\{x\}) \subset U$ and (i,j)-γ-$cl(\{y\}) \subset V$.*

Theorem 12 *A space X is (i,j)-γ-T_1 if and only if the singletons in X are (i,j)-γ-closed sets.*

Proof. Proof is evident since the family (i,j)-$\gamma O(X)$ is a topology. ■

Theorem 13 *A space X is (i,j)-γ-R_1 if and only if (i,j)-γ-$ker(\{x\}) \neq (i,j)$-γ-$ker(\{y\})$ for any x, y in X , there exist disjoint (i,j)-γ-open sets U and V such that γ-$cl(\{x\}) \subset U$ and γ-$cl(\{y\}) \subset V$.*

Definition 14 *A space X is said to be (i,j)-γ-T_2 if for any two distinct points x, y in X, there exist disjoint (i,j)-γ-open sets U, V such that $x \in U$ and $y \in V$.*

Theorem 15 *A space X is (i,j)-γ-T_2 if and only if it is (i,j)-γ-T_0 and (i,j)-γ-R_1.*

Proof. Necessity. If X is (i,j)-γ-T_2 then it is (i,j)-γ-T_1 and then (i,j)-γ-T_0. Since X is (i,j)-γ-T_1, by Theorem 12, (i,j)-γ-$cl(\{x\}) = \{x\}$ and (i,j)-γ-$cl(\{y\}) = \{y\}$ for any two distinct points x, y in X. Therefore, (i,j)-γ-$cl(\{x\}) \neq (i,j)$-γ-$cl(\{y\})$ for any two distinct points x, y in X and hence X is (i,j)-γ-R_1.

Sufficiency. If X is (i,j)-γ-T_0 and if x, y are two distinct points in X, there exists an (i,j)-γ-open set U containing only one of x and y but not the other. Let $x \in U$ and $y \notin U$, say. Then $y \notin (i,j)$-γ-$ker(\{x\})$ and so (i,j)-γ-$ker(\{x\}) \neq (i,j)\gamma$-$ker(\{y\})$ for any two distinct points x, y in X. Since X is (i,j)-γ-R_1, by Theorem 13, there exist disjoint (i,j)-γ-open sets U and V such that (i,j)-$\gamma cl(\{x\}) \subset U$ and (i,j)-γ-$cl(\{y\}) \subset V$. Thus $x \in U$ and $y \in V$ and $U \cap V = \emptyset$. Hence X is (i,j)-γ-T_2. ■

Definition 16 *A net $\{x_\alpha{:}\alpha \in D, \geq\}$ is said to be bitopologically converges to a point $x \in X$, denoted by $\{x_\alpha{:}\alpha \in D, \geq\} \xrightarrow{\gamma} x$ if the net is eventually in every (i,j)-γ-open set containing x, i, $j = 1,2$, $i \neq j$.*

Theorem 17 *If a map $f:(X,\tau_1,\tau_2) \to (Y,\sigma_1,\sigma_2)$ is p.γ-continuous then for each $x \in X$ and each net $\{x_\alpha:\alpha \in D, \geq\}$ in X, bitopologically γ-converging to x the image net $\{f(x_\alpha):\alpha \in D, \geq\}$ is bitopologically γ-convergent to $f(x)$ in Y.*

Proof. Let $V \subset Y$ be σ_i-open in Y containing $f(x)$, $i = 1, 2$. The bitopologically γ-convergence of the net $\{x_\alpha:\alpha \in D, \geq\}$ in X implies that there exists $\alpha_0 \in D$ such that for all $\alpha \geq \alpha_0$, $x_\alpha \in f^{-1}(V)$. Therefore, $f(x_\alpha) \in V$ for all $\alpha \geq \alpha_0$. Hence the net $\{f(x_\alpha):\alpha \in D, \geq\} \xrightarrow{\gamma} f(x)$. ■

Definition 18 *A space X is said to be (i,j)-γ-US if every bitopologically γ-convergent net $\{x_\alpha:\alpha \in D, \geq\}$ in X is bitopologically γ-convergent to a unique pioint in X.*

Proposition 19 *Every $(i,j)\gamma$-T_2 space is (i,j)-γ-US.*

Proof. If possible, let the net $\{x_\alpha:\alpha \in D, \geq\}$ in a $(i,j)\gamma$-T_2 space X be bitopologically γ-convergent to two distinct points x, y in X. Then the net is eventually in every (i,j)-γ-open set containing x and also in every (i,j)-γ-open set containing y. This contradicts that X is (i,j)-γ-T_2. ■

Proposition 20 *Every (i,j)-γ-US space is (i,j)-γ-T_1.*

Proof. Let x, $y \in X$, $x \neq y$. If $x_n = x$ for every x in the net $\{x_\alpha:\alpha \in D, \geq\}$ then it is evident that the net is bitopologically γ-convergent to x. Since X is (i,j)-γ-US, the net $\{x_\alpha:\alpha \in D, \geq\}$ cannot be bitopologically γ-convergent to y and hence there exists an (i,j)-γ-open set containing y but not x. A similar argument gives an (i,j)-γ-open set containing x but not y. Hence X is (i,j)-γ-T_1. ■

Remark 21 *The following diagram holds for a space X as shown in the Proposition 19 and 20.*

$$(i,j)\text{-}\gamma\text{-}T_2 \textit{ space} \Rightarrow (i,j)\text{-}\gamma\text{-}US \textit{ space} \Rightarrow (i,j)\text{-}\gamma\text{-}T_1 \textit{ space}$$

Theorem 22 *A space X is (i,j)-γ-T_2 if and only if it is (i,j)-γ-R_1 and (i,j)-γ-US.*

Proof. If X is (i,j)-γ-T_2, then it is (i,j)-γ-R_1, by Theorem 15 and by Proposition 19, X is (i,j)-γ-US.

Conversely, if X is (i,j)-γ-R_1 and (i,j)-γ-US then by Proposition 20, X is (i,j)-γ-T_1. Thus X is $(i,j)\gamma$-T_1 and (i,j)-γ-R_1. Hence by Theorem 15, X is (i,j)-γ-T_2. ■

4 Some bitopolgical γ- mappings

In this section we define (i, j)-γ-open mappings and (i, j)-γ-irresolute mappings.

Definition 23 *A map $f{:}(X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ is called (i, j)-γ-open if the image of each τ_i-open set in X is (i, j)-γ-open in Y, $i, j = 1, 2$, $i \neq j$.*

Recall that a map $f{:}(X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ is called (i, j)-pre-open if for each τ_i-open set in X, $f(U)$ is (i, j)-pre-open, $i, j = 1, 2$, $i \neq j$.

Remark 24 *Every (i, j)-γ-open map is (i, j)-pre-open but the converse is not true in general as shown in the following example.*

Example 25 *Let $X = \{a, b, c\}$, $\tau_1 = \{\emptyset, \{b, c\}, X\}$, $Y = \{a, b, c\}$, $\sigma_1 = \{\emptyset, \{a\}, X\}$ and $\sigma_2 = \{\emptyset, \{a, b\}, X\}$. Define a map $f{:}X \to Y$ as follows $f(a) = b$, $f(b) = a$, $f(c) = c$. Then f is $(1, 2)$-pre-open but not $(1, 2)$-γ-open since $f(\{b, c\}) = \{a, c\}$ which is $(1, 2)$-pre-open but not $(1, 2)$-γ-open.*

Recall that a space X is said to be pairwise Hausdorff[12] if for x, $y \in X$, $x \neq y$, there exist open sets U, V, $U \in \tau_1$, $V \in \tau_2$ such that $x \in U$ and $y \in V$.

Theorem 26 *Let $f{:}(X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ be a bijective (i, j)-γ-open map.If X be pairwise Hausdorff, then Y is (i, j)-γ-T_2.*

Proof. Let y_1 and y_2 be two distinct points in Y. Since f is bijective there exist x_1 and x_2 in X such that $f(x_1) \neq f(x_2)$. The space X is pairwise Hausdorff and so there exist disjoint sets U, V, $U \in \tau_1$ and $V \in \tau_2$ such that $x_1 \in U$and $x_2 \in V$. Then $f(x_1) \in f(U)$ and $f(x_2) \in f(V)$, $f(U)$ and $f(V)$ are (i, j)-γ-open sets and $f(U) \cap f(V) = \emptyset$. Thus Y is (i, j)-γ-T_2. ■

Definition 27 *A map $f{:}(X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ is called (i, j)-γ-irresolute if the inverse image of every (i, j)-γ-open set in Y is (i, j)-γ-open in X, $i, j = 1, 2$, $i \neq j$.*

Theorem 28 *If $f{:}(X, \tau_1, \tau_2) \to (Y, \sigma_1, \sigma_2)$ is an (i, j)-γ-irresolute bijective mapping and if Y is a (i, j)-γ-T_2 space then, X is (i, j)-γ-T_2.*

Proof. Let x_1, x_2 be two distinct points in X. Then there exist y_1, y_2 in Y such that $f(x_1) = y_1$ and $f(x_2) = y_2$ and $y_1 \neq y_2$. Since Y is (i, j)-γ-T_2, there exist disjoint (i, j)-γ-open sets U, V such that $y_1 \in U$ and $y_2 \in V$. As f is (i, j)-γ-irresolute, $f^{-1}(U)$ and $f^{-1}(V)$ are (i, j)-γ-open sets in X containing x_1 and x_2 respectively. Hence X is (i, j)-γ-T_2. ■

Theorem 29 *Let* $f{:}(X,\tau_1,\tau_2) \rightarrow (Y,\sigma_1,\sigma_2)$ *and* $g{:}(Y,\sigma_1,\sigma_2) \rightarrow (Z,\varrho_1,\varrho_2)$ *be two maps.Then*

(i) If f *is* (i,j)*-*γ*-irresolute and* g *is p.*γ*-continuous then,* $g \circ f$ *is p.*γ*-continuous*

(ii) If both f *and* g *are* (i,j)*-*γ*-irresolute then,* $g \circ f$ *is* (i,j)*-*γ*-irresolute.*

Proof. Obvious. ∎

ACKNOWLEDGEMENT

The second and third authors wish to acknowledge the support of University Grants Commission, New Delhi, India, FIP-X Plan.

Received: June 2008. Revised: August 2008.

References

[1] M.E. Abd El-Monsef, B.M. Taher and A.S. Salama, *Topological applications of rough set measures in information systems*, Preprint.

[2] D. Andrijevic, *On the Topology Generated by Pre-open sets*, Mate. Bech., **39** (1987), 367–376.

[3] A.V. Arhangel'Skii and P.J. Collins, *On submaximal spaces*, Topology Appl., **64** (1995), 219–241.

[4] H. Corson and E. Michael, *Metrizability of certain countable unions*, Illinois J. Math., **8** (1964), 351–360.

[5] S.N. El-Deeb, A. Hasanein, A.S. Mashhour and T. NOIRI, *On p-regular spaces*, Bull. Math. Soc. Sci. Math., R.S. Roumanie(N.S), **27** (75)(1983), 65-73.

[6] R. Devi, K. Bhuvaneswari and H. Maki, *Weak forms of g*ϱ*-closed sets, where* $\varrho \in \{\alpha,\alpha^\star,\alpha^{\star\star}\}$*, and digital plane*, Mem. Fac. Sci. Kochi Univ. Ser. A. Math., **25** (2004), 37–54.

[7] M.S. EL-NASCHIE, *On the certification of heterotic strings, M theory and* ϵ^∞ *theory*, Chaos, Solitons and Fractals, (2000), 2397–2408.

[8] J. Foran and P. Liebnitz, *A characterization of almost resolvable spaces*, Rend. Circ. Mat. Palermo, Serie II, Tomo XL, (1991), 136–141.

[9] A. Kar and P. Bhattacharyya, *Bitopological preopen sets, precontinuity and preopen mappings*, Indian J. Math.**34** (1992), 295–309.

[10] J.C. KELLY, *Bitopological Spaces*, Proc. Londo Math. Soc., **13** (1963), 71–89.

[11] A.S. Mashhour, I.A. Hasanein and S.N. El-Deeb, *On precontinuous and weak precontinuous mappings*, Proc. Math. Phys. Soc. Egypt, **53** (1982), 47–53.

[12] A.S. Mashhour, M.E. Abd El-Monsef, I.A. Hasanein and T. Noiri, *Strongly compact spaces*, Delta J. Sci., **8** (1984), 30–46.

[13] M.G. Murudeshwar and S.A. Naimpally, *Semi-Hausdorff spaces*, Canad. Math. Bull., **9** (1996), 353–356.

[14] Saeid Jafari, M. Lellis Thivagar, S. Athisaya Ponmani and R. Raja Rajeswari, *A Bitopological View of γ-open sets*, Communicated.

CUBO A Mathematical Journal
Vol.10, N$^{\underline{o}}$03, (145–159). October 2008

Asymptotic Constancy and Stability in Nonautonomous Stochastic Differential Equations

John A.D. Appleby*
School of Mathematical Sciences, Dublin City University,
Dublin 9, Ireland
email: `john.appleby@dcu.ie`
URL: `http://webpages.dcu.ie/~applebyj`

James P. Gleeson
Department of Applied Mathematics, University College Cork,
Cork, Ireland
email: `j.gleeson@ucc.ie`

and

Alexandra Rodkina
Department of Mathematics and Computer Science,
The University of the West Indies, Mona, Kingston 7, Jamaica
email: `alexandra.rodkina@uwimona.edu.jm`

ABSTRACT

This paper considers the asymptotic behaviour of a scalar non-autonomous stochastic differential equation which has zero drift, and whose diffusion term is a product of a function of time and space dependent function, and which has zero as a unique

*The first author is partially supported by an Albert College Fellowship awarded by Dublin City University's Research Advisory Panel. The third author is supported by the Mona Research Fellowship Programm awarded by the University of the West Indies, Jamaica and by London Mathematical Society

equilibrium solution. We classify the pathwise limiting behaviour of solutions; solution either tends to a non-trivial, non-equilibrium and random limit, or the solution hits zero in finite time. In the first case, the exact rate of decay can always be computed. These results can be inferred from the square integrability of the time dependent factor, and the asymptotic behaviour of the corresponding autonomous stochastic equation, where the time dependent multiplier is unity.

RESUMEN

Este artículo considera el comportamiento asintótico de una ecuación diferencial estocastica escalar no-autónoma la cual tiene cero desviación y cujo término de difusión es un producto de una función de equilibrio. Nosotros clasificamos el comportamiento limite por caminos de las soluciones; la solución atiende a un no equilibrio y límite random no trivial, o la solución encuentra cero en tiempo finito. En el primer caso, las tasas de decaimiento siempre pueden ser calculadas. Estos resultados pueden ser inferidos de la integrabilidad al cuadrado del factor dependiente del tiempo, y el comportamiento asintótico de la correspondiente ecuación estocatica autónoma, donde el multiplicador dependiente del tiempo es la unidad.

Key words and phrases: *Brownian motion, almost sure asymptotic stability, asymptotic constancy, stochastic differential equation, nonautonomous, Feller's test, explosions.*

Math. Subj. Class.: *60H10, 93E15.*

1 Introduction

This note considers the asymptotic behaviour of solutions of the "separable" stochastic differential equation

$$dX(t) = \sigma(t)g(X(t))\,dB(t). \tag{1.1}$$

A solution of this equation with initial condition ξ is denoted by $X(\cdot,\xi)$. It is presumed that zero is a point equilibrium, so $X(t,0)=0$ is a solution of (1.1). A standard deterministic change of time scale reduces this equation to an autonomous equation

$$d\tilde{X}(t) = g(\tilde{X}(t))\,d\tilde{B}(t), \tag{1.2}$$

from which it can be shown that the condition that $\sigma \in L^2([0,\infty);\mathbb{R})$ largely determines whether the solution tends to the equilibrium or to a non–trivial and non–equilibrium limit. Another feature which is examined is the relationship between the process $\tilde{X}$ hitting zero in a finite amount of time, or tending to zero as $t\to\infty$ (in the case when $\tilde{X}$ remains strictly positive) and the corresponding properties of X. As will be seen, a complete picture of the dynamics of (1.1) can be deduced in terms of conditions on g and σ.

Even though the time–change technique employed is well–known, some novel features appear in the analysis. First, we are unaware of an extensive literature concerning the pathwise convergence of solutions of stochastic differential equations to non–equilibrium limits. Second, we determine here sharp upper and lower estimates in terms of the rate at which the noise intensity fades on the almost sure rate of convergence of the solution to this non–equilibrium limit, in the case when $\sigma \in L^2([0,\infty);\mathbb{R})$. This requires a delicate use of the law of the iterated logarithm, partly correcting an error on the asymptotic behaviour of a tail martingale established in [1] and used in [2]. Finally, in the case where $\sigma \notin L^2([0,\infty);\mathbb{R})$, the results here, taken in conjunction with work in [3, 4] would enable exact almost sure rates of convergence to zero of solutions of (1.1) to be established.

2 Existence of solutions

In this paper we deal with highly nonlinear stochastic differential equations, SDE, whose solutions can hit zero at finite time, due to the non-Lipshitz behavior of the diffusion coefficients. Moreover, for non-autonomous equations, it is convenient for the completeness of our exposition, to state carefully and to prove an existence result. This result is a corollary of well-known existence result and martingale time changing theorem.

Let $(\Omega,\mathcal{F},(\mathcal{F}(t))_{t\geq 0},\mathbb{P})$ be a complete probability space with filtration $(\mathcal{F}(t))_{t\geq 0}$ satisfying the usual conditions (i.e. it is increasing and right continuous while $\mathcal{F}(0)$ contains all $\mathbb{P}$-null sets). Let $(B(t))_{t\geq 0}$ be a scalar standard Brownian motion, defined on the probability space $(\Omega,\mathcal{F},(\mathcal{F}(t))_{t\geq 0},\mathbb{P})$. Since we will consider equations with deterministic initial conditions, it is enough to work with the natural filtration of B: that is $\mathcal{F}(t)\equiv\mathcal{F}^B(t)$ where $\mathcal{F}^B(t)=\sigma(\tilde{B}(s):0\leq s\leq t)$.

Suppose that the function σ obeys

$$\sigma \in C([0,\infty);\mathbb{R}). \tag{2.1}$$

We define the local martingale $M=\{M(t),0\leq t<\infty,\mathcal{F}^B(t)\}$ by

$$M(t)=\int_0^t \sigma(s)\,dB(s), \quad t\geq 0, \tag{2.2}$$

with square variation $\langle M\rangle$ given by

$$\langle M\rangle(t)=\int_0^t \sigma^2(s)\,ds.$$

Let T^* be given by

$$T^*=\int_0^\infty \sigma^2(s)\,ds, \tag{2.3}$$

where we define $T^*=\infty$ if $\sigma\notin L^2([0,\infty);\mathbb{R})$.

Define also, for each $0 \le s \le T^*$, the $\mathcal{F}^B$—stopping time

$$T(s) = \inf\{t \ge 0 : \langle M \rangle(t) \ge s\}. \tag{2.4}$$

By the martingale time change theorem, there exists a standard Brownian motion $\{\tilde{B}(t); 0 \le t \le T^*; \mathcal{G}(t)\}$, such that

$$\tilde{B}(s) = M(T(s)), \quad \mathcal{G}(s) = \mathcal{F}^B(T(s)), \quad 0 \le s \le T^*,$$

and moreover

$$M(t) = \tilde{B}(\langle M \rangle(t)), \quad \mathcal{F}^B(t) = \mathcal{G}(\langle M \rangle(t)), \quad t \ge 0.$$

In what follows we presume that $g : \mathbb{R} \to \mathbb{R}$ obeys

$$g(0) = 0. \tag{2.5}$$

Since we do not want the equation to have any other equilibria in $(0, \infty)$ we ask that the non–degeneracy condition

$$g(x) > 0, \quad \text{for all } x \ne 0, \tag{2.6}$$

also be satisfied.

We are now in a position to state an existence result for the solution of the autonomous stochastic differential equation.

Proposition 2.1. *Suppose that g obeys (2.5) and (2.6) and that*

$$\text{there exists a strictly increasing function} \tag{2.7a}$$
$$q : [0, \infty) \to [0, \infty) \text{ with } q(0) = 0 \text{ such that}$$
$$\int_0^\epsilon \frac{1}{q^2(u)}\, du = \infty, \quad \text{for all } \epsilon > 0,$$

and that g and q both obey

$$|g(x) - g(y)| \le q(|x - y|), \quad x, y \in \mathbb{R}. \tag{2.7b}$$

Then there exists a unique strong non-exploding solution $\tilde{X}$ of

$$d\tilde{X}(t) = g(\tilde{X}(t))\, d\tilde{B}(t), \quad t \ge 0, \quad X(0) = \xi > 0, \tag{2.8}$$

on the complete probability space $(\Omega, \mathcal{F}, (\mathcal{G}(t))_{t \ge 0}, \mathbb{P})$.

This result is a corollary of the result of Yamada and Watanabe (see e.g. [7], Proposition 2.13, page 291, and [7], Theorem 5.4., page 332).

The main concern of this paper is the asymptotic behaviour of solutions of the nonautonomous equations

$$dX(t) = \sigma(t) g(X(t))\, dB(t), \quad t \ge 0, \quad X(0) = \xi > 0. \tag{2.9}$$

Before conducting this asymptotic analysis however, we must verify that this equation has a well-defined solution. This is accomplished by the following result.

Proposition 2.2. *Suppose that g and q obey (2.5), (2.6), (2.7a) and (2.7b) and σ obeys (2.1).*

Then there exists a unique strong non-exploding solution X of (2.9) on the complete probability space $(\Omega, \mathcal{F}, (\mathcal{F}^B(t))_{t\geq 0}, \mathbb{P})$.

Proof. By Proposition 2.1, $\tilde{X}$ is the unique strong solution of (2.8). Consider $\tilde{X}(t)$ for $t \in [0, T^*)$, where T^* is defined as in (2.3), and define

$$X(t) = \tilde{X}\left(\int_0^t \sigma^2(s)ds\right), \quad t \geq 0. \tag{2.10}$$

Then, as $\tilde{X}(s)$ is $\mathcal{G}(s)$-measurable, $\tilde{X}(t)$ is $\mathcal{G}\left(\int_0^t \sigma^2(s)ds\right)$-measurable. But $\mathcal{G}\left(\langle M\rangle(t)\right) = \mathcal{F}^B(t)$, so $X(t)$ is $\mathcal{F}^B(t)$-measurable. Now, by [7], Proposition 3.4.8, we get for $t \geq 0$

$$\int_0^{\langle M\rangle(t)} g(\tilde{X}(u))d\tilde{B}(u) = \int_0^t g(X(s))dM(s) = \int_0^t \sigma(s)g(X(s))dB(s),$$

and, therefore, as $\langle M\rangle(t) = \int_0^t \sigma^2(s)ds$, we get

$$\begin{aligned} X(t) &= \tilde{X}\left(\int_0^t \sigma^2(s)ds\right) = \tilde{X}\left(\langle M\rangle(t)\right) \\ &= \tilde{X}(0) + \int_0^{\langle M\rangle(t)} g(\tilde{X}(u))d\tilde{B}(u) = \xi + \int_0^t \sigma(s)g(X(s))dB(s). \end{aligned}$$

Therefore X defined by (2.10) is a strong solution of (2.9) on $(\Omega, \mathcal{F}, (\mathcal{F}^B(t))_{t\geq 0}, \mathbb{P})$.

To show uniqueness, suppose that there is another solution of (2.9), Y. Then $\tilde{Y}$ defined by $\tilde{Y}(s) = Y(T(s))$ obeys (2.8). Hence, as (2.8) has a unique solution, we have $\tilde{Y} = \tilde{X}$, and therefore it follows that $Y = X$.

This completes the proof. □

In this paper, we choose to write explicitly the dependence of solutions on their initial conditions, which are always assumed to be deterministic. Thus, the value at time $t \geq 0$ of the process Y with initial condition $Y(0) = \xi$ is denoted by $Y(t, \xi)$.

3 Main result

In this section, we state and discuss the main results of the paper concerning the asymptotic behaviour of non–autonomous equation. At the end of the section we present an example of a non–autonomous linear equation which can be analysed without the use of the theorems established here, but whose behaviour illustrates the results proven.

As seen in the proof of Proposition 2.2 the non–autonomous equation (2.9) is equivalent to (2.8), under a deterministic time change. However the subject of Proposition 2.2 is the existence for non–autonomous equation. The following Proposition by contrust, focusses on the relation between the solutions of the two equations.

Proposition 3.1. *Let $\xi > 0$ be deterministic. Suppose that g obeys (2.5), (2.6). Let σ obey (2.1). Suppose that $X(\cdot, \xi)$ is the unique strong solution of (2.9) with $X(0,\xi) = \xi$. Let T^* be given by* (2.3)*, and T be defined by* (2.4)*.*

(i) If $\sigma \in L^2([0,\infty);\mathbb{R})$, then there exists a standard Brownian motion $\tilde{B} = \{\tilde{B}(t); 0 \le t \le T^*; \mathcal{G}(t)\}$ where $\mathcal{G}(t) = \mathcal{F}^B(T(t))$ and $\tilde{B}(t) = B(T(t))$ such that the process $\tilde{X} = \{\tilde{X}(t); 0 \le t \le T^*; \mathcal{G}(t)\}$ defined by $\tilde{X}(t) = X(T(t))$ obeys (2.8)*.*

(ii) If $\sigma \notin L^2([0,\infty);\mathbb{R})$, then there exists a standard Brownian motion $\tilde{B} = \{\tilde{B}(t); 0 \le t < \infty; \mathcal{G}(t)\}$ where $\mathcal{G}(t) = \mathcal{F}^B(T(t))$ and $\tilde{B}(t) = B(T(t))$ such that the process $\tilde{X} = \{\tilde{X}(t); 0 \le t < \infty; \mathcal{G}(t)\}$ defined by $\tilde{X}(t) = X(T(t))$ obeys (2.8)*.*

Before stating the first result on asymptotic behaviour we present some notation and an important auxiliary result.

Suppose that $\tilde{X}(\cdot,\xi)$ is the solution of (2.8), where $\xi > 0$. Define

$$\tilde{S}_0(\xi) = \inf\{t \ge 0 : \tilde{X}(t,\xi) = 0\}. \tag{3.1}$$

Let us suppose that for $\delta > 0$ we may define the function $v : (0,\infty) \to (0,\infty)$ by

$$v(x) = 2\int_x^\delta \int_y^\delta \frac{dz}{g^2(z)}\,dy, \quad x > 0. \tag{3.2}$$

The following result is due to Feller (see e.g. [7], Theorem 5.5.29, page 348).

Proposition 3.2. *Let $\xi > 0$ be deterministic, and $\tilde{X}(\cdot,\xi)$ be a strong solution of (2.8). If $\tilde{S}_0(\xi)$ is as defined in (3.1), then*

$$\lim_{t \to \tilde{S}_0(\xi)} \tilde{X}(t,\xi) = 0, \qquad \sup_{0 \le t < \tilde{S}_0(\xi)} \tilde{X}(t,\xi) < \infty, \quad a.s.$$

Let v be defined by (3.2). Then

(i) $\lim_{x\to 0^+} v(x) < \infty$ *implies* $\tilde{S}_0(\xi) < \infty$*, a.s.;*

(ii) $\lim_{x\to 0^+} v(x) = \infty$ *implies* $\tilde{S}_0(\xi) = \infty$*, a.s.*

We define also

$$S_0(\xi) = \inf\{t \ge 0 : X(t,\xi) = 0\}. \tag{3.3}$$

We can determine whether $S_0(\xi)$ is finite or infinite with the help of Proposition 2.2.

We may now state the first main result on the asymptotic behaviour in this paper. It is a direct consequence of Proposition 3.1 and Proposition 3.2.

Theorem 3.3. *Let $\xi > 0$ be deterministic. Suppose that g obeys (2.5), (2.6). Let σ obey (2.1). Suppose that $X(\cdot, \xi)$ is the unique strong solution of (2.9) with $X(0, \xi) = \xi$.*

Let v be defined as in (3.2), and suppose that

$$\lim_{x \to 0^+} v(x) = \infty.$$

Then we have the following case distinction:

(a) If $\sigma \in L^2((0, \infty); \mathbb{R})$, then there exists an almost surely positive and $\mathcal{F}^B(\infty)$–measurable random variable $L = L(\xi, \omega)$ such that

$$\lim_{t \to \infty} X(t, \xi) = L(\xi) > 0, \quad a.s. \tag{3.4}$$

(b) If $\sigma \notin L^2((0, \infty); \mathbb{R})$, then

$$\lim_{t \to \infty} X(t, \xi) = 0, \quad a.s.$$

and $S_0(\xi)$ defined by (3.3) obeys $S_0(\xi) = \infty$, a.s.

The proof of this result and proofs of subsequent results in the this section, are postponed to the final section of the paper.

When $\sigma \in L^2([0, \infty); \mathbb{R})$ the rate at which convergence to $L(\xi)$ occurs can be determined exactly.

Theorem 3.4. *Let $\xi > 0$ be deterministic. Suppose that g obeys (2.5), (2.6), and let $g \in C^1((0, \infty); (0, \infty))$. Let σ obey (2.1) and $\sigma \in L^2([0, \infty); \mathbb{R})$. Suppose that $X(\cdot, \xi)$ is the unique strong solution of (2.9) with $X(0, \xi) = \xi$.*

Let v be defined as in (3.2), and suppose that

$$\lim_{x \to 0^+} v(x) = \infty.$$

Let $L(\xi)$ be the almost surely positive and $\mathcal{F}^B(\infty)$–measurable random variable defined by (3.4). *Then:*

(i) If σ obeys

$$\int_t^\infty \sigma^2(s)\, ds > 0, \quad \text{for all } t \geq 0, \tag{3.5}$$

then

$$\limsup_{t \to \infty} \frac{X(t, \xi) - L(\xi)}{\sqrt{2 \int_t^\infty \sigma^2(s)\, ds \log \log \left(\int_t^\infty \sigma^2(s)\, ds \right)^{-1}}} = g(L(\xi)), \quad a.s., \tag{3.6a}$$

$$\liminf_{t \to \infty} \frac{X(t, \xi) - L(\xi)}{\sqrt{2 \int_t^\infty \sigma^2(s)\, ds \log \log \left(\int_t^\infty \sigma^2(s)\, ds \right)^{-1}}} = -g(L(\xi)), \quad a.s. \tag{3.6b}$$

(ii) If σ does not obey (3.5) *i.e., if there exists $\tau \geq 0$ such that $\int_t^\infty \sigma^2(s)\,ds = 0$ for all $t \geq \tau$, then*

$$X(t,\xi) = X(\tau,\xi) = L(\xi), \quad \textit{for all } t \geq \tau, \textit{ a.s.}$$

Of course, in case (b) in Theorem 3.3, the rate of convergence cannot be so easily computed. However pathwise rates of decay to zero for nonlinear autonomous stochastic differential equations have been found in [3, 4], and could readily be applied here.

Finally, the distribution of the random limit L is in principle well–understood, by using the forward Kolmogorov equation for the process $\tilde{X}$.

Theorem 3.5. *Let $\xi > 0$ be deterministic. Suppose that g obeys (2.5), (2.6). Let σ obey (2.1) and $\sigma \in L^2([0,\infty);\mathbb{R})$. Let T^* be given by* (2.3)*. Suppose that $X(\cdot,\xi)$ is the unique strong solution of (2.9) with $X(0,\xi) = \xi$.*

Let v be defined as in (3.2), and suppose that

$$\lim_{x\to 0^+} v(x) = \infty,$$

and let $L(\xi)$ be the almost surely positive and $\mathcal{F}^B(\infty)$–measurable random variable defined by (3.4)*. Then*

$$\mathbb{P}[L(\xi) \leq x] = \int_0^x \Gamma(T^*;y)\,dy, \quad x \geq 0,$$

where

$$\frac{\partial \Gamma}{\partial t}(t;y) = \frac{1}{2}\frac{\partial^2}{\partial y^2}\left(g^2(y)\Gamma(t;y)\right), \quad (t,y) \in [0,T^*] \times (0,\infty),$$

and $\Gamma(0;y) = \delta_\xi(y)$, $y \in \mathbb{R}$, where δ_ξ is the δ-function.

The result holds because $L(\xi) = \lim_{t\to T^*+} \tilde{X}(t,\xi) = \tilde{X}(T^*,\xi)$. Moreover, as $\tilde{X}$ is a diffusion process with known infinitesimal generator and deterministic initial condition ξ, we can deduce its distribution function from the forward Kolmogorov equation, and therefore, the distribution of $L(\xi)$ is also known.

It remains merely to classify the behaviour in the case when $\lim_{x\to 0} v(x) < \infty$.

Theorem 3.6. *Let $\xi > 0$ be deterministic. Suppose that g obeys (2.5), (2.6). Let σ obey (2.1). Suppose that $X(\cdot,\xi)$ is the unique strong solution of (2.9) with $X(0,\xi) = \xi$.*

Let v be defined as in (3.2), and suppose that

$$\lim_{x\to 0^+} v(x) < \infty.$$

Then we have the following case distinction:

(a) If $\sigma \in L^2((0,\infty);\mathbb{R})$, then there exists an almost surely positive and $\mathcal{F}^B(\infty)$–measurable random variable $L = L(\xi,\omega)$ such that

$$\lim_{t\to\infty} X(t,\xi) = L(\xi) > 0, \quad \textit{a.s. on } \{\tilde{S}_0(\xi) \geq T^*\},$$

and

$$\lim_{t\to S_0(\xi)^-} X(t,\xi) = 0, \quad \text{a.s. on } \{\tilde{S}_0(\xi) < T^*\},$$

where $\tilde{S}_0(\xi)$ *defined by* (3.1) *and* $S_0(\xi)$ *is defined by* (3.3).

(b) *If* $\sigma \notin L^2((0,\infty);\mathbb{R})$, *then*

$$\lim_{t\to S_0(\xi)^-} X(t,\xi) = 0, \quad a.s.,$$

where $S_0(\xi)$ *defined by (3.3) obeys* $S_0(\xi) < +\infty$, *a.s.*

In case (a) the rate of convergence to the non–trivial random limit is the same as given in Theorem 3.4, but only a.s. on the event $\{\tilde{S}_0(\xi) > T^*\}$.

The probability of the event $\{\tilde{S}_0(\xi) < T^*\}$ can be computed for the process $\tilde{X}$ obeying (2.8), by considering the limit

$$\mathbb{P}[\tilde{S}_0(\xi) < T^*] = \lim_{a\to 0^+} \mathbb{P}[\tilde{S}_a(\xi) < T^*]$$

where for $\xi > a > 0$ we define $\tilde{S}_a(\xi) = \inf\{t \geq 0 : \tilde{X}(t,\xi) = a\}$. It is possible to compute the moment generating function of $\tilde{S}_a(\xi)$, $\lambda \mapsto \mathbb{E}[e^{-\lambda \tilde{S}_a(\xi)}]$ for $\lambda \geq 0$, by solving an appropriate Sturm–Liouville problem, from which the probability $\mathbb{P}[\tilde{S}_a(\xi) < T^*]$ can in principle be determined by inverse transform methods. The interested reader can refer to [6, Chapter 4.11] for further details on computation of the moment generating function.

3.1 An example

A simple example of a process which can be analysed completely *without* appealing to these results (but which is consistent with them) is the unique strong solution of

$$X(t) = \xi + \int_0^t \sigma(s)X(s)\,dB(s), \quad t \geq 0,$$

where $\xi > 0$ and $\sigma \in C([0,\infty);\mathbb{R})$. This equation has explicit solution

$$X(t,\xi) = \xi \exp\left(\int_0^t \sigma(s)\,dB(s) - \frac{1}{2}\int_0^t \sigma^2(s)\,ds\right), \quad t \geq 0.$$

Here we identify $g(x) = x$, $x \geq 0$, and have $\lim_{x\to 0^+} v(x) = \infty$. Hence Theorems 3.3, 3.4 and 3.5 can be applied to this stochastic differential equation.

In the case when $\sigma \in L^2([0,\infty);\mathbb{R})$, the martingale convergence theorem (cf., e.g., [8, Proposition IV.1.26]) ensures that $\lim_{t\to\infty} \int_0^t \sigma(s)\,dB(s)$ exists and is almost surely finite. Therefore, there is an almost surely positive and almost surely finite $\mathcal{F}^B(\infty)$—measurable random variable $L(\xi)$ given by

$$L(\xi) = \xi \exp\left(\int_0^\infty \sigma(s)\,dB(s) - \frac{1}{2}\int_0^\infty \sigma^2(s)\,ds\right)$$

such that

$$\lim_{t\to\infty} X(t,\xi) = L(\xi) > 0, \quad \text{a.s.}$$

This chimes with part (a) of Theorem 3.3.

In the case when $\sigma \notin L^2([0,\infty);\mathbb{R})$, we have that

$$\lim_{t\to\infty}\int_0^t \sigma^2(s)\,ds = +\infty, \quad \lim_{t\to\infty}\frac{\int_0^t \sigma(s)\,dB(s)}{\int_0^t \sigma^2(s)\,ds} = 0, \quad \text{a.s.}$$

The latter fact resulting from the Strong Law of Large Numbers for martingales (cf., e.g., [8, Exercise V.I.16]). Therefore

$$\lim_{t\to\infty}\frac{1}{\int_0^t \sigma^2(s)\,ds}\log X(t) = -\frac{1}{2}, \quad \text{a.s.}$$

Hence $\lim_{t\to\infty} X(t,\xi) = 0$, a.s., which agrees with part (b) of Theorem 3.3.

Moreover, in the case when $\sigma \in L^2([0,\infty);\mathbb{R})$, $L(\xi)$ is lognormally distributed; this is obvious by observation of the formula for $L(\xi)$, but can also be confirmed by solving the partial differential equation for the transition density Γ in Theorem 3.5.

In the case when $\sigma \in L^2([0,\infty);\mathbb{R})$, the rate of convergence in Theorem 3.4 can be obtained, if it is shown that

$$\limsup_{t\to\infty}\frac{\int_t^\infty \sigma(s)\,dB(s)}{\sqrt{2\int_t^\infty \sigma^2(s)\,ds\log\log\left(\int_t^\infty \sigma^2(s)\,ds\right)^{-1}}} = 1, \quad \text{a.s.,} \tag{3.7a}$$

$$\liminf_{t\to\infty}\frac{\int_t^\infty \sigma(s)\,dB(s)}{\sqrt{2\int_t^\infty \sigma^2(s)\,ds\log\log\left(\int_t^\infty \sigma^2(s)\,ds\right)^{-1}}} = -1, \quad \text{a.s.} \tag{3.7b}$$

This can be established as follows: define $T^* = \int_0^\infty \sigma^2(s)\,ds$, and let M be the local martingale defined in (2.2), and with square variation $\langle M\rangle$. Then, by the martingale time change theorem there exists a Brownian motion $\tilde{B}$ such that $M(t) = \tilde{B}(\langle M\rangle(t))$, $0 \le t < \infty$. Thus

$$\begin{aligned}
&\limsup_{t\to\infty}\frac{\int_t^\infty \sigma(s)\,dB(s)}{\sqrt{2\int_t^\infty \sigma^2(s)\,ds\log\log\left(\int_t^\infty \sigma^2(s)\,ds\right)^{-1}}}\\
&= \limsup_{t\to\infty}\frac{M(\infty)-M(t)}{\sqrt{2(\langle M\rangle(\infty)-\langle M\rangle(t))\log\log\left(\langle M\rangle(\infty)-\langle M\rangle(t)\right)^{-1}}}\\
&= \limsup_{t\to\infty}\frac{\tilde{B}(\langle M\rangle(\infty))-\tilde{B}(\langle M\rangle(t))}{\sqrt{2(\langle M\rangle(\infty)-\langle M\rangle(t))\log\log\left(\langle M\rangle(\infty)-\langle M\rangle(t)\right)^{-1}}}\\
&= \limsup_{s\to T^{*-}}\frac{\tilde{B}(T^*)-\tilde{B}(s)}{\sqrt{2(T^*-s)\log\log(T^*-s)^{-1}}}\\
&= \limsup_{u\to 0^+}\frac{\tilde{B}(T^*)-\tilde{B}(T^*-u)}{\sqrt{2u\log\log(u)^{-1}}} = 1, \quad \text{a.s.,}
\end{aligned}$$

since $\bar{B}$ defined by $\bar{B}(t) = \tilde{B}(T^*) - \tilde{B}(T^* - t)$, $0 \le t \le T^*$, is also a standard Brownian motion, and therefore subject to the Law of the Iterated Logarithm. (3.7) was stated in [2], but a proof was not supplied there. A variant of this result is proven in [1] but this proof contains an error as it is *incorrectly* stated there that

$$\int_t^\infty X(s)\,dB(s) = \int_0^{1/t} \frac{1}{s} X(1/s)\,dW(s)$$

for a process $X \in L^2([0,\infty);\mathbb{R})$ a.s., where W is the standard Brownian motion given by $W(t) = tB(1/t)$ for $t > 0$ and $W(0) = 0$.

4 Proofs

4.1 Proof of Theorem 3.3

By Proposition 3.2, it follows that $\lim_{x\to 0^+} v(x) = \infty$ implies $\tilde{S}_0(\xi) = +\infty$, a.s.

In case (a), when $\sigma \in L^2([0,\infty);\mathbb{R})$, $T(s) \to \infty$ as $s \to T^{*-}$. Hence, as $\tilde{X}(s) = X(T(s))$ for $s \in [0, T^*]$,

$$\lim_{t\to\infty} X(t) = \lim_{s\to T^{*-}} X(T(s)) = \lim_{s\to T^{*-}} \tilde{X}(s) = \tilde{X}(T^*) > 0, \quad \text{a.s.},$$

because $T^* < +\infty = \tilde{S}_0(\xi)$, a.s.

In case (b), when $\sigma \notin L^2([0,\infty);\mathbb{R})$, $T(s) \to \infty$ as $s \to \infty$. Hence, as $\tilde{X}(s) = X(T(s))$ for $s \in [0,\infty)$, we have $S_0(\xi) = T(\tilde{S}_0(\xi)) = +\infty$, a.s., and

$$\lim_{t\to\infty} X(t) = \lim_{s\to\infty} X(T(s)) = \lim_{s\to\infty} \tilde{X}(s) = \lim_{s\to\tilde{S}_0(\xi)} \tilde{X}(s) = 0, \quad \text{a.s.},$$

because $\tilde{S}_0(\xi) = \infty$, a.s.

4.2 Proof of Theorem 3.4

The proof of part (ii) is straightforward, because for $t \ge \tau$ we have

$$X(t) = X(\tau) + \int_\tau^t \sigma(s) g(X(s))\,dB(s) = X(\tau),$$

as $\int_\tau^\infty \sigma^2(s)\,ds = 0$ and the continuity of σ imply that $\sigma(t) = 0$ for all $t \in [\tau,\infty)$.

To prove part (i) we proceed as follows. Because $\sigma \in L^2([0,\infty);\mathbb{R})$, and $\lim_{x\to 0^+} v(x) = \infty$, by Proposition 3.1 and Proposition 3.2, the process $\tilde{X} = \{X(t); 0 \le t \le T^*; \mathcal{G}(t)\}$ defined by $\tilde{X}(t) = X(T(t))$ is strictly positive a.s. and obeys

$$\tilde{X}(t) = \xi + \int_0^t g(\tilde{X}(s))\,d\tilde{B}(s), \quad 0 \le t \le T^*,$$

where T^* and T are defined by (2.3) and (2.4).

Since g in $C^1((0,\infty);(0,\infty))$ we may define the function $h \in C^2((0,\infty);\mathbb{R})$ by

$$h(x) = \int_1^x \frac{1}{g(u)}\,du, \quad x \in \mathbb{R}.$$

Then the process $\tilde{Y} = \{\tilde{Y}(t); 0 \le t \le T^*; \mathcal{G}(t)\}$ defined by $\tilde{Y}(t) = h(\tilde{X}(t))$ is well–defined. Since $h \in C^2((0,\infty);\mathbb{R})$ and $\tilde{X}(t) > 0$ for all $t \in [0,T^*]$ a.s. $\tilde{Y}$ is an Itô–process, which by Itô's rule, has decomposition for $0 \le t \le T^*$ given by

$$\tilde{Y}(t) = h(\tilde{X}(t)) = h(\xi) + \tilde{B}(t) - \frac{1}{2}\int_0^t g'(\tilde{X}(s))\,ds.$$

Since $\tilde{X}$ is almost surely positive and continuous, $g \in C^1((0,\infty);(0,\infty))$, it follows that

$$\lim_{t\to T^*-} \frac{1}{T^*-t}\int_t^{T^*} g'(\tilde{X}(s))\,ds = g'(\tilde{X}(T^*)), \quad \text{a.s.}$$

Therefore

$$\lim_{t\to T^*-} \frac{1}{\sqrt{2(T^*-t)\log\log(1/(T^*-t))}}\int_t^{T^*} g'(\tilde{X}(s))\,ds = 0, \quad \text{a.s.}$$

Then for $t \in [0,T^*]$

$$h(\tilde{X}(T^*)) - h(\tilde{X}(t)) = \tilde{B}(T^*) - \tilde{B}(t) - \frac{1}{2}\int_t^{T^*} g'(\tilde{X}(s))\,ds,$$

and

$$\begin{aligned}
&\limsup_{t\to T^*-} \frac{h(\tilde{X}(T^*)) - h(\tilde{X}(t))}{\sqrt{2(T^*-t)\log\log(1/(T^*-t))}} \\
&= \limsup_{t\to T^*-} \frac{\tilde{B}(T^*) - \tilde{B}(t)}{\sqrt{2(T^*-t)\log\log(1/(T^*-t))}} \\
&= \limsup_{s\to 0+} \frac{\tilde{B}(T^*) - \tilde{B}(T^*-s)}{\sqrt{2s\log\log(1/s)}}.
\end{aligned}$$

Now, because $\bar{B} = \{\bar{B}(t) : 0 \le t \le T^*; \mathcal{F}^{\bar{B}}(t)\}$ defined by $\bar{B}(t) = \tilde{B}(T^*) - \tilde{B}(T^*-t)$ is a standard Brownian motion, by the Law of the Iterated Logarithm for Brownian motion we have

$$\limsup_{s\to 0^*+} \frac{\tilde{B}(T^*) - \tilde{B}(T^*-s)}{\sqrt{2s\log\log(1/s)}} = \limsup_{s\to 0^*+} \frac{\bar{B}(s)}{\sqrt{2s\log\log(1/s)}} = 1, \quad \text{a.s.}$$

Hence

$$\limsup_{t\to T^*-} \frac{h(\tilde{X}(T^*)) - h(\tilde{X}(t))}{\sqrt{2(T^*-t)\log\log(1/(T^*-t))}} = 1, \quad \text{a.s.}$$

Since h is in $C^1((0,\infty);\mathbb{R})$ and $\tilde{X}$ has continuous sample paths, we have

$$\lim_{t\to T^*-} \frac{h(\tilde{X}(T^*)) - h(\tilde{X}(t))}{\tilde{X}(T^*) - \tilde{X}(t)} = h'(\tilde{X}(T^*)) = \frac{1}{g(\tilde{X}(T^*))}, \quad \text{a.s.}$$

Hence

$$\begin{aligned}
&\limsup_{t\to T^*-} \frac{\tilde{X}(T^*)-\tilde{X}(t)}{\sqrt{2(T^*-t)\log\log(1/(T^*-t))}}\\
&= \limsup_{t\to T^*-} \frac{\tilde{X}(T^*)-\tilde{X}(t)}{h(\tilde{X}(T^*))-h(\tilde{X}(t))} \frac{h(\tilde{X}(T^*))-h(\tilde{X}(t))}{\sqrt{2(T^*-t)\log\log(1/(T^*-t))}}\\
&= g(\tilde{X}(T^*)), \quad \text{a.s.}
\end{aligned}$$

Therefore, as $T(s)\to\infty$ as $s\uparrow T^*$, and $\langle M\rangle(T(s))=s$ for $s\in[0,T^*]$, we have

$$\begin{aligned}
&\limsup_{t\to\infty} \frac{L(\xi)-X(t)}{\sqrt{2\int_t^\infty \sigma^2(s)\,ds \log\log(1/\int_t^\infty \sigma^2(s)\,ds)}}\\
&= \limsup_{t\to\infty} \frac{X(\infty)-X(t)}{\sqrt{2(\langle M\rangle(\infty)-\langle M\rangle(t))\log\log(1/\langle M\rangle(\infty)-\langle M\rangle(t))}}\\
&= \limsup_{s\uparrow T^*} \frac{X(T(T^*))-X(T(s))}{\sqrt{2(\langle M\rangle(T(T^*))-\langle M\rangle(T(s)))\log\log(1/\langle M\rangle(T(T^*))-\langle M\rangle(T(s)))}}\\
&= \limsup_{s\uparrow T^*} \frac{\tilde{X}(T^*)-\tilde{X}(s)}{\sqrt{2(T^*-s)\log\log(1/(T^*-s))}}\\
&= g(\tilde{X}(T^*))\\
&= g(L(\xi)), \quad \text{a.s.}
\end{aligned}$$

An analogous argument gives

$$\liminf_{t\to T^*-} \frac{\tilde{X}(T^*)-\tilde{X}(t)}{\sqrt{2(T^*-t)\log\log(1/(T^*-t))}} = -g(\tilde{X}(T^*)), \quad \text{a.s.,}$$

from which we can infer that

$$\liminf_{t\to\infty} \frac{L(\xi)-X(t)}{\sqrt{2\int_t^\infty \sigma^2(s)\,ds \log\log(1/\int_t^\infty \sigma^2(s)\,ds)}} = -g(L(\xi)), \quad \text{a.s.,}$$

as required.

4.3 Proof of Theorem 3.6

By Proposition 3.2, it follows that $\lim_{x\to 0^+} v(x) < +\infty$ implies $\tilde{S}_0(\xi) < +\infty$, a.s.

In case (b) when $\sigma \notin L^2([0,\infty);\mathbb{R})$, $T(s)\to\infty$ as $s\to\infty$. Hence, as $\tilde{X}(s) = X(T(s))$ for $s\in[0,\infty)$, we have $S_0(\xi) = T(\tilde{S}_0(\xi)) < +\infty$, a.s., and

$$\lim_{s\to S_0(\xi)^-} X(s) = \lim_{t\to \tilde{S}_0(\xi)^-} X(T(t)) = \lim_{t\to \tilde{S}_0(\xi)^-} \tilde{X}(t) = 0, \quad \text{a.s.,}$$

because $\tilde{S}_0(\xi) < +\infty$, a.s.

In case (a) when $\sigma \in L^2([0,\infty);\mathbb{R})$, $T(s) \to \infty$ as $s \to T^{*-}$. Define the event

$$A = \{\omega : \tilde{S}_0(\xi) \geq T^*\}.$$

Then because $\tilde{X}(s) = X(T(s))$ for $s \in [0,T^*]$, we have $S_0(\xi) = T(\tilde{S}_0(\xi))$, and so

$$A = \{\omega : \tilde{S}_0(\xi) \geq T^*\} = \{\omega : S_0(\xi) = +\infty\}.$$

Clearly, as $\tilde{S}_0(\xi) \geq T^*$ on A, we have

$$\lim_{t\to T^{*-}} \tilde{X}(t) = \tilde{X}(T^*) > 0, \quad \text{a.s. on } A.$$

Thus, as $\tilde{X}(s) = X(T(s))$ for $s \in [0,T^*]$,

$$\lim_{t\to\infty} X(t) = \lim_{s\to T^{*-}} X(T(s)) = \lim_{s\to T^{*-}} \tilde{X}(s) = \tilde{X}(T^*) > 0, \quad \text{a.s. on } A.$$

Hence

$$\left\{\lim_{t\to\infty} X(t,\xi) = L(\xi) > 0,\ S_0(\xi) = +\infty\right\} = \{\tilde{S}_0(\xi) \geq T^*\}, \quad \text{a.s.}$$

On the other hand, consider the event $\bar{A}$, where

$$\bar{A} = \{\omega : \tilde{S}_0(\xi) < T^*\} = \{\omega : S_0(\xi) < +\infty\},$$

by virtue of the fact that $S_0(\xi) = T(\tilde{S}_0(\xi)) < T(T^*) = +\infty$. Then

$$\lim_{t\to\tilde{S}_0(\xi)^-} \tilde{X}(t) = 0, \quad \text{a.s. on } \bar{A}.$$

Thus, as $\tilde{X}(s) = X(T(s))$ for $s \in [0,T^*]$,

$$\lim_{t\to S_0(\xi)^-} X(t) = \lim_{s\to\tilde{S}_0(\xi)^-} X(T(s)) = \lim_{s\to\tilde{S}_0(\xi)^-} \tilde{X}(s) = 0, \quad \text{a.s. on } \bar{A}.$$

Hence

$$\left\{\lim_{t\to S_0(\xi)^-} X(t,\xi) = 0, S_0(\xi) < +\infty\right\} = \{\tilde{S}_0(\xi) < T^*\}, \quad \text{a.s.}$$

Received: July 2008. Revised: August 2008.

References

[1] J.A.D. Appleby *Almost sure subexponential decay rates of scalar Itô-Volterra equations*, Electron. J. Qual. Theory Differ. Equ., Proc. 7th Coll. QTDE, Paper no.1, 1–32, 2004.

[2] J.A.D. Appleby and D. Mackey, *Almost sure polynomial asymptotic stability of scalar stochastic differential equations with damped stochastic perturbations*, Electron. J. Qual. Theory Differ. Equ., Proc. 7th Coll. QTDE, Paper no.2, 1–33, 2004.

[3] J.A.D. Appleby, X. Mao and A. Rodkina, *Pathwise superexponential decay rates of solutions of autonomous stochastic differential equations*, Stochastics, **77**(3), 241–270, 2005.

[4] J.A.D. Appleby, A. Rodkina and H. Schurz, *Pathwise non-exponential decay rates of solutions of scalar nonlinear stochastic differential equation*, Disc. Con. Dynam. Sys. Ser. B., **6**(4), 667–696, 2006.

[5] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, North Holland, New York, 2nd edition, 1981.

[6] K. Itô and H.P. McKean, *Diffusion processes and their sample paths*, Springer, Berlin, 2nd edition, 1974.

[7] I. Karatzas and S.E. Shreve, *Brownian Motion and Stochastic Calculus*, volume 113 of Graduate Texts in Mathematics, Springer, New York, 1991.

[8] D. Revuz and M. Yor, *Continuous Martingales and Brownian Motion.* Third edition, Springer, New York, 1999.

CUBO A Mathematical Journal
Vol.10, Nº03, (161–170). October 2008

Existence and Uniqueness of Pseudo Almost Automorphic Solutions to Some Classes of Partial Evolution Equations

J. Blot, D. Pennequin
Université Paris 1 Panthéon-Sorbonne, Laboratoire Marin Mersenne,
Centre P.M.F., 90 rue de Tolbiac,
75647 PARIS Cedex 13, FRANCE
emails: `blot@univ-paris1.fr, pennequi@univ-paris1.fr`

AND

Gaston M. N'Guérékata
Department of Mathematics, Morgan State University,
1700 E. Cold Spring Lane, Baltimore, M.D. 21251 – USA
email: `Gaston.N'Guerekata@morgan.edu`

ABSTRACT

We are concerned in this paper with the partial differential equation $\frac{d}{dt}[u(t)+f(t,u(t))] = Au(t) \quad t \in \mathbb{R}$, where A is a (generally unbounded) linear operator which generates a semigroup of bounded linear operators $(T(t))_{t\geq 0}$. Under appropriate sufficient conditions, we prove the existence and uniqueness of a pseudo almost automorphic mild solution to the equation.

RESUMEN

Nosotros consideramos en este artículo la ecuación diferencial parcial $\frac{d}{dt}[u(t)+f(t,u(t))] = Au(t) \quad t \in \mathbb{R}$, donde A es un (generalmente no acotado) operador lineal que genera un semigrupo de operadores lineales acotados $(T(t))_{t\geq 0}$. Bajo condiciones suficientes apropriadas, provamos la existencia y unicidad de una solución blanda pseudo casi automorfica para tal ecuación.

Key words and phrases: *Pseudo Almost automorphic function, exponentially stable semigroup, partial differential equations.*

Math. Subj. Class.: *34G10, 47A55.*

1 Introduction

Since the publication of the monograph [12], the study of almost automorphic function (a concept introduced by S. Bochner in the literature in the mid sixties as a generalization of almost periodicity in the sense of Bohr) has regained great interest. Several extensions of the concept were introduced including asymptotic almost automorphy by N'Guérékata ([10]), p-almost automorphy by Diagana ([2]), and Stepanov-like almost automorphy by N'Guérékata and Pankov ([14]). Recently, J. Liang et al. have suggested the notion of pseudo almost automorphic functions, i.e. functions that can be written uniquely as a sum of an almost automorphic function and an ergodic term, i.e. a function with vanishing mean (cf [6], and [7], [8]). This latter turns out to be more general than asymptotic almost automorphy. However it seems to be more complicated.

There has been a considerable interest in the existence of (these various types of) almost automorphic solutions of evolution equations. Semigroups theory and fixed point techniques have been frequently used for semilinear evolution equations. In [3] the authors studied the existence and uniqueness of an almost automorphic mild solution to the equation

$$\frac{d}{dt}\left[u(t)+f(t,u(t))\right]=Au(t)+g(t,u(t)) \quad t\in\mathbb{R}, \tag{1.1}$$

where the functions $f(t,u)$ and $g(t,u)$ are almost automorphic in t, for each u. This latter motivated our recent paper [15], where we study the existence and uniqueness of a pseudo almost automorphic mild solution the semilinear evolution equations of the form

$$\frac{du}{dt}=Au(t)+g(t,u(t)), \qquad t\in\mathbb{R}, \tag{1.2}$$

where A is an unbounded sectorial operator with not necessarily dense domain in a Banach space $\mathbb{X}$ and $g:\mathbb{R}\times\mathbb{X}_\alpha\to\mathbb{X}$, where $\mathbb{X}_\alpha$, $\alpha\in(0,1)$, is any intermediate Banach space between $D(A)$ and $\mathbb{X}$.

In this paper, we study pseudo almost automorphic solutions to perturbations to Equations:

$$\frac{du}{dt}=Au(t) \quad t\in\mathbb{R}, \tag{1.3}$$

consisting of the class of abstract partial evolution equations of the form

$$\frac{d}{dt}\left[u(t)+f(t,u(t))\right]=Au(t) \quad t\in\mathbb{R}, \tag{1.4}$$

where A is the infinitesimal generator of an exponentially stable C_0-semigroup acting on $\mathbb{X}$, B, C are two densely defined closed linear operators on $\mathbb{X}$, and f is continuous functions.

Under some appropriate assumptions, we establish the existence and uniqueness of an almost automorphic (mild) solution to Eq. (1.4) using the Banach fixed-point principle.

We start this work by presenting some properties of pseuso almost automorphic functions in Section 2 including an application to a Volterra-like integral equation. Our main result (Theoreme 3.3) is presented in Section 3.

2 Preliminaries

In this work, $(\mathbb{X}, \|\cdot\|)$ will stand for a Banach space. The collection of all bounded linear operators from $\mathbb{X}$ is denoted by $B(\mathbb{X})$ — this is a Banach space when it is equipped with its natural norm $\|A\|_{B(\mathbb{X})} := \sup\limits_{x\in\mathbb{X}, x\neq 0} \dfrac{\|Ax\|_{\mathbb{X}}}{\|x\|_X}$.

The fields of real and complex numbers, are respectively denoted by $\mathbb{C}$ and $\mathbb{R}$. We let $BC(\mathbb{R},\mathbb{X})$ denote the space of all $\mathbb{X}$-valued bounded continuous functions $\mathbb{R}\to\mathbb{X}$– it is a Banach space when equipped with the sup norm $\|u\|_\infty := \sup\limits_{t\in\mathbb{R}} \|u(t)\|_{\mathbb{X}}$ for each $u\in B(\mathbb{R},\mathbb{X})$.

We will use the following well-known concepts in the sequel.

Definition 2.1. A continuous function $f:\mathbb{R}\mapsto\mathbb{X}$ is said to be almost automorphic if for every sequence of real numbers $(s'_n)_{n\in\mathbb{N}}$ there exists a subsequence $(s_n)_{n\in\mathbb{N}}\subset(s'_n)_{n\in\mathbb{N}}$ such that

$$g(t) := \lim_{n\mapsto\infty} f(t+s_n)$$

is well-defined for each $t\in\mathbb{R}$, and

$$f(t) = \lim_{n\mapsto\infty} g(t-s_n)$$

for each $t\in\mathbb{R}$.

Similarly,

Definition 2.2. A continuous function $f:\mathbb{R}\times\mathbb{X}\mapsto\mathbb{X}$ is said to be almost automorphic in $t\in\mathbb{R}$ for each $u\in\mathbb{X}$ if every sequence of real numbers $(\sigma_n)_{n\in\mathbb{N}}$ contains a subsequence $(s_n)_{n\in\mathbb{N}}$ such that

$$g(t,u) := \lim_{n\mapsto\infty} f(t+s_n,u)$$

is well defined for each $t\in\mathbb{R}$ and each $u\in\mathbb{X}$ and,

$$f(t,u) = \lim_{n\mapsto\infty} g(t-s_n,u)$$

exists for each $t\in\mathbb{R}$ and $u\in\mathbb{X}$.

The following natural properties hold: If $f, h : \mathbb{R} \mapsto \mathbb{X}$ are almost automorphic functions and if $\lambda \in \mathbb{R}$, then $f + h$, λf, and f_λ are almost automorphic, where $f_\lambda(t) := f(t + \lambda)$. Moreover, $R(f) := \{f(t), t \in \mathbb{R}\}$ is relatively compact.

Since the range of an almost automorphic function f is relatively compact on $\mathbb{X}$, then it is bounded. Almost automorphic functions constitute a Banach space $AA(\mathbb{X})$ when it is endowed with the sup norm. This naturally generalizes the concept of (Bochner) almost periodic functions.

Definition 2.3. Let $\mathbb{X}$ be a Banach space.

1. A bounded continuous function with vanishing mean value can be defined as

$$AA_0(\mathbb{R}, \mathbb{X}) = \left\{ \phi \in BC(\mathbb{R}, \mathbb{X}) : \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} \|\phi(\sigma)\| d\sigma = 0 \right\}.$$

2. Similarly we define $AA_0(\mathbb{R} \times \mathbb{X}, \mathbb{X})$ to be the collection of all functions $f : t \mapsto f(t, x) \in BC(\mathbb{R} \times \mathbb{X}, \mathbb{X})$ satisfying

$$\lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} \|f(\sigma, x)\| d\sigma = 0$$

uniformly for x in any bounded subset of $\mathbb{X}$.

Now we describe the sets $PAA(\mathbb{R}, \mathbb{X})$ and $PAA(\mathbb{R} \times \mathbb{X}, \mathbb{X})$ of pseudo almost automorphic functions:

$$PAA(\mathbb{R}, \mathbb{X}) = \left\{ \begin{array}{l} f = g + \phi \in BC(\mathbb{R}, \mathbb{X}), \\ g \in AA(\mathbb{R}, \mathbb{X}) \ \ and \ \ \phi \in AA_0(\mathbb{R}, \mathbb{X}) \end{array} \right\};$$

$$PAA(\mathbb{R} \times \mathbb{X}, \mathbb{X}) = \left\{ \begin{array}{l} f = g + \phi \in BC(\mathbb{R} \times \mathbb{X}, \mathbb{X}), \\ g \in AA(\mathbb{R} \times \mathbb{X}, \mathbb{X}) \ \ and \ \ \phi \in AA_0(\mathbb{R} \times \mathbb{X}, \mathbb{X}) \end{array} \right\}.$$

In both cases above, g and ϕ are called respectively the principal and the ergodic terms of f.

We have the following elementary properties of pseudo almost automorphic functions.

Theorem 2.4. *([8] Theorem 2.2).*

$PAA(\mathbb{R}, \mathbb{X})$ *is a Banach space under the supremum norm.*

Let now $f, h : \mathbb{R} \to \mathbb{R}$ and consider the convolution

$$(f \star h)(t) := \int_{\mathbb{R}} f(s) h(t - s) ds, \quad t \in \mathbb{R},$$

if the integral exists.

Remark 2.5. The operator $J : PAA(\mathbb{R}, \mathbb{X}) \to PAA(\mathbb{R}, \mathbb{X})$ such that $(Jx)(t) := x(-t)$ is well-defined and linear. Moreover it is an isometry and $J^2 = I$.

Remark 2.6. The operator T_a defined by $(T_a x)(t) := x(t+a)$ for a fixed $a \in \mathbb{R}$ leaves $PAA(\mathbb{R}, \mathbb{X})$ invariant.

Let us now discuss conditions which do ensure the pseudo almost automorphy of the convolution function $f \star h$ of f with h where f is pseudo almost automorphic and h is a Lebesgue mesurable function satisfying additional assumptions.

Let $f : \mathbb{R} \to X$ and $h : \mathbb{R} \to \mathbb{R}$; the convolution function (if it does exist) of f with h denoted $f \star h$ is defined by:

$$(f \star h)(t) := \int_{\mathbb{R}} f(\sigma)h(t-\sigma)d\sigma = \int_{\mathbb{R}} f(t-\sigma)h(\sigma)d\sigma = (h \star f)(t), \quad \text{for all } t \in \mathbb{R}.$$

Hence, if $f \star h$ is well-defined, then $f \star h = h \star f$.

Let $\varphi \in L^1$ and $\lambda \in \mathbb{C}$. It is well-known that the operator $A_{\varphi,\lambda}$ defined by

$$A_{\varphi,\lambda} u = \lambda u + \varphi \star u \tag{2.1}$$

acts continuously in $BC(\mathbb{R}, \mathbb{X})$ i.e., there exists $K > 0$ such that

$$\|A_{\varphi,\lambda} u\|_{BC(\mathbb{R},\mathbb{X})} \leq K \|u\|_{BC(\mathbb{R},\mathbb{X})}, \forall u \in BC(\mathbb{R}, \mathbb{X}) \tag{2.2}$$

Moreover $A_{\varphi,\lambda}$ leaves $BC(\mathbb{R}, \mathbb{X})$ invariant.

Now denote $\mathcal{M} := \{PAP(\mathbb{R}, \mathbb{X}), PAA(\mathbb{R}, \mathbb{X})\}$ where $PAP(\mathbb{R}, X)$ is the Banach space of all pseudo almost periodic functions $f : \mathbb{R} \to \mathbb{X}$. Then we have.

Theorem 2.7. *For* $\Omega \in \mathcal{M}$,

$$A_{\varphi,\lambda}(\Omega) \subset \Omega.$$

Proof. . It is an immediate consequence of the remarks above. ☐

Application: A Volterra-like equation

Consider the equation

$$x(t) = g(t) + \int_{-\infty}^{+\infty} a(t-\sigma)x(\sigma)d\sigma, \ t \in \mathbb{R}, \tag{2.3}$$

where $g : \mathbb{R} \to \mathbb{R}$ is a continuous function and $a \in L^1(\mathbb{R})$.

Theorem 2.8. *Suppose* $g \in PAA(\mathbb{R}, \mathbb{X})$ *and* $\|a\|_{L^1} < 1$. *Then* (2.3) *above has a unique pseudo almost automorphic solution.*

Proof. It is clear that the operator

$$x \in PAA(\mathbb{R}, \mathbb{X}) \to \int_{-\infty}^{+\infty} a(t-\sigma)x(\sigma)d\sigma \in PAA(\mathbb{R}, \mathbb{X})$$

is well-defined. Now consider $\Gamma : PAA(\mathbb{R},\mathbb{X}) \to PAA(\mathbb{R},\mathbb{X})$ such that

$$(\Gamma x)(t) = g(t) + \int_{-\infty}^{+\infty} a(t-\sigma)x(\sigma)d\sigma, \quad t \in \mathbb{R}.$$

We can easily show that

$$\|(\Gamma x) - (\Gamma y)\| \le \|a\|_{L^1}\|x-y\|.$$

The conclusion is immediate by the principle of contraction. □

3 Main results

This section is devoted to the proof of the main result of the paper, that is, the existence and uniqueness of an almost automorphic (mild) solution to Eq. (1.4). For that we need to establish a few preliminary results.

Definition 3.1. A function $u \in BC(\mathbb{R},\mathbb{X})$ is said to be a mild solution to Eq. (1.4) if the function $s \to AT(t-s)f(s,u(s))$ is integrable on $(-\infty,t)$ for each $t \in \mathbb{R}$ and

$$u(t) = -f(t,u(t)) - \int_{-\infty}^{t} AT(t-s)f(s,u(s))ds$$

for each $t \in \mathbb{R}$.

We now make the following assumptions.

(H.1) The operator A is the infinitesimal generator of an exponentially stable semigroup $(T(t))_{t\ge 0}$ such that there exist constants $M > 0$ and $\delta > 0$ with

$$\|T(t)\|_{B(\mathbb{X})} \le Me^{-\delta t}, \quad \forall t \ge 0.$$

Furthermore, the function $\sigma \to AT(\sigma)$ defined from $(0,\infty)$ into $B(\mathbb{X})$ is strongly (Lebesgue) measurable and there exist a function $\gamma : (0,\infty) \to [0,\infty)$ such that $\sup_{s\ge s_0}\gamma(s) < \infty$ for any $s_0 > 0$, and a constant $\omega > 0$ with $\rho := \displaystyle\int_0^{\infty} e^{-\omega s}\gamma(s)ds < \infty$ such that

$$\|AT(s)\|_{B(\mathbb{X})} \le e^{-\omega s}\,.\,\gamma(s), \quad s > 0.$$

(H.2) The function $f : \mathbb{R}\times\mathbb{X} \mapsto \mathbb{X}$, $(t,u) \mapsto f(t,u)$ is jointly continuous and

$$\|f(t,u) - f(t,v)\|_{\mathbb{X}} \le k(t)\,.\,\|u-v\|, \quad \text{and}$$

for all $t \in \mathbb{R}$, and $\forall u,v \in \mathbb{X}$. Here $k \in L^1(\mathbb{R},\mathbb{R}^+)$.

(H.3) $f = g + \psi \in PAA(\mathbb{R}\times\mathbb{X},\mathbb{X})$, where g and ψ are the principal and the ergodic terms of f respectively and $f(t,u)$ and $g(t,u)$ are uniformly continuous on every bounded subset $K \subset \mathbb{X}$ uniformly in $t \in \mathbb{R}$.

Lemma 3.2. *Suppose that assumptions (H.1)-(H.2)-(H.3) hold. Define the nonlinear operator* Λ_1 *by: For each* $\xi \in PAA(\mathbb{X})$,

$$(\Lambda_1\xi)(t) = \int_{-\infty}^{t} AT(t-s)f(s,\xi(s))ds$$

Then Λ_1 *maps* $PAA(\mathbb{X})$ *into itself.*

Proof. Set h defined by: $h(.) = f(.,\xi(.))$. Since $h \in PAA(\mathbb{R},\mathbb{X})$ using [6, Theorem 2.4] with assumption (H.3), we can write $h = \beta + \phi$ where β is the principal part and ϕ the ergodic term of h. Using the same argument as in [11], we can prove that $t \mapsto \int_{-\infty}^{t} AT(t-s)\beta(s)ds$ is in $AA(\mathbb{X})$. Now, set:

$$\nu(t) = -\int_{-\infty}^{t} AT(t-s)\phi(s)ds.$$

We have:

$$\frac{1}{2T}\int_{-T}^{T}\|\nu(t)\|_X dt \leq \frac{1}{2T}\int_{-T}^{T}\int_{-\infty}^{t}\|A(t-s)\phi(s)\|_{\mathbb{X}}dsdt$$
$$\leq \frac{1}{2T}\int_{-T}^{T}\int_{-\infty}^{t} e^{-\omega(t-s)}\gamma(t-s)\|\phi(s)\|_{\mathbb{X}}dsdt.$$

Let's write:

$$\frac{1}{2T}\int_{-T}^{T}\int_{-\infty}^{t} e^{-\omega(t-s)}\gamma(t-s)\|\phi(s)\|_{\mathbb{X}}dsdt = I_1 + I_2,$$

where:

$$I_1 = \frac{1}{2T}\int_{-T}^{T}\int_{-\infty}^{-T} e^{-\omega(t-s)}\gamma(t-s)\|\phi(s)\|_{\mathbb{X}}dsdt$$

and

$$I_2 = \frac{1}{2T}\int_{-T}^{T}\int_{-T}^{t} e^{-\omega(t-s)}\gamma(t-s)\|\phi(s)\|_{\mathbb{X}}dsdt.$$

We prove know that $I_1 \to 0$ and $I_2 \to 0$ as $T \to \infty$.

Indeed, for I_1, let $s_0 > 0$ and set $M(s_0) = \sup_{s \geq s_0}\gamma(s)$, and $K = \sup_{t\in\mathbb{R}}\|\phi(t)\|_{\mathbb{X}}$. We have:

$$I_1 \leq K\frac{1}{2T}\int_{-T}^{T}\int_{-\infty}^{-T} e^{-\omega(t-s)}\gamma(t-s)dsdt = \frac{K}{2T}\int\int_{D} e^{-\omega(t-s)}\gamma(t-s)dsdt,$$

where $D = \{(s,t) \in \mathbb{R}^2, |t| \leq T, \quad s \leq -T\}$. We introduce also:

$$D_1 = \{(s,t) \in D, t-s \geq s_0\}, \quad D_2 = D \setminus D_1.$$

We have:

$$\int\int_{D_1} e^{-\omega(t-s)}\gamma(t-s)dsdt \leq M(s_0)\int\int_{D_1} e^{-\omega(t-s)}dsdt$$
$$\leq M(s_0)\int\int_{D} e^{-\omega(t-s)}dsdt = \frac{M(s_0)e^{-\omega T}}{\omega}\int_{-T}^{T} e^{-\omega t}dt \leq 2T\frac{M(s_0)e^{-\omega T}}{\omega}.$$

Moreover,

$$\int\int_{D_2} e^{-\omega(t-s)}\gamma(t-s)dsdt \le \int_{-T-s_0}^{-T}\int_{-T}^{s+s_0} e^{-\omega(t-s)}\gamma(t-s)dtds$$
$$\le \int_{-T-s_0}^{-T}\int_{-T-s}^{s_0} e^{-\omega\sigma}\gamma(\sigma)d\sigma ds$$
$$\le \int_{-T-s_0}^{-T}\int_{0}^{s_0} e^{-\omega\sigma}\gamma(\sigma)d\sigma ds$$
$$\le s_0\int_0^{s_0} e^{-\omega\sigma}\gamma(\sigma)d\sigma.$$

So, for any $T \ge 1$, we have:

$$I_1 \le \frac{K}{2T}\left(2T\frac{M(s_0)e^{-\omega T}}{\omega} + s_0\int_0^{s_0} e^{-\omega\sigma}\gamma(\sigma)d\sigma\right)$$
$$\le K\left(e^{-\omega T}\frac{M(s_0)}{\omega} + s_0\int_0^{s_0} e^{-\omega\sigma}\gamma(\sigma)d\sigma\right).$$

Let $\epsilon > 0$. We can find $s_0 > 0$ such that $Ks_0\int_0^{s_0} e^{-\omega\sigma}\gamma(\sigma)d\sigma < \varepsilon/2$. Let us take such an s_0. After, for T sufficiently large, $Ke^{-\omega T}\frac{M(s_0)}{\omega} < \epsilon/2$, and so, for sufficiently large T, $I_1 \le \epsilon$.
Now, we consider I_2. We have:

$$I_2 = \frac{1}{2T}\int_{-T}^{T}\|\phi(s)\|_{\mathbb{X}}ds\int_s^T e^{-\omega(t-s)}\gamma(t-s)dt$$
$$\le \frac{1}{2T}\int_{-T}^{T}\|\phi(s)\|_{\mathbb{X}}ds\int_0^{T-s} e^{-\omega\sigma}\gamma(\sigma)d\sigma$$
$$\le \rho\frac{1}{2T}\int_{-T}^{T}\|\phi(s)\|_{\mathbb{X}}ds \to 0 \qquad \text{as} \qquad T\to\infty.$$

□

Now we are ready to state and prove the following.

Theorem 3.3. *Suppose that assumptions (H.1)-(H.2) hold. Then Eq. (1.4) has a unique pseudo almost automorphic (mild) solution if*

Proof. Define the nonlinear operator $\Gamma : AA(\mathbb{X}) \mapsto AA(\mathbb{X})$ by:

$$\Gamma(u) : t \mapsto -f(t,u(t)) - \int_{-\infty}^{t} AT(t-s)f(s,u(s))ds.$$

We have:

$$\|\Gamma(u)(t)-\Gamma(v)(t)\|_{\mathbb{X}} \le \|f(t,u(t))-f(t,v(t))\|_{\mathbb{X}} + \int_{-\infty}^{t}\|AT(t-s)(f(s,u(s))-f(s,v(s)))\|_{\mathbb{X}}ds$$

$$\leq k(t)\|u(t)-v(t)\|_{\mathbb{X}} + \int_{-\infty}^{t} e^{-\omega(t-s)}\gamma(t-s)k(s)\|u(x)-v(s)\|_{\mathbb{X}}ds$$
$$\leq \left[k(t) + \int_{-\infty}^{t} e^{-\omega(t-s)}\gamma(t-s)k(s)ds\right]\|u-v\|_{\infty}$$
$$\leq (1+\rho)\|k\|_{\infty}\|u-v\|_{\infty}.$$

So, we obtain:

$$\|\Gamma(u)-\Gamma(v)\|_{\infty} \leq (1+\rho)\|k\|_{\infty}\|u-v\|_{\infty},$$

and we can conclude using the Banach's fixed point principle.

□

Acknowledgements. This paper was written when the second author was visiting the "Laboratoire Marin Mersenne" of the University of Paris 1 Panthéon-Sorbonne in June 2008. He is grateful to Professors Blot and Pennequin for their invitation.

Received: June 2008. Revised: August 2008.

References

[1] S. Bochner, *A New Approach to Almost Periodicity*, Proc. Nat. Acad. Sci. U.S.A., **48**(1962), pp. 2039–2043.

[2] T. Diagana, *Existence of p-almost automorphic mild solutions to some abstract differential equations*, Intern. J. Evol. Equ., **1**(1) (2005), 57–67.

[3] T. Diagana and G.M. N'Guérékata, *Almost automorphic solutions to some classes of partial evolution equations*, Appl. Math. Lett., **20**(2007), no. 4, 462–466.

[4] T. Diagana, G.M. N'Guérékata and N.V. Minh, *Almost Automorphic Solutions of Evolution Equations*, Proc. Amer. Math. Soc., **132**(2004), no. 11, pp. 3289–3298.

[5] J.A. Goldstein and G.M. N'Guérékata, *Almost Automorphic Solutions of Semilinear Evolution Equations*, Proc. Amer. Math. Soc., **133**(2005), no. 8, pp. 2401–2408.

[6] J. Liang, J. Zhang and T.-J. Xiao, *Composition of pseudo almost automorphic and asymptotically almost automorphic functions*, J. Math. Anal. Appl., J. Math. Anal. Appl., **340**(2008), No.2, 1493–1499.

[7] J. Liang, G.M. N'Guérékata, T.-J. Xiao and J. Zhang, *Some properties of pseudo almost automorphic functions and applications to abstract differential equations*, Nonlinear Analysis TMA, (to appear).

[8] T.-J. Xiao, J. Liang and J. Zhang, *Pseudo almost automorphic functions to semilinear differential equations in Banach spaces*, Semigroup Forum, **76**(2008), No.3, 518–524.

[9] A. Lunardi, *Analytic semigroups and optimal regularity in parabolic problems*, PNLDE Vol.16, Birkhäauser Verlag, Basel, 1995.

[10] G.M. N'Guérékata, *Quelques remarques sur les fonctions asymptotiquement presque automorphes*, Ann. Sci. math. Québec, **7**(2) (1983), 185–191.

[11] G.M. N'Guérékata, *Existence and Uniqueness of Almost Automorphic Mild Solutions to Some Semilinear Abstract Differential Equations*, Semigroup Forum, **69**(2004), pp. 80–86.

[12] G.M. N'Guérékata, *Almost Automorphic Functions and Almost Periodic Functions in Abstract Spaces*, Kluwer Academic / Plenum Publishers, New York-London-Moscow, 2001.

[13] G.M. N'Guérékata, *Topics in Almost Automorphy*, Springer-Verlag, New York, 2005.

[14] G.M. N'Guérékata and A. Pankov, *Stepanov-like almost automorphic functions and monotone evolution equations*, Nonlinear Analysis, **68**(9) (2008), 2658–2667.

[15] G.M. N'Guérékata and D. Pennequin, *Pseudo almost automorphic solutions for hyperbolic semilinear evolution equations in intermediate Banach spaces*, Dyn. Contin. Disc. Impuls. Syst., Ser. A, Proceedings of the 6th. Conference on Differential Equation and Dynamical Systems, Baltimore, MD, USA, May 22–26, 2008, (to appear).

[16] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer-Verlag, 1983.

[17] M. Renardy and R.C. Rogers, *An Introduction to Partial Differential Equations*, Texts in Appl. Math. **13**, Springer-Verlag, New York-Berlin- Heidelberg-London-Paris, 1992.

CUBO A Mathematical Journal
Vol.10, Nº03, (171–194). October 2008

Rosso-Yamane Theorem on PBW Basis of $U_q(A_N)$*

Yuqun Chen, Hongshan Shao
School of Mathematical Sciences, South China Normal University,
Guangzhou 510631, P. R. China
emails: `yqchen@scnu.edu.cn, shaohongshan118@163.com`

AND

K.P. Shum
Department of Mathematics, The University of Hong Kong,
Pokfulam Road, Hong Kong, China (SAR)
email: `kpshum@maths.hku.hk`

ABSTRACT

Let $U_q(A_N)$ be the Drinfeld-Jimbo quantum group of type A_N. In this paper, by using Gröbner-Shirshov bases, we give a simple (but not short) proof of the Rosso-Yamane Theorem on PBW basis of $U_q(A_N)$.

RESUMEN

Sea $U_q(A_N)$ el grupo cuántico de Drinfel-Jimbo de tipo A_N. En este artículo, usando bases de Gröbner-Shirshov damos una demonstración simple (pero no corta) del Teorema de Rosso-Yamane sobre bases PBW de $U_q(A_N)$.

Key words and phrases: *Quantum group, Quantum enveloping algebra, Gröbner-Shirshov basis.*

Math. Subj. Class.: *20G42, 16S15, 13P10.*

*Supported by the NNSF of China (No.10771077) and the NSF of Guangdong Province (No.06025062).

1 Introduction

Since any algebra (commutative, associative, Lie), as well as any module over an algebra, can be presented by generators and defining relations, it is important to have a general method to deal with these presentations. Such a method now exists and is called the Gröbner bases method (due to B. Buchberger [18], [19]), or standand bases method (due to H. Hironaka [21]), or Gröbner-Shirshov bases method (due to A. I. Shirshov [35]). The original Shirshov's paper [35] is for Lie algebra presentations, but it can be easily adopted for associative algebra presentations as well, see L. A. Bokut [3] and G. Bergman [1].

Let, for example, $L = Lie(X|[x_ix_j] - \sum \alpha_{ij}^k x_k,\ i > j,\ x_i, x_j, x_k \in X)$ be a Lie algebra over a field (or a commutative ring) k presented by a k-basis X and the multiplication table. Then $S = \{[x_ix_j] - \sum \alpha_{ij}^k x_k|\ i > j,\ x_i, x_j, x_k \in X\}$ is a Gröbner-Shirshov basis (subset) of the free Lie algebra $Lie(X)$ over k. On the other hand, the universal enveloping algebra $U(L) = k\langle X|x_ix_j - x_jx_i - \sum \alpha_{ij}^k x_k,\ i > j,\ x_i, x_j, x_k \in X\rangle$ is the associative algebra presented by the same set X and the defining relations $S^{(-)}$ (we rewrite S using $[xy] = xy - yx$). There is a general but not difficult result that for any $S \subset Lie(X)$, S is a Gröbner-Shirshov basis in the sense of Lie algebras if and only if $S^{(-)} \subset k\langle X\rangle$ is a Gröbner-Shirshov basis in the sense of associative algebras (see, for example, [9] and [7]). This means that in our case, $S^{(-)}$ is a Gröbner-Shirshov basis (subset) in $k\langle X\rangle$. By Composition-Diamond lemma (see below), the S-irreducible words on X, $Irr(S) = \{x_{i_1} \ldots x_{i_k},\ i_1 \leq \ldots \leq i_k,\ k \geq 0\}$ form a k-basis of $U(L)$. This is a conceptional proof of the PBW-Theorem by using Gröbner-Shirshov bases theory.

There are many results on Gröbner-Shirshov bases for associative and Lie algebras, as well as for semigroups, groups, conformal algebras, dialgebras, and so on, see, for example, surveys [14], [15], [25] and [8]. Let us mention those for simple Lie algebras and Lie superalgebras via Serre's presentations ([10], [11], [12], [13], [9]), for modules over simple Lie algebras and Iwahori-Hecke algebras ([23], [24], [25]), for Kac-Moody algebras of types $A_n^{(1)}$, $B_n^{(1)}$, $C_n^{(1)}$, $D_n^{(1)}$ ([31], [32], [33]), for Coxeter groups ([17]), for braid groups via Artin-Burau, Artin-Garside and Briman-Ko-Lee presentations ([4], [5] and [6]).

Drinfeld-Jimbo ([20], [22]) presentations for quantized enveloping algebras $U_q(g)$, where g is a semisimple Lie algebra, are a natural source of associative presentations. M. Rosso [34] and I. Yamane [36] found the PBW-basis of $U_q(A_N)$. G. Lusztig [29] and [30], and M. Kashiwara [26] and [27] found the bases of $U_q(g)$ for any semisimple algebra g, as well as for their representations. Their approach work equally well for quantized enveloping algebras associated with arbitrary symmetrizable Cartan matrix, not just those corresponding to finite dimensional Lie algebras. V. K. Kharchenko [28] found the approach to linear bases of quantized enveloping algebras via the so called character Hopf algebras.

It the paper [16], Gröbner-Shirshov bases approach was applied to study $U_q(g)$ for any symmetrizable Cartan matrix. Using this approach, they got a new proof of the triangular decomposition of $U_q(g)$ (see, for example, Jantzen [37]). For $U_q(A_N)$, it was proved by Bokut and Malcolmson [16] that the Jimbo relations (see [36]) of $U_q^+(A_N)$ constitute a Gröbner-Shirshov basis of $U_q^+(A_N)$ in Jimbo generators $x_{ij}, 1 \leq i, j \leq N+1$ (see below).

In this paper, we give an elementary proof that Jimbo relations S is a Gröbner-Shirshov basis of $U_q^+(A_N)$. For such a purpose, we just check all possible compositions of polynomials from S and proved that all them can be resolved. Also in §1 in this paper, we are giving necessary definitions and Composition-Diamond lemma following Shirshov [35].

2 Preliminaries

We first cite some concepts and results from the literature which are related to the Gröbner-Shirshov bases for associative algebras.

Let k be a field, $k\langle X\rangle$ the free associative algebra over k generated by X and X^* the free monoid generated by X, where the empty word is the identity which is denoted by 1. For a word $w \in X^*$, we denote the length of w by $deg(w)$. Let X^* be a well ordered set. Let $f \in k\langle X\rangle$ with the leading word $\bar{f}$. Then we call f monic if $\bar{f}$ has coefficient 1.

Definition 2.1. *([35], see also [2], [3]) Let f and g be two monic polynomials in $k\langle X\rangle$ and $<$ a well ordering on X^*. Then, there are two kinds of compositions:*

(i) If w is a word such that $w = \bar{f}b = a\bar{g}$ for some $a, b \in X^*$ with $deg(\bar{f}) + deg(\bar{g}) > deg(w)$, then the polynomial $(f, g)_w = fb - ag$ is called the intersection composition of f and g with respect to w.

(ii) If $w = \bar{f} = a\bar{g}b$ for some $a, b \in X^*$, then the polynomial $(f, g)_w = f - agb$ is called the inclusion composition of f and g with respect to w.

Definition 2.2. *([2], [3], cf. [35]) Let $S \subset k\langle X\rangle$ such that every $s \in S$ is monic. Then the composition $(f, g)_w$ is called trivial modulo (S, w) if $(f, g)_w = \sum \alpha_i a_i s_i b_i$, where each $\alpha_i \in k$, $a_i, b_i \in X^*$, $s_i \in S$ and $\overline{a_i s_i b_i} < w$. If this is the case, then we write*

$$(f, g)_w \equiv 0 \quad mod(S, w).$$

In general, for $p, q \in k\langle X\rangle$, we write $p \equiv q \quad mod(S, w)$ which means that $p - q = \sum \alpha_i a_i s_i b_i$, where each $\alpha_i \in k, a_i, b_i \in X^*$, $s_i \in S$ and $\overline{a_i s_i b_i} < w$.

Definition 2.3. *([2], [3], cf. [35]) We call the set S with respect to the well ordering $<$ a Gröbner-Shirshov set (basis) in $k\langle X\rangle$ if any composition of polynomials in S is trivial modulo S.*

If a subset S of $k\langle X\rangle$ is not a Gröbner-Shirshov basis, then we can add to S all nontrivial compositions of polynomials of S, and by continuing this process (maybe infinitely) many times, we eventually obtain a Gröbner-Shirshov basis S^c. Such a process is called the Shirshov algorithm.

A well ordering $>$ on X^* is called a monomial order if it is compatible with the multiplication of words, that is, for $u, v \in X^*$, we have

$$u > v \Rightarrow w_1 u w_2 > w_1 v w_2, \ for\ all\ w_1,\ w_2 \in X^*.$$

A standard example of monomial order on X^* is the deg-lex order to compare two words first by degree and then lexicographically, where X is a well ordered set.

The following lemma was first proved by Shirshov [35] for free Lie algebras (with deg-lex order) in 1962 (see also Bokut [2]). In 1976, Bokut [3] specialized the approach of Shirshov to associative algebras (see also Bergman [1]). For the case of commutative polynomials, this lemma is known as the Buchberger's Theorem in [18] and [19].

Lemma 2.4. *(Composition-Diamond Lemma) Let k be a field, $k\langle X|S\rangle = k\langle X\rangle/Id(S)$ and $<$ a monomial order on X^*, where $Id(S)$ is the ideal of $k\langle X\rangle$ generated by S. Then the following statements are equivalent:*

(i) S is a Gröbner-Shirshov basis.

(ii) $f \in Id(S) \Rightarrow \bar{f} = a\bar{s}b$ for some $s \in S$ and $a, b \in X^*$.

(iii) $Irr(S) = \{u \in X^* | u \neq a\bar{s}b, s \in S, a, b \in X^*\}$ is a basis of the algebra $k\langle X|S\rangle$. □

3 Rosso-Yamane theorem on PBW basis of $U_q(A_N)$

Let k be a field, $A = (a_{ij})$ an integral symmetrizable $N \times N$ Cartan matrix so that $a_{ii} = 2$, $a_{ij} \leq 0$ $(i \neq j)$ and there exists a diagonal matrix D with diagonal entries d_i which are nonzero integers such that the product DA is symmetric. Let q be a nonzero element of k such that $q^{4d_i} \neq 1$ for each i. Then the quantum enveloping algebra is (see [20], [22])

$$U_q(A) = k\langle X \cup H \cup Y | S^+ \cup K \cup T \cup S^- \rangle,$$

where

$$\begin{aligned}
X &= \{x_i\},\\
H &= \{h_i^{\pm 1}\},\\
Y &= \{y_i\},\\
S^+ &= \{\sum_{\nu=0}^{1-a_{ij}} (-1)^\nu \begin{pmatrix} 1-a_{ij} \\ \nu \end{pmatrix}_t x_i^{1-a_{ij}-\nu} x_j x_i^\nu,\ where\ i \neq j,\ t = q^{2d_i}\},\\
S^- &= \{\sum_{\nu=0}^{1-a_{ij}} (-1)^\nu \begin{pmatrix} 1-a_{ij} \\ \nu \end{pmatrix}_t y_i^{1-a_{ij}-\nu} y_j y_i^\nu,\ where\ i \neq j,\ t = q^{2d_i}\},\\
K &= \{h_i h_j - h_j h_i, h_i h_i^{-1} - 1, h_i^{-1} h_i - 1, x_j h_i^{\pm 1} - q^{\mp d_i a_{ij}} h^{\pm 1} x_j, h_i^{\pm 1} y_j - y_j h^{\pm 1}\},\\
T &= \{x_i y_j - y_j x_i - \delta_{ij} \frac{h_i^2 - h_i^{-2}}{q^{2d_i} - q^{-2d_i}}\} \quad \text{and}\\
\begin{pmatrix} m \\ n \end{pmatrix}_t &= \begin{cases} \prod_{i=1}^{n} \frac{t^{m-i+1} - t^{i-m-1}}{t^i - t^{-i}} & m > n > 0,\\ 1 & n = 0 \quad \text{or} \quad m = n. \end{cases}
\end{aligned}$$

Let

$$A = A_N = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & -1 & \cdots & 0 \\ 0 & -1 & 2 & \cdots & 0 \\ . & . & . & . & . \\ 0 & 0 & 0 & \cdots & 2 \end{pmatrix} \quad \text{and} \quad q^8 \neq 1.$$

It is reminded that in this case, the diagonal matrix D is identity.

We introduce some new variables defined by Jimbo (see [36]) which generate $U_q(A_N)$:

$$\widetilde{X} = \{x_{ij}, 1 \leq i < j \leq N+1\},$$

where

$$x_{ij} = \begin{cases} x_i & j = i+1,\\ q x_{i,j-1} x_{j-1,j} - q^{-1} x_{j-1,j} x_{i,j-1} & j > i+1. \end{cases}$$

We now order the set $\widetilde{X}$ in the following way.

$$x_{mn} > x_{ij} \Longleftrightarrow (m,n) >_{lex} (i,j).$$

Let us recall from Yamane [36] the following notation:

$$\begin{aligned} C_1 &= \{((i,j),(m,n))|i=m<j<n\}, \\ C_2 &= \{((i,j),(m,n))|i<m<n<j\}, \\ C_3 &= \{((i,j),(m,n))|i<m<j=n\}, \\ C_4 &= \{((i,j),(m,n))|i<m<j<n\}, \\ C_5 &= \{((i,j),(m,n))|i<j=m<n\}, \\ C_6 &= \{((i,j),(m,n))|i<j<m<n\}. \end{aligned}$$

Let the set $\widetilde{S}^+$ consist of Jimbo relations:

$$\begin{aligned} x_{mn}x_{ij} &- q^{-2}x_{ij}x_{mn} && ((i,j),(m,n)) \in C_1 \cup C_3, \\ x_{mn}x_{ij} &- x_{ij}x_{mn} && ((i,j),(m,n)) \in C_2 \cup C_6, \\ x_{mn}x_{ij} &- x_{ij}x_{mn} + (q^2-q^{-2})x_{in}x_{mj} && ((i,j),(m,n)) \in C_4, \\ x_{mn}x_{ij} &- q^2x_{ij}x_{mn} + qx_{in} && ((i,j),(m,n)) \in C_5. \end{aligned}$$

It is easily seen that $U_q^+(A_N) = k\langle \widetilde{X} | \widetilde{S^+} \rangle$.

The following theorem is from [16].

Theorem 3.1. *([16] Theorem 4.1) Let the notation be as before. Then, with the deg-lex order on $\widetilde{X}^*$, $\widetilde{S}^+$ is a Gröbner-Shirshov basis for $k\langle \widetilde{X} | \widetilde{S^+} \rangle = U_q^+(A_N)$.*

Proof. We will prove that all compositions in $\widetilde{S}^+$ are trivial modulo $\widetilde{S}^+$. We consider the following cases.

Case 1. $f = x_{mn}x_{ij} - q^{-2}x_{ij}x_{mn},\ g = x_{ij}x_{kl} - q^{-2}x_{kl}x_{ij},\ w = x_{mn}x_{ij}x_{kl}$.

In the case, we have

$$(f,g)_w = -q^{-2}x_{ij}x_{mn}x_{kl} + q^{-2}x_{mn}x_{kl}x_{ij}.$$

There are four subcases to consider.

	$((i,j),(m,n)) \in C_1$	$((i,j),(m,n)) \in C_3$
$((k,l),(i,j)) \in C_1$	1.1. $((k,l),(m,n)) \in C_1$	1.3. $((k,l),(m,n)) \in C_4, C_5$ *or* C_6
$((k,l),(i,j)) \in C_3$	1.2. $((k,l),(m,n)) \in C_4$	1.4. $((k,l),(m,n)) \in C_3$

1.1. $((i,j),(m,n)) \in C_1,\ ((k,l),(i,j)) \in C_1$ and $((k,l),(m,n)) \in C_1$.

Then, we have

$$\begin{aligned} (f,g)_w &\equiv -q^{-4}x_{ij}x_{kl}x_{mn} + q^{-4}x_{kl}x_{mn}x_{ij} \\ &\equiv -q^{-6}x_{kl}x_{ij}x_{mn} + q^{-6}x_{kl}x_{ij}x_{mn} \\ &\equiv 0. \end{aligned}$$

1.2. $((i,j),(m,n)) \in C_1,\ ((k,l),(i,j)) \in C_3$ and $((k,l),(m,n)) \in C_4$.

Then, we have $(i,j) = (m,l),\ ((k,n),(i,j)) \in C_2$ and

$$\begin{aligned}(f,g)_w &\equiv -q^{-2}x_{ij}[x_{kl}x_{mn} - (q^2 - q^{-2})x_{kn}x_{ml}] + q^{-2}[x_{kl}x_{mn} - (q^2 - q^{-2})x_{kn}x_{ml}]x_{ij} \\ &\equiv -q^{-4}x_{kl}x_{ij}x_{mn} + q^{-2}(q^2 - q^{-2})x_{kn}x_{ij}x_{ml} \\ &\quad +q^{-4}x_{kl}x_{ij}x_{mn} - q^{-2}(q^2 - q^{-2})x_{kn}x_{ij}x_{ml} \\ &\equiv 0.\end{aligned}$$

1.3. $((i,j),(m,n)) \in C_3,\ ((k,l),(i,j)) \in C_1$ and $((k,l),(m,n)) \in C_4,\ C_5$ or C_6.

1.3.1. If $((k,l),(m,n)) \in C_4$ $(m < l)$, then $(k,n) = (i,j),\ ((i,j),(m,l)) \in C_2$ and

$$\begin{aligned}(f,g)_w &\equiv -q^{-2}x_{ij}[x_{kl}x_{mn} - (q^2 - q^{-2})x_{kn}x_{ml}] + q^{-2}[x_{kl}x_{mn} - (q^2 - q^{-2})x_{kn}x_{ml}]x_{ij} \\ &\equiv -q^{-4}x_{kl}x_{ij}x_{mn} + q^{-2}(q^2 - q^{-2})x_{ij}x_{kn}x_{ml} + q^{-4}x_{kl}x_{ij}x_{mn} \\ &\quad -q^{-2}(q^2 - q^{-2})x_{kn}x_{ij}x_{ml} \\ &\equiv 0.\end{aligned}$$

1.3.2. If $((k,l),(m,n)) \in C_5$ $(m = l)$, then $(k,n) = (i,j)$ and

$$\begin{aligned}(f,g)_w &\equiv -q^{-2}x_{ij}(q^2x_{kl}x_{mn} - qx_{kn}) + q^{-2}(q^2x_{kl}x_{mn} - qx_{kn})x_{ij} \\ &\equiv -x_{ij}x_{kl}x_{mn} + q^{-1}x_{ij}x_{kn} + x_{kl}x_{mn}x_{ij} - q^{-1}x_{kn}x_{ij} \\ &\equiv -q^{-2}x_{kl}x_{ij}x_{mn} + q^{-2}x_{kl}x_{ij}x_{mn} \\ &\equiv 0.\end{aligned}$$

1.3.3. If $((k,l),(m,n)) \in C_6$ $(m > l)$, then

$$\begin{aligned}(f,g)_w &\equiv -q^{-2}x_{ij}x_{kl}x_{mn} + q^{-2}x_{kl}x_{mn}x_{ij} \\ &\equiv -q^{-4}x_{kl}x_{ij}x_{mn} + q^{-4}x_{kl}x_{ij}x_{mn} \\ &\equiv 0.\end{aligned}$$

1.4. $((i,j),(m,n)) \in C_3,\ ((k,l),(i,j)) \in C_3$ and $((k,l),(m,n)) \in C_3$.

This case is similar to 1.1.

Case 2. $f = x_{mn}x_{ij} - q^{-2}x_{ij}x_{mn},\ g = x_{ij}x_{kl} - x_{kl}x_{ij},\ w = x_{mn}x_{ij}x_{kl}$.

In the case, we have

$$(f,g)_w = -q^{-2}x_{ij}x_{mn}x_{kl} + x_{mn}x_{kl}x_{ij}.$$

There are also four subcases to consider.

	$((i,j),(m,n)) \in C_1$	$((i,j),(m,n)) \in C_3$
$((k,l),(i,j)) \in C_2$	2.1. $((k,l),(m,n)) \in C_2, C_3$ *or* C_4	2.3. $((k,l),(m,n)) \in C_2$
$((k,l),(i,j)) \in C_6$	2.2. $((k,l),(m,n)) \in C_6$	2.4. $((k,l),(m,n)) \in C_6$

2.1. $((i,j),(m,n)) \in C_1,\ ((k,l),(i,j)) \in C_2$ and $((k,l),(m,n)) \in C_2, C_3$ or C_4.

2.1.1. If $((k,l),(m,n)) \in C_2\ (n < l)$, then

$$\begin{aligned}(f,g)_w &\equiv -q^{-2}x_{ij}x_{kl}x_{mn} + x_{kl}x_{mn}x_{ij}\\ &\equiv -q^{-2}x_{kl}x_{ij}x_{mn} + q^{-2}x_{kl}x_{ij}x_{mn}\\ &\equiv 0.\end{aligned}$$

2.1.2. If $((k,l),(m,n)) \in C_3\ (n = l)$, then

$$\begin{aligned}(f,g)_w &\equiv -q^{-4}x_{ij}x_{kl}x_{mn} + q^{-2}x_{kl}x_{mn}x_{ij}\\ &\equiv -q^{-4}x_{kl}x_{ij}x_{mn} + q^{-4}x_{kl}x_{ij}x_{mn}\\ &\equiv 0.\end{aligned}$$

2.1.3. If $((k,l),(m,n)) \in C_4\ (n > l)$, then $((k,n),(i,j)) \in C_2,\ ((i,j),(m,l)) \in C_1$ and

$$\begin{aligned}(f,g)_w &\equiv -q^{-2}x_{ij}[x_{kl}x_{mn} - (q^2-q^{-2})x_{kn}x_{ml}] + [x_{kl}x_{mn} - (q^2-q^{-2})x_{kn}x_{ml}]x_{ij}\\ &\equiv -q^{-2}x_{kl}x_{ij}x_{mn} + q^{-2}(q^2-q^{-2})x_{kn}x_{ij}x_{ml} + q^{-2}x_{kl}x_{ij}x_{mn}\\ &\quad -q^{-2}(q^2-q^{-2})x_{kn}x_{ij}x_{ml}\\ &\equiv 0.\end{aligned}$$

For the cases 2.2, 2.3 and 2.4, the proofs are similar to 2.1.1.

Case 3. $f = x_{mn}x_{ij} - q^{-2}x_{ij}x_{mn},\ g = x_{ij}x_{kl} - x_{kl}x_{ij} + (q^2-q^{-2})x_{kj}x_{il},\ w = x_{mn}x_{ij}x_{kl}$. In the case, we have

$$(f,g)_w = -q^{-2}x_{ij}x_{mn}x_{kl} + x_{mn}x_{kl}x_{ij} - (q^2-q^{-2})x_{mn}x_{kj}x_{il}.$$

There are two subcases to consider.

	$((i,j),(m,n)) \in C_1$	$((i,j),(m,n)) \in C_3$
$((k,l),(i,j)) \in C_4$	3.1. $((k,l),(m,n)),((k,j),(m,n)) \in C_4$	3.2. $((k,l),(m,n)) \in C_4, C_5$ *or* C_6 $((k,j),(m,n)) \in C_3$

3.1. $((i,j),(m,n)) \in C_1,\ ((k,l),(i,j)) \in C_4$ and $(k,l),(m,n)),((k,j),(m,n)) \in C_4$.

Then, we have $((k,n),(i,j)) \in C_2,\ ((i,l),(m,n)) \in C_1,\ ((i,l),(m,j)) \in C_1,\ ((m,l),(i,j))$

$\in C_1$ and

$$\begin{aligned}(f,g)_w &\equiv -q^{-2}x_{ij}[x_{kl}x_{mn}-(q^2-q^{-2})x_{kn}x_{ml}]+[x_{kl}x_{mn}-(q^2-q^{-2})x_{kn}x_{ml}]x_{ij}\\ &\quad -(q^2-q^{-2})[x_{kj}x_{mn}-(q^2-q^{-2})x_{kn}x_{mj}]x_{il}\\ &\equiv -q^{-2}[x_{kl}x_{ij}-(q^2-q^{-2})x_{kj}x_{il}]x_{mn}+q^{-2}(q^2-q^{-2})x_{kn}x_{ij}x_{ml}+q^{-2}x_{kl}x_{ij}x_{mn}\\ &\quad -(q^2-q^{-2})x_{kn}x_{ml}x_{ij}-q^{-2}(q^2-q^{-2})x_{kj}x_{il}x_{mn}+q^{-2}(q^2-q^{-2})^2x_{kn}x_{il}x_{mj}\\ &\equiv q^{-4}(q^2-q^{-2})x_{kn}x_{ml}x_{ij}-(q^2-q^{-2})x_{kn}x_{ml}x_{ij}+q^{-2}(q^2-q^{-2})x_{kn}x_{ml}x_{ij}\\ &\equiv 0.\end{aligned}$$

3.2. $((i,j),(m,n))\in C_3,\ ((k,l),(i,j))\in C_4,\ (k,l),(m,n))\in C_4,C_5$ *or* C_6 and $((k,j),(m,n))\in C_3$.

3.2.1. If $((k,l),(m,n))\in C_4$ $(l>m)$ and $((k,j),(m,n))\in C_3$, then $((k,n),(i,j))\in C_3,\ ((i,j),(m,l))\in C_2,\ ((i,l),(m,n))\in C_4$ and

$$\begin{aligned}(f,g)_w &\equiv -q^{-2}x_{ij}[x_{kl}x_{mn}-(q^2-q^{-2})x_{kn}x_{ml}]+[x_{kl}x_{mn}-(q^2-q^{-2})x_{kn}x_{ml}]x_{ij}\\ &\quad -q^{-2}(q^2-q^{-2})x_{kj}x_{mn}x_{il}\\ &\equiv -q^{-2}[x_{kl}x_{ij}-(q^2-q^{-2})x_{kj}x_{il}]x_{mn}+q^{-4}(q^2-q^{-2})x_{kn}x_{ij}x_{ml}+q^{-2}x_{kl}x_{ij}x_{mn}\\ &\quad -(q^2-q^{-2})x_{kn}x_{ij}x_{ml}-q^{-2}(q^2-q^{-2})x_{kj}[x_{il}x_{mn}-(q^2-q^{-2})x_{in}x_{ml}]\\ &\equiv 0.\end{aligned}$$

3.2.2. If $((k,l),(m,n))\in C_5$ $(l=m)$ and $((k,j),(m,n))\in C_3$, then $((k,l),(i,j))\in C_4,\ ((k,n),(i,j))\in C_3,\ ((i,l),(m,n))\in C_5$ and

$$\begin{aligned}(f,g)_w &\equiv -q^{-2}x_{ij}(q^2x_{kl}x_{mn}-qx_{kn})+(q^2x_{kl}x_{mn}-qx_{kn})x_{ij}-q^{-2}(q^2-q^{-2})x_{kj}x_{mn}x_{il}\\ &\equiv -[x_{kl}x_{ij}-(q^2-q^{-2})x_{kj}x_{il}]x_{mn}+q^{-3}x_{kn}x_{ij}+x_{kl}x_{ij}x_{mn}-qx_{kn}x_{ij}\\ &\quad -q^{-2}(q^2-q^{-2})x_{kj}[q^2x_{il}x_{mn}-qx_{in}]\\ &\equiv q^{-3}x_{kn}x_{ij}-qx_{kn}x_{ij}+q^{-1}(q^2-q^{-2})x_{kn}x_{ij}\\ &\equiv 0.\end{aligned}$$

3.2.3. If $((k,l),(m,n))\in C_6$ $(l<m)$ and $((k,j),(m,n))\in C_3$, then $((i,l),(m,n))\in C_6$ and

$$\begin{aligned}(f,g)_w &\equiv -q^{-2}x_{ij}x_{kl}x_{mn}+x_{kl}x_{mn}x_{ij}-q^{-2}(q^2-q^{-2})x_{kj}x_{mn}x_{il}\\ &\equiv -q^{-2}[x_{kl}x_{ij}-(q^2-q^{-2})x_{kj}x_{il}]x_{mn}+q^{-2}x_{kl}x_{ij}x_{mn}-q^{-2}(q^2-q^{-2})x_{kj}x_{il}x_{mn}\\ &\equiv 0.\end{aligned}$$

Case 4. $f=x_{mn}x_{ij}-q^{-2}x_{ij}x_{mn},\ g=x_{ij}x_{kl}-q^2x_{kl}x_{ij}+qx_{kj},\ w=x_{mn}x_{ij}x_{kl}$.

In the case, we have

$$(f,g)_w=-q^{-2}x_{ij}x_{mn}x_{kl}+q^2x_{mn}x_{kl}x_{ij}-qx_{mn}x_{kj}.$$

There are two subcases to consider.

	$((i,j),(m,n)) \in C_1$	$((i,j),(m,n)) \in C_3$
$((k,l),(i,j)) \in C_5$	4.1. $((k,l),(m,n)) \in C_5$ $((k,j),(m,n)) \in C_4$	4.2. $((k,l),(m,n)) \in C_6$ $((k,j),(m,n)) \in C_3$

4.1. $((i,j),(m,n)) \in C_1,\ ((k,l),(i,j)) \in C_5, ((k,l),(m,n)) \in C_5$ and $((k,j),(m,n)) \in C_4$.

Then, we have $((k,n),(i,j)) \in C_2$ $(m = i)$ and

$$\begin{aligned}(f,g)_w &\equiv -q^{-2}x_{ij}(q^2x_{kl}x_{mn} - qx_{kn}) + q^2(q^2x_{kl}x_{mn} - qx_{kn})x_{ij} \\ &\quad -q[x_{kj}x_{mn} - (q^2 - q^{-2})x_{kn}x_{mj}] \\ &\equiv -(q^2x_{kl}x_{ij} - qx_{kj})x_{mn} + q^{-1}x_{kn}x_{ij} + q^2x_{kl}x_{ij}x_{mn} \\ &\quad -q^3x_{kn}x_{ij} - qx_{kj}x_{mn} + q(q^2 - q^{-2})x_{kn}x_{mj} \\ &\equiv 0.\end{aligned}$$

4.2. $((i,j),(m,n)) \in C_3,\ ((k,l),(i,j)) \in C_5, ((k,l),(m,n)) \in C_6$ and $((k,j),(m,n)) \in C_3$.

Then, we have

$$\begin{aligned}(f,g)_w &\equiv -q^{-2}x_{ij}x_{kl}x_{mn} + q^2x_{kl}x_{mn}x_{ij} - q^{-1}x_{kj}x_{mn} \\ &\equiv -q^{-2}(q^2x_{kl}x_{ij} - qx_{kj})x_{mn} + x_{kl}x_{ij}x_{mn} - q^{-1}x_{kj}x_{mn} \\ &\equiv 0.\end{aligned}$$

Case 5. $f = x_{mn}x_{ij} - x_{ij}x_{mn},\ g = x_{ij}x_{kl} - q^{-2}x_{kl}x_{ij},\ w = x_{mn}x_{ij}x_{kl}$.

In the case, we have

$$(f,g)_w = -x_{ij}x_{mn}x_{kl} + q^{-2}x_{mn}x_{kl}x_{ij}.$$

There are four subcases to consider.

	$((i,j),(m,n)) \in C_2$	$((i,j),(m,n)) \in C_6$
$((k,l),(i,j)) \in C_1$	5.1. $((k,l),(m,n)) \in C_2, C_3, C_4, C_5$ *or* C_6	5.3. $((k,l),(m,n)) \in C_6$
$((k,l),(i,j)) \in C_3$	5.2. $((k,l),(m,n)) \in C_2$	5.4. $((k,l),(m,n)) \in C_6$

5.1. $((i,j),(m,n)) \in C_2,\ ((k,l),(i,j)) \in C_1$, and $((k,l),(m,n)) \in C_2, C_3, C_4, C_5$ or C_6.

5.1.1. If $((k,l),(m,n)) \in C_2$ $(l > n)$, then we have $((k,l),(i,j)) \in C_1$ and

$$\begin{aligned}(f,g)_w &\equiv -x_{ij}x_{kl}x_{mn} + q^{-2}x_{kl}x_{mn}x_{ij} \\ &\equiv -q^{-2}x_{kl}x_{ij}x_{mn} + q^{-2}x_{kl}x_{ij}x_{mn} \\ &\equiv 0.\end{aligned}$$

5.1.2. If $((k,l),(m,n)) \in C_3$ $(l = n)$, then

$$\begin{aligned}(f,g)_w &\equiv -q^{-2}x_{ij}x_{kl}x_{mn} + q^{-4}x_{kl}x_{mn}x_{ij} \\ &\equiv -q^{-4}x_{kl}x_{ij}x_{mn} + q^{-4}x_{kl}x_{ij}x_{mn} \\ &\equiv 0.\end{aligned}$$

5.1.3. If $((k,l),(m,n)) \in C_4$ $(m < l < n)$, then we have $((k,l),(i,j)) \in C_1$, $((k,n),(i,j)) \in C_1$, $((i,j),(m,l)) \in C_2$ and

$$\begin{aligned}(f,g)_w &\equiv -x_{ij}[x_{kl}x_{mn} - (q^2-q^{-2})x_{kn}x_{ml}] + q^{-2}[x_{kl}x_{mn} - (q^2-q^{-2})x_{kn}x_{ml}]x_{ij} \\ &\equiv -x_{ij}x_{kl}x_{mn} + (q^2-q^{-2})x_{ij}x_{kn}x_{ml} + q^{-2}x_{kl}x_{mn}x_{ij} - q^{-2}(q^2-q^{-2})x_{kn}x_{ml}x_{ij} \\ &\equiv -q^{-2}x_{kl}x_{ij}x_{mn} + q^{-2}(q^2-q^{-2})x_{kn}x_{ij}x_{ml} + q^{-2}x_{kl}x_{ij}x_{mn} \\ &\quad -q^{-2}(q^2-q^{-2})x_{kn}x_{ij}x_{ml} \\ &\equiv 0.\end{aligned}$$

5.1.4. If $((k,l),(m,n)) \in C_5$ $(m = l)$, then we have $((k,n),(i,j)) \in C_1$ and

$$\begin{aligned}(f,g)_w &\equiv -x_{ij}(q^2x_{kl}x_{mn} - qx_{kn}) + q^{-2}(q^2x_{kl}x_{mn} - qx_{kn})x_{ij} \\ &\equiv -q^2x_{ij}x_{kl}x_{mn} + qx_{ij}x_{kn} + x_{kl}x_{mn}x_{ij} - q^{-1}x_{kn}x_{ij} \\ &\equiv -x_{kl}x_{ij}x_{mn} + q^{-1}x_{kn}x_{ij} + x_{kl}x_{ij}x_{mn} - q^{-1}x_{kn}x_{ij} \\ &\equiv 0.\end{aligned}$$

5.1.5. If $((k,l),(m,n)) \in C_6$ $(l < m)$, the proof is similar to 5.1.1.

For the cases of 5.2, 5.3 and 5.4, the proofs are also similar to 5.1.1.

Case 6. $f = x_{mn}x_{ij} - x_{ij}x_{mn}$, $g = x_{ij}x_{kl} - x_{kl}x_{ij}$, $w = x_{mn}x_{ij}x_{kl}$.

In the case, we have

$$(f,g)_w = -x_{ij}x_{mn}x_{kl} + x_{mn}x_{kl}x_{ij}.$$

There are four subcases to consider.

	$((i,j),(m,n)) \in C_2$	$((i,j),(m,n)) \in C_6$
$((k,l),(i,j)) \in C_2$	6.1. $((k,l),(m,n)) \in C_2$	6.3. $((k,l),(m,n)) \in C_2, C_3, C_4, C_5$ *or* C_6
$((k,l),(i,j)) \in C_6$	6.2. $((k,l),(m,n)) \in C_6$	6.4. $((k,l),(m,n)) \in C_6$

6.1. $((i,j),(m,n)) \in C_2$, $((k,l),(i,j)) \in C_2$ and $((k,l),(m,n)) \in C_2$.

Then, we have

$$\begin{aligned}(f,g)_w &\equiv -x_{ij}x_{kl}x_{mn} + x_{kl}x_{mn}x_{ij} \\ &\equiv -x_{kl}x_{ij}x_{mn} + x_{kl}x_{ij}x_{mn} \\ &\equiv 0.\end{aligned}$$

6.2. $((i,j),(m,n)) \in C_2,\ ((k,l),(i,j)) \in C_6$ and $((k,l),(m,n)) \in C_6$.

This case is similar to 6.1.

6.3. $((i,j),(m,n)) \in C_6,\ ((k,l),(i,j)) \in C_2$ and $((k,l),(m,n)) \in C_2,,C_3,C_4,C_5$ or C_6.

6.3.1. If $((k,l),(m,n)) \in C_2\ (l > n)$, the proof is similar to 6.1.

6.3.2. If $((k,l),(m,n)) \in C_3\ (l = n)$, then

$$\begin{aligned}(f,g)_w &\equiv -q^{-2}x_{ij}x_{kl}x_{mn} + q^{-2}x_{kl}x_{mn}x_{ij}\\ &\equiv -q^{-2}x_{kl}x_{ij}x_{mn} + q^{-2}x_{kl}x_{ij}x_{mn}\\ &\equiv 0.\end{aligned}$$

6.3.3. If $((k,l),(m,n)) \in C_4\ (m < l < n)$, then we have $((k,n),(i,j)) \in C_2,\ ((i,j),(m,n)) \in C_6$ and

$$\begin{aligned}(f,g)_w &\equiv -x_{ij}[x_{kl}x_{mn} - (q^2-q^{-2})x_{kn}x_{ml}] + [x_{kl}x_{mn} - (q^2-q^{-2})x_{kn}x_{ml}]x_{ij}\\ &\equiv -x_{ij}x_{kl}x_{mn} + (q^2-q^{-2})x_{ij}x_{kn}x_{ml} + x_{kl}x_{mn}x_{ij} - (q^2-q^{-2})x_{kn}x_{ml}x_{ij}\\ &\equiv -x_{kl}x_{ij}x_{mn} + (q^2-q^{-2})x_{kn}x_{ij}x_{ml} + x_{kl}x_{ij}x_{mn} - (q^2-q^{-2})x_{kn}x_{ij}x_{ml}\\ &\equiv 0.\end{aligned}$$

6.3.4. If $((k,l),(m,n)) \in C_5\ (m = l)$, then we have $((k,n),(i,j)) \in C_2$ and

$$\begin{aligned}(f,g)_w &\equiv -x_{ij}(q^2x_{kl}x_{mn} - qx_{kn}) + (q^2x_{kl}x_{mn} - qx_{kn})x_{ij}\\ &\equiv -q^2x_{ij}x_{kl}x_{mn} + qx_{ij}x_{kn} + q^2x_{kl}x_{mn}x_{ij} - qx_{kn}x_{ij}\\ &\equiv -q^2x_{kl}x_{ij}x_{mn} + qx_{kn}x_{ij} + q^2x_{kl}x_{ij}x_{mn} - qx_{kn}x_{ij}\\ &\equiv 0.\end{aligned}$$

6.3.5. If $((k,l),(m,n)) \in C_6\ (l < m)$, the proof is similar to 6.1.

6.4. $((i,j),(m,n)) \in C_6,\ ((k,l),(i,j)) \in C_6$ and $((k,l),(m,n)) \in C_6$.

This case is also similar to 6.1.

Case 7. $f = x_{mn}x_{ij} - x_{ij}x_{mn},\ g = x_{ij}x_{kl} - x_{kl}x_{ij} + (q^2-q^{-2})x_{kj}x_{il},\ w = x_{mn}x_{ij}x_{kl}$.

In the case, we have

$$(f,g)_w = -x_{ij}x_{mn}x_{kl} + x_{mn}x_{kl}x_{ij} - (q^2-q^{-2})x_{mn}x_{kj}x_{il}.$$

There are two subcases to consider.

	$((i,j),(m,n)) \in C_2$	$((i,j),(m,n)) \in C_6$
$((k,l),(i,j)) \in C_4$	7.1. $((k,l),(m,n)) \in C_2,C_3,C_4,C_5$ *or* C_6 $((k,j),(m,n)) \in C_2$	7.2. $((k,l),(m,n))$, $((k,j),(m,n)) \in C_6$

7.1. $((i,j),(m,n)) \in C_2,\ ((k,l),(i,j)) \in C_4,\ ((k,l),(m,n)) \in C_2, C_3, C_4, C_5$ or C_6 and $((k,j),(m,n)) \in C_2$.

7.1.1. If $((k,l),(m,n)) \in C_2$ $(n < l)$ and $((k,j),(m,n)) \in C_2$, then we have $((i,l),(m,n)) \in C_2$ and

$$\begin{aligned}(f,g)_w &\equiv -x_{ij}x_{kl}x_{mn} + x_{kl}x_{mn}x_{ij} - (q^2-q^{-2})x_{kj}x_{mn}x_{il}\\ &\equiv -[x_{kl}x_{ij} - (q^2-q^{-2})x_{kj}x_{il}]x_{mn} + x_{kl}x_{ij}x_{mn} - (q^2-q^{-2})x_{kj}x_{il}x_{mn}\\ &\equiv 0.\end{aligned}$$

7.1.2. If $((k,l),(m,n)) \in C_3$ $(n = l)$ and $((k,j),(m,n)) \in C_2$, then $((i,l),(m,n)) \in C_3$ and

$$\begin{aligned}(f,g)_w &\equiv -q^{-2}x_{ij}x_{kl}x_{mn} + q^{-2}x_{kl}x_{mn}x_{ij} - (q^2-q^{-2})x_{kj}x_{mn}x_{il}\\ &\equiv -q^{-2}[x_{kl}x_{ij} - (q^2-q^{-2})x_{kj}x_{il}]x_{mn} + q^{-2}x_{kl}x_{ij}x_{mn} - q^{-2}(q^2-q^{-2})x_{kj}x_{il}x_{mn}\\ &\equiv 0.\end{aligned}$$

7.1.3. If $((k,l),(m,n)) \in C_4$ $(m < l < n)$ and $((k,j),(m,n)) \in C_2$, then we obtain $((k,n),(i,j)) \in C_4,\ ((i,j),(m,l)) \in C_2,\ ((i,l),(m,n)) \in C_4$ and

$$\begin{aligned}(f,g)_w &\equiv -x_{ij}[x_{kl}x_{mn} - (q^2-q^{-2})x_{kn}x_{ml}] + [x_{kl}x_{mn} - (q^2-q^{-2})x_{kn}x_{ml}]x_{ij}\\ &\quad -(q^2-q^{-2})x_{kj}x_{mn}x_{il}\\ &\equiv -x_{ij}x_{kl}x_{mn} + (q^2-q^{-2})x_{ij}x_{kn}x_{ml} + x_{kl}x_{mn}x_{ij} - (q^2-q^{-2})x_{kn}x_{ml}x_{ij}\\ &\quad -(q^2-q^{-2})x_{kj}[x_{il}x_{mn} - (q^2-q^{-2})x_{in}x_{ml}]\\ &\equiv -[x_{kl}x_{ij} - (q^2-q^{-2})x_{kj}x_{il}]x_{mn} + (q^2-q^{-2})[x_{kn}x_{ij} - (q^2-q^{-2})x_{kj}x_{in}]x_{ml}\\ &\quad +x_{kl}x_{ij}x_{mn} - (q^2-q^{-2})x_{kn}x_{ij}x_{ml}\\ &\equiv 0.\end{aligned}$$

7.1.4. If $((k,l),(m,n)) \in C_5$ $(m = l)$ and $((k,j),(m,n)) \in C_2$, then $((k,n),(i,j)) \in C_4,\ ((i,l),(m,n)) \in C_5$ and

$$\begin{aligned}(f,g)_w &\equiv -x_{ij}(q^2x_{kl}x_{mn} - qx_{kn}) + (q^2x_{kl}x_{mn} - qx_{kn})x_{ij} - (q^2-q^{-2})x_{kj}x_{mn}x_{il}\\ &\equiv -q^2x_{ij}x_{kl}x_{mn} + qx_{ij}x_{kn} + q^2x_{kl}x_{mn}x_{ij} - qx_{kn}x_{ij}\\ &\quad -(q^2-q^{-2})x_{kj}(q^2x_{il}x_{mn} - qx_{in})\\ &\equiv -q^2[x_{kl}x_{ij} - (q^2-q^{-2})x_{kj}x_{il}]x_{mn} + q[x_{kn}x_{ij} - (q^2-q^{-2})x_{kj}x_{in}]\\ &\quad +q^2x_{kl}x_{ij}x_{mn} - qx_{kn}x_{ij} - q^2(q^2-q^{-2})x_{kj}x_{il}x_{mn} + q(q^2-q^{-2})x_{kj}x_{in}\\ &\equiv 0.\end{aligned}$$

7.1.5. If $((k,l),(m,n)) \in C_6$ $(l < m)$ and $((k,j),(m,n)) \in C_2$, then $((i,l),(m,n)) \in C_6$. This case is similar to 7.1.1.

7.2. $((i,j),(m,n)) \in C_6,\ ((k,l),(i,j)) \in C_4,\ ((k,l),(m,n)),((k,j),(m,n)) \in C_6$.

This case is also similar to 7.1.1.

Case 8. $f = x_{mn}x_{ij} - x_{ij}x_{mn},\ g = x_{ij}x_{kl} - q^2 x_{kl}x_{ij} + qx_{kj},\ w = x_{mn}x_{ij}x_{kl}$.

In the case, we have

$$(f,g)_w = -x_{ij}x_{mn}x_{kl} + q^2 x_{mn}x_{kl}x_{ij} + qx_{mn}x_{kj}.$$

There are two subcases to consider.

	$((i,j),(m,n)) \in C_2$	$((i,j),(m,n)) \in C_6$
$((k,l),(i,j)) \in C_5$	8.1. $((k,l),(m,n)) \in C_6$ $((k,j),(m,n)) \in C_2$	8.2. $((k,l),(m,n)),((k,j),(m,n)) \in C_6$

8.1. $((i,j),(m,n)) \in C_2,\ ((k,l),(i,j)) \in C_5,\ ((k,l),(m,n)) \in C_6$ and $((k,j),(m,n)) \in C_2$. Then, we have

$$\begin{aligned}(f,g)_w &\equiv -x_{ij}x_{kl}x_{mn} + q^2 x_{kl}x_{mn}x_{ij} + qx_{kj}x_{mn}\\ &\equiv -(q^2 x_{kl}x_{ij} - qx_{kj})x_{mn} + q^2 x_{kl}x_{ij}x_{mn} + qx_{kj}x_{mn}\\ &\equiv 0.\end{aligned}$$

8.2. $((i,j),(m,n)) \in C_6,\ ((k,l),(i,j)) \in C_5,\ ((k,l),(m,n)),((k,j),(m,n)) \in C_6$.

This case is similar to 8.1.

Case 9. $f = x_{mn}x_{ij} - x_{ij}x_{mn} + (q^2 - q^{-2})x_{in}x_{mj},\ g = x_{ij}x_{kl} - q^{-2}x_{kl}x_{ij},\ w = x_{mn}x_{ij}x_{kl}$.

In the case, we have

$$(f,g)_w = -x_{ij}x_{mn}x_{kl} + (q^2 - q^{-2})x_{in}x_{mj}x_{kl} + q^{-2}x_{mn}x_{kl}x_{ij}.$$

There are two subcases to consider.

	$((i,j),(m,n)) \in C_4$
$((k,l),(i,j)) \in C_1$	9.1. $((k,l),(m,n)),((k,l),(m,j)) \in C_4,\ C_5$ *or* C_6
$((k,l),(i,j)) \in C_3$	9.2. $((k,l),(m,n)) \in C_4\ \ ((k,l),(m,j)) \in C_3$

9.1. $((i,j),(m,n)) \in C_4,\ ((k,l),(i,j)) \in C_1$ and $((k,l),(m,n)),((k,l),(m,j)) \in C_4,\ C_5$ or C_6.

9.1.1. If $((k,l),(m,n)),((k,l),(m,j)) \in C_4$ $(l > m)$, then we have $((i,j),(k,n)) \in C_1$, $((k,n),(m,l)) \in C_2$, $((k,j),(i,n)) \in C_1$, $((k,l),(i,n)) \in C_1$, $((i,j),(m,l)) \in C_2$ and

$$\begin{aligned}
(f,g)_w &\equiv -x_{ij}[x_{kl}x_{mn} - (q^2-q^{-2})x_{kn}x_{ml}] + (q^2-q^{-2})x_{in}[x_{kl}x_{mj} - (q^2-q^{-2})x_{kj}x_{ml}] \\
&\quad +q^{-2}[x_{kl}x_{mn} - (q^2-q^{-2})x_{kn}x_{ml}]x_{ij} \\
&\equiv -x_{ij}x_{kl}x_{mn} + (q^2-q^{-2})x_{ij}x_{kn}x_{ml} + (q^2-q^{-2})x_{in}x_{kl}x_{mj} \\
&\quad -(q^2-q^{-2})^2x_{in}x_{kj}x_{ml} + q^{-2}x_{kl}x_{mn}x_{ij} - q^{-2}(q^2-q^{-2})x_{kn}x_{ml}x_{ij} \\
&\equiv -q^{-2}x_{kl}x_{ij}x_{mn} + (q^2-q^{-2})x_{ij}x_{kn}x_{ml} + q^{-2}(q^2-q^{-2})x_{kl}x_{in}x_{mj} \\
&\quad -q^{-2}(q^2-q^{-2})^2x_{kj}x_{in}x_{ml} + q^{-2}x_{kl}[x_{ij}x_{mn} - (q^2-q^{-2})x_{in}x_{mj}] \\
&\quad -q^{-2}(q^2-q^{-2})x_{kn}x_{ij}x_{ml} \\
&\equiv (q^2-q^{-2})x_{ij}x_{kn}x_{ml} - q^{-2}(q^2-q^{-2})^2x_{kj}x_{in}x_{ml} - q^{-4}(q^2-q^{-2})x_{ij}x_{kn}x_{ml} \\
&\equiv 0.
\end{aligned}$$

9.1.2. If $((k,l),(m,n)),((k,l),(m,j)) \in C_5$ $(l = m)$, then we have $((i,j),(k,n)) \in C_1$, $((k,l),(i,n),((k,j),(i,n)) \in C_1$ and

$$\begin{aligned}
(f,g)_w &\equiv -x_{ij}(q^2x_{kl}x_{mn} - qx_{kn}) + (q^2-q^{-2})x_{in}(q^2x_{kl}x_{mj} - qx_{kj}) \\
&\quad +q^{-2}(q^2x_{kl}x_{mn} - qx_{kn})x_{ij} \\
&\equiv -q^2x_{ij}x_{kl}x_{mn} + qx_{ij}x_{kn} + q^2(q^2-q^{-2})x_{in}x_{kl}x_{mj} - q(q^2-q^{-2})x_{in}x_{kj} \\
&\quad +x_{kl}x_{mn}x_{ij} - q^{-1}x_{kn}x_{ij} \\
&\equiv -x_{kl}x_{ij}x_{mn} + qx_{ij}x_{kn} + (q^2-q^{-2})x_{kl}x_{in}x_{mj} - q^{-1}(q^2-q^{-2})x_{kj}x_{in} \\
&\quad +x_{kl}[x_{ij}x_{mn} - (q^2-q^{-2})x_{in}x_{mj}] - q^{-3}x_{ij}x_{kn} \\
&\equiv qx_{ij}x_{kn} - qx_{kj}x_{in} + q^{-3}x_{kj}x_{in} - q^{-3}x_{ij}x_{kn} \\
&\equiv 0.
\end{aligned}$$

9.1.3. If $((k,l),(m,n)),((k,l),(m,j)) \in C_6$ $(l < m)$, then we have $((k,l),(i,n)) \in C_1$ and

$$\begin{aligned}
(f,g)_w &\equiv -x_{ij}x_{kl}x_{mn} - (q^2-q^{-2})x_{in}x_{kl}x_{mj} + q^{-2}x_{kl}x_{mn}x_{ij} \\
&\equiv -q^{-2}x_{kl}x_{ij}x_{mn} - q^{-2}(q^2-q^{-2})x_{kl}x_{in}x_{mj} + q^{-2}x_{kl}[x_{ij}x_{mn} - (q^2-q^{-2})x_{in}x_{mj}] \\
&\equiv 0.
\end{aligned}$$

9.2. $((i,j),(m,n)) \in C_4$, $((k,l),(i,j)) \in C_3$, $((k,l),(m,n)) \in C_4$ and $((k,l),(m,j)) \in C_3$.

Then, we have $((k,n),(i,j)) \in C_2,\ ((k,l),(i,n)) \in C_4,\ ((i,j),(m,l)) \in C_3$ and

$$\begin{aligned}(f,g)_w &\equiv -x_{ij}[x_{kl}x_{mn}-(q^2-q^{-2})x_{kn}x_{ml}]+q^{-2}(q^2-q^{-2})x_{in}x_{kl}x_{mj}\\ &\quad +q^{-2}[x_{kl}x_{mn}-(q^2-q^{-2})x_{kn}x_{ml}]x_{ij}\\ &\equiv -x_{ij}x_{kl}x_{mn}+(q^2-q^{-2})x_{ij}x_{kn}x_{ml}+q^{-2}(q^2-q^{-2})x_{in}x_{kl}x_{mj}+q^{-2}x_{kl}x_{mn}x_{ij}\\ &\quad -q^{-2}(q^2-q^{-2})x_{kn}x_{ml}x_{ij}\\ &\equiv -q^{-2}x_{kl}x_{ij}x_{mn}+(q^2-q^{-2})x_{kn}x_{ij}x_{ml}+q^{-2}(q^2-q^{-2})[x_{kl}x_{in}\\ &\quad -(q^2-q^{-2})x_{kn}x_{il}]x_{mj}+q^{-2}x_{kl}[x_{ij}x_{mn}-(q^2-q^{-2})x_{in}x_{mj}]\\ &\quad -q^{-4}(q^2-q^{-2})x_{kn}x_{ij}x_{ml}\\ &\equiv (q^2-q^{-2})x_{kn}x_{ij}x_{ml}-q^{-2}(q^2-q^{-2})x_{kn}x_{il}x_{mj}-q^{-4}(q^2-q^{-2})x_{kn}x_{ij}x_{ml}\\ &\equiv 0.\end{aligned}$$

Case 10. $f=x_{mn}x_{ij}-x_{ij}x_{mn}+(q^2-q^{-2})x_{in}x_{mj},\ g=x_{ij}x_{kl}-x_{kl}x_{ij},\ w=x_{mn}x_{ij}x_{kl}$.
In the case, we have

$$(f,g)_w=-x_{ij}x_{mn}x_{kl}+(q^2-q^{-2})x_{in}x_{mj}x_{kl}+x_{mn}x_{kl}x_{ij}.$$

There are two subcases to consider.

	$((i,j),(m,n)) \in C_4$
$((k,l),(i,j)) \in C_2$	10.1. $((k,l),(m,n)) \in C_2, C_3$ *or* C_4 $((k,l),(m,j)) \in C_2$
$((k,l),(i,j)) \in C_6$	10.2. $((k,l),(m,n)),((k,l),(m,j)) \in C_6$

10.1. $((i,j),(m,n)) \in C_4,\ ((k,l),(i,j)) \in C_2,\ ((k,l),(m,n)) \in C_2, C_3$ or C_4 and $((k,l),(m,j)) \in C_2$.

10.1.1. If $((k,l),(m,n)) \in C_2$ $(l>n)$, then we have $((k,l),(i,n)) \in C_2$ and

$$\begin{aligned}(f,g)_w &\equiv -x_{ij}x_{kl}x_{mn}+(q^2-q^{-2})x_{in}x_{kl}x_{mj}+x_{kl}x_{mn}x_{ij}\\ &\equiv -x_{kl}x_{ij}x_{mn}+(q^2-q^{-2})x_{kl}x_{in}x_{mj}+x_{kl}[x_{ij}x_{mn}-(q^2-q^{-2})x_{in}x_{mj}]\\ &\equiv 0.\end{aligned}$$

10.1.2. If $((k,l),(m,n)) \in C_3$ $(l=n)$, then we have $((k,l),(i,n)) \in C_3$ and

$$\begin{aligned}(f,g)_w &\equiv -q^{-2}x_{ij}x_{kl}x_{mn}+(q^2-q^{-2})x_{in}x_{kl}x_{mj}+q^{-2}x_{kl}x_{mn}x_{ij}\\ &\equiv -q^{-2}x_{kl}x_{ij}x_{mn}+q^{-2}(q^2-q^{-2})x_{kl}x_{in}x_{mj}+q^{-2}x_{kl}[x_{ij}x_{mn}-(q^2-q^{-2})x_{in}x_{mj}]\\ &\equiv 0.\end{aligned}$$

10.1.3. If $((k,l),(m,n)) \in C_4$ $(l < n)$, then we have $((k,n),(i,j)) \in C_2$, $((k,l),(i,n)) \in C_4$, $((i,j),(m,l)) \in C_4$ and

$$\begin{aligned}
(f,g)_w &\equiv -x_{ij}[x_{kl}x_{mn} - (q^2-q^{-2})x_{kn}x_{ml}] + (q^2-q^{-2})x_{in}x_{kl}x_{mj} \\
&\quad +[x_{kl}x_{mn} - (q^2-q^{-2})x_{kn}x_{ml}]x_{ij} \\
&\equiv -x_{ij}x_{kl}x_{mn} + (q^2-q^{-2})x_{ij}x_{kn}x_{ml} + (q^2-q^{-2})x_{in}x_{kl}x_{mj} \\
&\quad +x_{kl}x_{mn}x_{ij} - (q^2-q^{-2})x_{kn}x_{ml}x_{ij} \\
&\equiv -x_{kl}x_{ij}x_{mn} + (q^2-q^{-2})x_{kn}x_{ij}x_{ml} + (q^2-q^{-2})[x_{kl}x_{in} - (q^2-q^{-2})x_{kn}x_{il}]x_{mj} \\
&\quad +x_{kl}[x_{ij}x_{mn} - (q^2-q^{-2})x_{in}x_{mj}] - (q^2-q^{-2})x_{kn}[x_{ij}x_{ml} - (q^2-q^{-2})x_{il}x_{mj}] \\
&\equiv 0.
\end{aligned}$$

10.2. $((i,j),(m,n)) \in C_4$, $((k,l),(i,j)) \in C_6$, $((k,l),(m,n)),(k,l),(m,j)) \in C_6$.

This case is similar to 10.1.

Case 11. $f = x_{mn}x_{ij} - x_{ij}x_{mn} + (q^2-q^{-2})x_{in}x_{mj}$, $g = x_{ij}x_{kl} - x_{kl}x_{ij} + (q^2-q^{-2})x_{kj}x_{il}$, $w = x_{mn}x_{ij}x_{kl}$.

In the case, we have

$$(f,g)_w = -x_{ij}x_{mn}x_{kl} + (q^2-q^{-2})x_{in}x_{mj}x_{kl} + x_{mn}x_{kl}x_{ij} - (q^2-q^{-2})x_{mn}x_{kj}x_{il},$$

with

	$((i,j),(m,n)) \in C_4$
$((k,l),(i,j)) \in C_4$	$((k,l),(m,n)),((k,l),(m,j)) \in C_4, C_5$ *or* C_6

11.1. If $((k,l),(m,n))$, $((k,l),(m,j)) \in C_4$ $(l > m)$, then we have $((k,n),(i,j)) \in C_2$, $((k,l),(i,n)) \in C_4$, $((k,j),(i,n)) \in C_4$, $((i,j),(m,l)) \in C_2$, $((i,l),(m,n)) \in C_4$,
$((i,l),(m,j)) \in C_4$ and

$$\begin{aligned}
(f,g)_w &\equiv -x_{ij}[x_{kl}x_{mn} - (q^2-q^{-2})x_{kn}x_{ml}] + (q^2-q^{-2})x_{in}[x_{kl}x_{mj} - (q^2-q^{-2})x_{kj}x_{ml}] \\
&\quad +[x_{kl}x_{mn} - (q^2-q^{-2})x_{kn}x_{ml}]x_{ij} - (q^2-q^{-2})[x_{kj}x_{mn} - (q^2-q^{-2})x_{kn}x_{mj}]x_{il} \\
&\equiv -x_{ij}x_{kl}x_{mn} + (q^2-q^{-2})x_{ij}x_{kn}x_{ml} + (q^2-q^{-2})x_{in}x_{kl}x_{mj} \\
&\quad -(q^2-q^{-2})x_{in}x_{kj}x_{ml} + x_{kl}x_{mn}x_{ij} - (q^2-q^{-2})x_{kn}x_{ml}x_{ij} \\
&\quad -(q^2-q^{-2})x_{kj}x_{mn}x_{il} + (q^2-q^{-2})^2x_{kn}x_{mj}x_{il} \\
&\equiv -[x_{kl}x_{ij} - (q^2-q^{-2})x_{kj}x_{il}]x_{mn} + (q^2-q^{-2})x_{kn}x_{ij}x_{ml} + (q^2-q^{-2})[x_{kj}x_{in} \\
&\quad -(q^2-q^{-2})x_{kn}x_{il}]x_{mj} - (q^2-q^{-2})[x_{kj}x_{in} - (q^2-q^{-2})x_{kn}x_{ij}]x_{ml} \\
&\quad +x_{kl}[x_{ij}x_{mn} - (q^2-q^{-2})x_{in}x_{mj}] - (q^2-q^{-2})x_{kn}x_{ij}x_{ml} \\
&\quad -(q^2-q^{-2})x_{kj}[x_{il}x_{mn} - (q^2-q^{-2})x_{in}x_{ml}] \\
&\quad +(q^2-q^{-2})^2x_{kn}[x_{il}x_{mj} - (q^2-q^{-2})x_{ij}x_{ml}] \\
&\equiv 0.
\end{aligned}$$

11.2. If $((k,l),(m,n)),\ ((k,l),(m,j)) \in C_5\ (l=m)$, then we have $((k,n),(i,j)) \in C_2,\ ((k,l),(i,n)) \in C_4,\ ((k,j),(i,n)) \in C_4,\ ((i,l),(m,n)) \in C_5,\ ((i,l),(m,j)) \in C_5$ and

$$\begin{aligned}
(f,g)_w &\equiv -x_{ij}(q^2x_{kl}x_{mn} - qx_{kn}) + (q^2-q^{-2})x_{in}(q^2x_{kl}x_{mj} - qx_{kj}) + (q^2x_{kl}x_{mn} - qx_{kn})x_{ij} \\
&\quad -(q^2-q^{-2})[x_{kj}x_{mn} - (q^2-q^{-2})x_{kn}x_{mj}]x_{il} \\
&\equiv -q^2x_{ij}x_{kl}x_{mn} + qx_{ij}x_{kn} + q^2(q^2-q^{-2})x_{in}x_{kl}x_{mj} - q(q^2-q^{-2})x_{in}x_{kj} \\
&\quad +q^2x_{kl}x_{mn}x_{ij} - qx_{kn}x_{ij} - (q^2-q^{-2})x_{kj}x_{mn}x_{il} + (q^2-q^{-2})^2x_{kn}x_{mj}x_{il} \\
&\equiv -q^2[x_{kl}x_{ij} - (q^2-q^{-2})x_{kj}x_{il}]x_{mn} + qx_{kn}x_{ij} + q^2(q^2-q^{-2})[x_{kl}x_{in} \\
&\quad -(q^2-q^{-2})x_{kn}x_{il}]x_{mj} - q(q^2-q^{-2})[x_{kj}x_{in} - (q^2-q^{-2})x_{kn}x_{ij}] \\
&\quad +q^2x_{kl}[x_{ij}x_{mn} - (q^2-q^{-2})x_{in}x_{mj}] - qx_{kn}x_{ij} \\
&\quad -(q^2-q^{-2})x_{kj}[q^2x_{il}x_{mn} - qx_{in}] + (q^2-q^{-2})^2x_{kn}[q^2x_{il}x_{mj} - qx_{ij}] \\
&\equiv 0.
\end{aligned}$$

11.3. If $((k,l),(m,n)),\ ((k,l),(m,j)) \in C_6\ (l<m)$, then $((k,j),(m,n)) \in C_4,\ ((k,l),(i,n)) \in C_4,\ ((i,l),(m,n)),((i,l),(m,j)) \in C_6$ and

$$\begin{aligned}
(f,g)_w &\equiv -x_{ij}x_{kl}x_{mn} + (q^2-q^{-2})x_{in}x_{kl}x_{mj} + x_{kl}x_{mn}x_{ij} \\
&\quad -(q^2-q^{-2})[x_{kj}x_{mn} - (q^2-q^{-2})x_{kn}x_{mj}]x_{il} \\
&\equiv -[x_{kl}x_{ij} - (q^2-q^{-2})x_{kj}x_{il}]x_{mn} + (q^2-q^{-2})[x_{kl}x_{in} - (q^2-q^{-2})x_{kn}x_{il}]x_{mj} \\
&\quad +x_{kl}[x_{ij}x_{mn} - (q^2-q^{-2})x_{in}x_{mj}] - (q^2-q^{-2})x_{kj}x_{il}x_{mn} + (q^2-q^{-2})^2x_{kn}x_{il}x_{mj} \\
&\equiv 0.
\end{aligned}$$

Case 12. $f = x_{mn}x_{ij} - x_{ij}x_{mn} + (q^2-q^{-2})x_{in}x_{mj},\ g = x_{ij}x_{kl} - q^2x_{kl}x_{ij} + qx_{kj},\ w = x_{mn}x_{ij}x_{kl}$, with

	$((i,j),(m,n)) \in C_4$
$((k,l),(i,j)) \in C_5$	$((k,l),(m,n)),((k,l),(m,j)) \in C_6$ $((k,j),(m,n)) \in C_4\quad ((k,l),(i,n)) \in C_5$

In the case, we can deduce that

$$\begin{aligned}
(f,g)_w &= -x_{ij}x_{mn}x_{kl} + (q^2-q^{-2})x_{in}x_{mj}x_{kl} + q^2x_{mn}x_{kl}x_{ij} - qx_{mn}x_{kj} \\
&\equiv -x_{ij}x_{kl}x_{mn} + (q^2-q^{-2})x_{in}x_{kl}x_{mj} + q^2x_{kj}x_{mn}x_{ij} \\
&\quad -q[x_{kj}x_{mn} - (q^2-q^{-2})x_{kn}x_{mj}] \\
&\equiv -(q^2x_{kl}x_{ij} - qx_{kj})x_{mn} + (q^2-q^{-2})(q^2x_{kl}x_{in} - qx_{kn})x_{mj} \\
&\quad +q^2x_{kl}[x_{ij}x_{mn} - (q^2-q^{-2})x_{in}x_{mj}] - qx_{kj}x_{mn} + q(q^2-q^{-2})x_{kn}x_{mj} \\
&\equiv -q^2x_{kl}x_{ij}x_{mn} + qx_{kj}x_{mn} + q^2(q^2-q^{-2})x_{kl}x_{in}x_{mj} - q(q^2-q^{-2})x_{kn}x_{mj} \\
&\quad +q^2x_{kl}x_{ij}x_{mn} - q^2(q^2-q^{-2})x_{kl}x_{in}x_{mj} - qx_{kj}x_{mn} + q(q^2-q^{-2})x_{kn}x_{mj} \\
&\equiv 0.
\end{aligned}$$

Case 13. $f = x_{mn}x_{ij} - q^2 x_{ij}x_{mn} + qx_{in},\ g = x_{ij}x_{kl} - q^{-2}x_{kl}x_{ij},\ w = x_{mn}x_{ij}x_{kl}$.

In the case, we have

$$(f,g)_w = -q^2 x_{ij}x_{mn}x_{kl} + qx_{in}x_{kl} + q^{-2}x_{mn}x_{kl}x_{ij}.$$

There are two subcases to consider.

	$((i,j),(m,n)) \in C_5$
$((k,l),(i,j)) \in C_1$	13.1. $((k,l),(m,n)) \in C_6 \quad ((k,l),(i,n)) \in C_1$
$((k,l),(i,j)) \in C_3$	13.2. $((k,l),(m,n)) \in C_5 \quad ((k,l),(i,n)) \in C_4$

13.1. $((i,j),(m,n)) \in C_5,\ ((k,l),(i,j)) \in C_1,\ ((k,l),(m,n)) \in C_6$ and $((k,l),(i,n)) \in C_1$. Then, we have

$$\begin{aligned}(f,g)_w &= -q^2 x_{ij}x_{kl}x_{mn} + q^{-1}x_{kl}x_{in} + q^{-2}x_{kl}x_{mn}x_{ij}\\ &\equiv -x_{kl}x_{ij}x_{mn} + q^{-1}x_{kl}x_{in} + q^{-2}x_{kl}(q^2 x_{ij}x_{mn} - qx_{in})\\ &\equiv -x_{kl}x_{ij}x_{mn} + q^{-1}x_{kl}x_{in} + x_{kl}x_{ij}x_{mn} - q^{-1}x_{kl}x_{in}\\ &\equiv 0.\end{aligned}$$

13.2. $((i,j),(m,n)) \in C_5,\ ((k,l),(i,j)) \in C_3,\ ((k,l),(m,n)) \in C_5$ and $((k,l),(i,n)) \in C_4$. Then, we have $((k,n),(i,j)) \in C_2$ and

$$\begin{aligned}(f,g)_w &\equiv -q^2 x_{ij}(q^2 x_{kl}x_{mn} - qx_{kn}) + q[x_{kl}x_{in} - (q^2 - q^{-2})x_{kn}x_{il}]\\ &\quad -q^{-2}(q^2 x_{kl}x_{mn} - qx_{kn})x_{ij}\\ &\equiv -q^4 x_{ij}x_{kl}x_{mn} + q^3 x_{ij}x_{kn} + qx_{kl}x_{in} - q(q^2 - q^{-2})x_{kn}x_{il} + x_{kl}x_{mn}x_{ij}\\ &\quad -q^{-1}x_{kn}x_{ij}\\ &\equiv -q^2 x_{kl}x_{ij}x_{mn} + q^3 x_{kn}x_{ij} + qx_{kl}x_{in} - q^3 x_{kn}x_{il}\\ &\quad +q^{-1}x_{kn}x_{il} + q^2 x_{kl}x_{ij}x_{mn} - qx_{kl}x_{in} - q^{-1}x_{kn}x_{ij}\\ &\equiv 0.\end{aligned}$$

Case 14. $f = x_{mn}x_{ij} - q^2 x_{ij}x_{mn} + qx_{in},\ g = x_{ij}x_{kl} - x_{kl}x_{ij},\ w = x_{mn}x_{ij}x_{kl}$.

In the case, we have

$$(f,g)_w = -q^2 x_{ij}x_{mn}x_{kl} + qx_{in}x_{kl} + x_{mn}x_{kl}x_{ij}.$$

There are two subcases to consider.

	$((i,j),(m,n)) \in C_5$
$((k,l),(i,j)) \in C_2$	14.1. $((k,l),(m,n)),((k,l),(i,n)) \in C_2, C_3$ *or* C_4
$((k,l),(i,j)) \in C_6$	14.2. $((k,l),(m,n)),((k,l),(i,n)) \in C_6$

14.1. $((i,j),(m,n)) \in C_5,\ ((k,l),(i,j)) \in C_2$ and $((k,l),(m,n)),((k,l),(i,n)) \in C_2, C_3$ or C_4.

14.1.1. If $((k,l),(m,n))$ and $((k,l),(i,n)) \in C_2\ (l > n)$, then

$$\begin{aligned}(f,g)_w &= -q^2 x_{ij}x_{kl}x_{mn} + qx_{kl}x_{in} + x_{kl}x_{mn}x_{ij} \\ &\equiv -q^2 x_{kl}x_{ij}x_{mn} + qx_{kl}x_{in} + x_{kl}(q^2 x_{ij}x_{mn} - qx_{in}) \\ &\equiv 0.\end{aligned}$$

14.1.2. If $((k,l),(m,n))$ and $((k,l),(i,n)) \in C_3\ (l = n)$, then

$$\begin{aligned}(f,g)_w &= -x_{ij}x_{kl}x_{mn} + q^{-1}x_{kl}x_{in} + q^{-2}x_{kl}x_{mn}x_{ij} \\ &\equiv -x_{kl}x_{ij}x_{mn} + q^{-1}x_{kl}x_{in} + x_{kl}x_{ij}x_{mn} - q^{-1}x_{kl}x_{in} \\ &\equiv 0.\end{aligned}$$

14.1.3. If $((k,l),(m,n)),\ ((k,l),(i,n)) \in C_4\ (l < n)$, then we have $((k,n),(i,j)) \in C_2,\ ((i,j),(m,l)) \in C_5$ and

$$\begin{aligned}(f,g)_w &\equiv -q^2 x_{ij}[x_{kl}x_{mn} - (q^2 - q^{-2})x_{kn}x_{ml}] + q[x_{kl}x_{in} - (q^2 - q^{-2})x_{kn}x_{il}] \\ &\quad +[x_{kl}x_{mn} - (q^2 - q^{-2})x_{kn}x_{ml}]x_{ij} \\ &\equiv -q^2 x_{ij}x_{kl}x_{mn} + q^2(q^2 - q^{-2})x_{ij}x_{kn}x_{ml} + qx_{kl}x_{in} - q(q^2 - q^{-2})x_{kn}x_{il} \\ &\quad +x_{kl}x_{mn}x_{ij} - (q^2 - q^{-2})x_{kn}x_{ml}x_{ij} \\ &\equiv -q^2 x_{kl}x_{ij}x_{mn} + q^2(q^2 - q^{-2})x_{kn}x_{ij}x_{ml} + qx_{kl}x_{in} - q(q^2 - q^{-2})x_{kn}x_{il} \\ &\quad +x_{kn}(q^2 x_{ij}x_{mn} - qx_{in}) - (q^2 - q^{-2})x_{kn}(q^2 x_{ij}x_{mn} - qx_{il}) \\ &\equiv 0.\end{aligned}$$

14.2. $((i,j),(m,n)) \in C_5,\ ((k,l),(i,j)) \in C_6$ and $((k,l),(m,n)),((k,l),(i,n)) \in C_6$.

This case is similar to 14.1.1.

Case 15. $f = x_{mn}x_{ij} - q^2 x_{ij}x_{mn} + qx_{in},\ g = x_{ij}x_{kl} - x_{kl}x_{ij} + (q^2 - q^{-2})x_{kj}x_{il},\ w = x_{mn}x_{ij}x_{kl}$, with

	$((i,j),(m,n)) \in C_5$
$((k,l),(i,j)) \in C_4$	$((k,l),(m,n)) \in C_6$ $((k,l),(i,n)) \in C_4$ $((k,j),(m,n)) \in C_5$ $((i,l),(m,n)) \in C_6$

Then, we have

$$\begin{aligned}
(f,g)_w &= -q^2x_{ij}x_{mn}x_{kl}+qx_{in}x_{kl}+x_{mn}x_{kl}x_{ij}-(q^2-q^{-2})x_{mn}x_{kj}x_{il}\\
&\equiv -q^2x_{ij}x_{kl}x_{mn}+q[x_{kl}x_{in}-(q^2-q^{-2})x_{kn}x_{il}]+x_{kl}x_{mn}x_{ij}\\
&\quad -(q^2-q^{-2})(q^2x_{kj}x_{mn}-qx_{kn})x_{il}\\
&\equiv -q^2[x_{kl}x_{ij}-(q^2-q^{-2})x_{kj}x_{il}]x_{mn}+qx_{kl}x_{in}-q(q^2-q^{-2})x_{kn}x_{il}\\
&\quad +x_{kl}(q^2x_{ij}x_{mn}-qx_{in})-q^2(q^2-q^{-2})x_{kj}x_{mn}x_{il}+q(q^2-q^{-2})x_{kn}x_{il}\\
&\equiv -q^2x_{kl}x_{ij}x_{mn}+q^2(q^2-q^{-2})x_{kj}x_{il}x_{mn}+qx_{kl}x_{in}-q(q^2-q^{-2})x_{kn}x_{il}\\
&\quad +q^2x_{kl}x_{ij}x_{mn}-qx_{kl}x_{in}-q^2(q^2-q^{-2})x_{kj}x_{il}x_{mn}+q(q^2-q^{-2})x_{kn}x_{il}\\
&\equiv 0.
\end{aligned}$$

Case 16. $f=x_{mn}x_{ij}-q^2x_{ij}x_{mn}+qx_{in},\ g=x_{ij}x_{kl}-q^2x_{kl}x_{ij}+qx_{kj},\ w=x_{mn}x_{ij}x_{kl}$, with

	$((i,j),(m,n))\in C_5$
$((k,l),(i,j))\in C_5$	$((k,l),(m,n))\in C_6\quad ((k,l),(i,n)),((k,j),(m,n))\in C_5$

In the case, we have

$$\begin{aligned}
(f,g)_w &= -q^2x_{ij}x_{mn}x_{kl}+qx_{in}x_{kl}+q^2x_{mn}x_{kl}x_{ij}-qx_{mn}x_{kj}\\
&\equiv -q^2x_{ij}x_{kl}x_{mn}+q(q^2x_{kl}x_{in}-qx_{kn})+q^2x_{kl}x_{mn}x_{ij}-q(q^2x_{kj}x_{mn}-qx_{kn})\\
&\equiv -q^2(q^2x_{kl}x_{ij}-qx_{kj})x_{mn}+q^3x_{kl}x_{in}-q^2x_{kn}+q^2x_{kl}(q^2x_{ij}x_{mn}-qx_{in})\\
&\quad -q^3x_{kj}x_{mn}+q^2x_{kn}\\
&\equiv 0.
\end{aligned}$$

Thus, $\widetilde{S}^+$ is a Gröbner-Shirshov basis. This completes the proof of Theorem 3.1. □

Similarly, with the deg-lex order on $\widetilde{Y}^*$, $\widetilde{S}^-$ is a Gröbner-Shirshov basis for $U_q^-(A_N)=k\langle\widetilde{Y}|\widetilde{S}^-\rangle$.

We now use the same notation as before. Order the generators by: $x_i>x_j,\ h_i>h_i^{-1}>h_j>h_j^{-1},\ y_i>y_j$ if $i>j$, and $x_i>h_j^{\pm1}>y_m$ for all $i,\ j,\ m$. Then we obtain a well ordering (deg-lex) on $\widetilde{X}\cup H\cup\widetilde{Y}$. Thus, by Theorem 3.1, we re-obtain the following theorem in [16].

Theorem 3.2. *([16] Theorem 2.7) Let the notation be as before. Then with the deg-lex order on* $\{\widetilde{X}\cup H\cup\widetilde{Y}\}^*$, $\widetilde{S}^+\cup T\cup K\cup\widetilde{S}^-$ *is a Gröbner-Shirshov basis for* $U_q(A_N)=k\langle\widetilde{X}\cup H\cup\widetilde{Y}|\widetilde{S}^+\cup T\cup K\cup\widetilde{S}^-\rangle$.

Acknowledgement: The authors would like to express their deepest gratitude to Professor L. A. Bokut for his kind guidance, useful discussions and enthusiastic encouragement.

Received: June 2008. Revised: August 2008.

References

[1] G.M. Bergman, *The diamond lemma for ring theory*, Adv. in Math., **29**(1978), 178–218.

[2] L.A. Bokut, *Unsolvability of the word problem, and subalgebras of finitely presented Lie algebras*, Izv. Akad. Nauk. SSSR Ser. Mat., **36**(1972), 1173–1219.

[3] L.A. Bokut, *Imbeddings into simple associative algebras*, Algebra i Logika, **15**(1976), 117–142.

[4] L.A. Bokut, *Gröbner-Shirshov bases for the braid groups in the Briman-Ko-Lee generators*, sumbitted.

[5] L.A. Bokut, *Gröbner-Shirshov bases for braid groups in Artin-Garside generators*, J. Symbolic Compu., Doi:10.1016/j.jsc.2007.02.003. Available online 17 November 2007.

[6] L.A. Bokut, V.V. Chainikov and K.P. Shum, *Markov and Artin normal form theorem for braid groups*, Comm. Algebra, **35**(2007), 2105–2115.

[7] L.A. Bokut and Y. Chen, *Gröbner-Shirshov bases for free Lie algebras: after A.I. Shirshov*, SEA. Bull. Math., **31**(2007), 1057–1076.

[8] L. Bokut and Y. Chen, *Gronbner-Shirshov bases: some new results*, Proceedings of the Second International Congress in Algebra and Combinatorics, World Scientific, 2008, 35–56.

[9] L.A. Bokut, S.J. Kang, K.H. Lee and P. Malcolmson, *Gröbner-Shirshov bases for Lie superalgebras and their universal enveloping algebras*, J. Algebra, **217**(2) (1999), 461–495.

[10] L.A. Bokut and A.A. Klein, *Serre relations and Gröbner-Shirshov bases for simple Lie algebras I*, Internat. J. Algebra Comput., **6**(4) (1996), 389–400.

[11] L.A. Bokut and A.A. Klein, *Serre relations and Gröbner-Shirshov bases for simple Lie algebras II*, Internat. J. Algebra Comput., **6**(4) (1996), 401–412.

[12] L.A. Bokut and A.A. Klein, *Gröbner-Shirshov bases for exceptional Lie algebras I*, J. Pure Applied Algebra, **133**(1998), 51–57.

[13] L.A. Bokut and A.A. Klein, *Gröbner-Shirshov bases for exceptional Lie algebras E6, E7, E8*, Algebra and Combinatorics (Hong Kong), 37–46, Springer, Singapore, 1999.

[14] L.A. Bokut and P. Kolesnikov, *Gröbner-Shirshov bases: from their incipiency to the prensent*, Journal of Mathematical Sciences, **116**(1) (2003), 2894–2916.

[15] L.A. Bokut and P. Kolesnikov, *Gröbner-Shirshov bases, conformal algebras and pseudo-algebras*, Journal of Mathematical Sciences, **131**(2005), 5962–6003.

[16] L.A. Bokut and P. Malcolmson, *Gröbner-Shirshov basis for quantum enveloping algebras*, Israel Journal of Mathematics, **96** (1996), 97–113.

[17] L.A. Bokut and L.-S. Shiao, *Gröbner-Shirshov bases for Coxeter groups*, Comm. Algebra, **29**(9) (2001), 4305–4319.

[18] B. Buchberger, *An algorithm for finding a basis for the residue class ring of a zero-dimensional polynomial ideal*, [in German], Ph.D thesis, University of Innsbruck, Austria, 1965.

[19] B. Buchberger, *An algorithmical criteria for the solvability of algebraic systems of equations*, [in German], Aequationes Math., **4**(1970), 374–383.

[20] V.G. Drinfeld, *Hopf algebra and the quantum Yang-Baxter equation*, Doklady Adakemii Nauk SSSR, **283**(5) (1985), 1060–1064.

[21] H. Hironaka, *Resolution of singularities of an algebraic vareity over a field of characteristic zero I, II*, Ann. of Math., **79**(1964), 109–203; ibid. **79**(1964), 205–326.

[22] M. Jimbo, *A q-difference analogue of $U(G)$ and the Yang-Baxter equation*, Letters in Mathematical Physics, **10**(1) (1985), 63–69.

[23] S.-J. Kang and K.-H. Lee, *Gröbner-Shirshov bases for repesentation theory*, J. Korean Math. Soc., **37**(1) (2000), 55–72.

[24] S.-J. Kang and K.-H. Lee, *Gröbner-Shirshov basis theory for irreducible $sl_{(n+1)}-$modules*, J. Algebra, **232**(2000), 1–20.

[25] S.-J. Kang, I.-S. Lee, K.-H. Lee and H. OH, *Hecke algebras, Speche modules and Gröbner-Shirshov bases*, J. Algebra, **252**(2002), 258–292.

[26] M. Kashiwara, *Crystalizing the q-analogue of universal enveloping algebras*, Commun. Math. Phys., **133**(1990), 249–260.

[27] M. Kashiwara, *On crystal bases of the q-analogue of universal enveloping algebras*, Duke Math. J., **63**(1991), 465–516.

[28] V.K. Kharchenko, *A quantum analog of the Poincare-Birkhoff-Witt theorem*, Algebra i Logika, **38**(4) (1999), 259–276.

[29] G. Lusztig, *Canonical bases arising from quantized enveloping algebra*, J. Amer. Math. Soc., **3**(2) (1990), 447–498.

[30] G. Lusztig, *Canonical bases arising from quantized enveloping algebra II*, Progr. Theor. Phys. Suppl., **102** (1990), 175-201.

[31] E.N. Poroshenko, *Gröbner-Shirshov bases for the Kac-Moody algebras of the type $A_n^{(1)}$*, Comm. Algebra, **30**(6) (2002), 2617–2637.

[32] E.N. Poroshenko, *Gröbner-Shirshov bases for the Kac-Moody algebras of the type $C_n^{(1)}$ and $D_n^{(1)}$*, Vestn. Novosib. Gos. Univ., Ser. Mat. Mekh. Inform., **2**(1)(2002), 58–70, (in Russian).

[33] E.N. Poroshenko, *Gröbner-Shirshov bases for the Kac-Moody algebras of the type* $B_n^{(1)}$, Int. J. Math. Game Theory Algebra, **13**(2) (2003), 117–128.

[34] M. Rosso, *An analogue of the Poincare-Birkhoff-Witt theorem and the universal R-matrix of* $U_q(sl(N+1))$, Comm. Math. Phys., **124**(2) (1989), 307–318.

[35] A.I. Shirshov, *Some algorithmic problem for Lie algebras*, Sibirsk. Mat. Z., **3**(1962), 292–296, (in Russian); English translation in SIGSAM Bull., **33**(2) (1999), 3–6.

[36] I. Yamane, *A Poincare-Birkhoff-Witt theorem for quantized universal enveloping algebras of type* A_N, Publ., RIMS. Kyoto Univ., **25**(3) (1989), 503–520.

[37] J.C. Jantzen, *Lectures on quantum groups*, Graduate Texts in Mathematics, Vol. 6, AMS, 1996.

CUBO A Mathematical Journal
Vol.10, Nº03, (195–210). October 2008

On the Structure of Primitive n-Sum Groups

Eloisa Detomi and Andrea Lucchini
Dipartimento di Matematica Pura ed Applicata,
Via Trieste 63, 35121 Padova, Italy
email: detomi@math.unipd.it
email: lucchini@math.unipd.it

ABSTRACT

For a finite group G, let $\sigma(G)$ be least cardinality of a collection of proper subgroups whose set-theoretical union is all of G. We study the structure of groups G containing no normal non-trivial subgroup N such that $\sigma(G/N) = \sigma(G)$.

RESUMEN

Para un grupo G, sea $\sigma(G)$ la menor cardinalidad de la colección de subgrupos propios cujas union (de conjuntos) es todo G. Nosotros estudiamos la estructura de grupos G contiendo no trivial no normales subgrupos N tal que $\sigma(G/N) = \sigma(G)$.

Key words and phrases: *n-sum groups; minimal coverings; monolithic groups.*

Math. Subj. Class.: *20D60.*

1 Introduction

If G is a non cyclic finite group, then there exists a finite collection of proper subgroups whose set-theoretical union is all of G; such a collection is called a *cover* for G. A minimal cover is one of least cardinality and the size of a minimal cover of G is denoted by $\sigma(G)$ (and for convenience we shall write $\sigma(G) = \infty$ if G is cyclic). The study of minimal covers was introduced by J.H.E. Cohn [8]; following his notation, we say that a finite group G is an n-sum group if $\sigma(G) = n$ and that a group G is a *primitive n-sum group* if $\sigma(G) = n$ and G has no normal non-trivial subgroup N such that $\sigma(G/N) = n$. We will say that G is *σ-primitive* if it is a primitive n-sum group for some integer n. Notice that if N is a normal subgroup of G, then $\sigma(G) \le \sigma(G/N)$; indeed a cover of G/N can be lifted to a cover of G.

It is clear that if G is a non cyclic monolithic primitive group (i.e. G has a unique minimal normal subgroup and the Frattini subgroup of G is trivial) and $G/\operatorname{soc}(G)$ is cyclic, then G is a σ-primitive group.

Moreover Cohn proved that an abelian σ-primitive group is the direct product of two cyclic groups of order p, a prime number.

Tomkinson [14] showed that in a finite solvable group G, $\sigma(G) = |V| + 1$, where V is a chief factor of G with least order among chief factors of G with multiple complements. This allows to prove (see for example [5]) that a σ-primitive solvable group G is as described above, i.e. either G is abelian or G is monolithic and $G/\operatorname{soc}(G)$ is cyclic.

However there exist examples of σ-primitive groups with $G/\operatorname{soc}(G)$ non cyclic: actually with $G/\operatorname{soc}(G) \cong \operatorname{Alt}(p)$ for some prime p (see Corollary 9 and Corollary 12).

The aim of this paper is to collect information on the structure of the σ-primitive groups. In particular we prove that *if G is σ-primitive, then G contains at most one abelian minimal normal subgroup; moreover two non-abelian minimal normal subgroups of G are not G-equivalent* (we refer to an equivalence relation among the chief factors of a finite group introduced in [10] and [9], whose main properties are summarized at the beginning of Section 2). *Furthermore if G is non-abelian, then all the solvable factor groups of $G/\operatorname{soc}(G)$ are cyclic.*

No example is known of a non-abelian σ-primitive group containing two distinct minimal normal subgroups. This leads to conjecture that a non-abelian σ-primitive group is monolithic. We prove a partial result supporting this conjecture.

Theorem 1. *Let G be a σ-primitive group with no abelian minimal normal subgroups. Then either G is a primitive monolithic group and $G/\operatorname{soc}(G)$ is cyclic, or $G/\operatorname{soc}(G)$ is non-solvable and all the non-abelian composition factors of $G/\operatorname{soc}(G)$ are alternating groups of odd degree.*

A better knowledge of the σ-primitive groups is useful in dealing with several questions about the minimal covers. For example, confirming a conjecture of Tomkinson, we prove:

Theorem 2. *There is no finite group G with $\sigma(G) = 11$.*

Another application concerns the study of $\sigma(G)$ when $G = H \times K$ is a direct product of two finite groups. Cohn proved that if H and K have coprime order, then $\sigma(H \times K) = \min\{\sigma(H), \sigma(K)\}$. We prove the following more general result:

Theorem 3. *Let $G = H \times K$ be the direct product of two subgroups. If no alternating group* $\mathrm{Alt}(n)$ *with n odd is a homomorphic image of both H and K, then either $\sigma(G) = \min\{\sigma(H), \sigma(K)\}$ or $\sigma(G) = p + 1$ and the cyclic group of order p is a homomorphic image of both H and K.*

2 Preliminary results and easy remarks

For the reader's convenience, we recall the definition of an equivalence relation among the elements of the set $\mathcal{CF}(G)$ of the chief factors of G, that was introduced in [10] and studied in details in [9]. A group G is said to be primitive if it has a maximal subgroup with trivial core. The socle $\mathrm{soc}(G)$ of a primitive group G can be either an abelian minimal normal subgroup (I), or a non-abelian minimal normal subgroup (II), or the product of two non-abelian minimal normal subgroups (III); we say respectively that G is *primitive of type I, II, III* and in the first two cases we say that G is *monolithic*. Two chief factors of a finite group G are said to be G-*equivalent* if either they are G-isomorphic between them or to the two minimal normal subgroups of a primitive epimorphic image of type III of G. This means that two G-equivalent chief factors of G are either G-isomorphic between them or to two chief factors of G having a common complement (which is a maximal subgroup of G). A chief factor H/K is called Frattini if $H/K \leq \Phi(G)$. For any $A \in \mathcal{CF}(G)$ we denote by $I_G(A)$ the set of those elements of G which induces by conjugation an inner automorphism in A. Moreover we denote by $R_G(A)$ the intersection of the normal subgroups N of G contained in $I_G(A)$ and with the property that $I_G(A)/N$ is non-Frattini and G-equivalent to A. We collect here a sequence of basic properties of the subgroups $I_G(A)$ and $R_G(A)$, proved and discussed in [9]:

Proposition 4. *Let $A \in \mathcal{CF}(G)$ and let $I/R = I_G(A)/R_G(A)$. Then:*

1. *either $R = I$, in which case we set $\delta_G(A) = 0$, or $I/R = \mathrm{soc}(G/R)$ and it is a direct product of $\delta_G(A)$ minimal normal subgroups G-equivalent to A;*
2. *each chief series of G contains exactly $\delta_G(A)$ non-Frattini chief factors G-equivalent to A;*
3. *if A is abelian, then I/R has a complement in G/R;*
4. *if $\delta_G(A) \geq 2$, then any two different minimal normal subgroups of I/R have a common complement, which is a maximal subgroup;*
5. *a chief factor H/K of G is non-Frattini and G-equivalent to A if and only if $RH/RK \neq 1$ and $RH \leq I$.*

Note that if $\delta_G(A) = 1$, then $G/R_G(A)$ is a monolithic primitive group (the monolithic primitive group associated to A).

In the rest of the section we will discuss some basis results on the relation between $\sigma(G)$ and $\sigma(G/N)$ when N is a minimal normal subgroup of G. We start summarizing some known properties of σ.

Lemma 5. *Let N be a minimal normal subgroup of a group G. If $\sigma(G) < \sigma(G/N)$, then*

1. *if N has c complements which are maximal subgroups, then $c+1 \leq \sigma(G)$;*
2. *if $N = S^r$ where S is a non-abelian simple group and $l(S)$ is the minimal index of a maximal subgroup of S, then $l(S)^r + 1 \leq \sigma(G)$.*

Proof. (1) (See e.g. [14, Proof of Theorem 2.2]) Let $\mathcal{M} = \{M_i \mid i = 1, \dots, \sigma(G)\}$ be a set of maximal subgroups whose union covers G and let M be a complement of N. Clearly $M = \bigcup_{1 \leq i \leq \sigma(G)} M \cap M_i$, however $\sigma(M) = \sigma(G/N) > \sigma(G)$, hence $M = M \cap M_i$ for some i; in particular if M is a maximal subgroup of G, then $M = M_i \in \mathcal{M}$. So $\mathcal{M}$ contains all the c complements of N which are maximal; since the union of these complements does not cover N, we need at least $c+1$ subgroups in $\mathcal{M}$.

(2) Let $l_G(N)$ be the smallest index of a proper subgroup of G supplementing N. By Lemma 3.2 in [14] a minimal cover $\mathcal{M}$ of G contains at least $l_G(N)$ subgroups which supplement N. On the other hand, if all the subgroups in $\mathcal{M}$ are supplements of N, then by [8, Lemma 1] we have $l_G(N) \leq \sigma(G) - 1$. In any case we conclude $\sigma(G) \geq l_G(N) + 1 \geq l(S)^r + 1$. □

Corollary 6. *Let N be a minimal normal subgroup of a group G. If $\sigma(G) < \sigma(G/N)$, then*

1. *if N is abelian, complemented and non-central, then $|N| + 1 \leq \sigma(G)$;*
2. *if $N = S^r$ where S is a non-abelian simple group, then $5^r + 1 \leq \sigma(G)$.*

Proposition 7. *Let N be a non-solvable normal subgroup of a finite group G. Then $\sigma(G) \leq |N| - 1$.*

Proof. Consider the centralizers in G of the nontrivial elements of N: if there exists an element $g \in G$ which does not belong to $\bigcap_{1 \neq n \in N} C_G(n)$ then the subgroup $\langle g \rangle$ acts fixed point freely on N. By the classification of finite simple groups (see e.g. [15]), it follows that N is solvable, a contradiction. Hence $\sigma(G) \leq |N| - 1$. □

Corollary 8. *If N is a non-abelian minimal normal subgroup of G and $\delta_G(N) > 1$, then $\sigma(G) = \sigma(G/N)$.*

Proof. Assume by contradiction that $\sigma(G) < \sigma(G/N)$. Since $\delta_G(N) > 1$, there exists a maximal subgroup M of G, such that G/M_G is a primitive group of type III and M/M_G is a common complement of the two minimal normal subgroups of the socle $H/M_G \times NM_G/M_G$ of G/M_G. In particular M is a non-normal complement of N and it has $|N|$ conjugates, hence $|N| + 1 \leq \sigma(G)$ by Lemma 5. This contradicts Proposition 7. □

Corollary 9. *Let p a large prime not of the form $(q^k - 1)/(q - 1)$ where q is a prime power and k an integer; then $\sigma(\mathrm{Alt}(5) \wr \mathrm{Alt}(p)) < \sigma(\mathrm{Alt}(p))$.*

Proof. By Proposition 7, $\sigma(\mathrm{Alt}(5) \wr \mathrm{Alt}(p)) < |\mathrm{Alt}(5)|^p$. On the other hand, by Theorem [12, 4.4], $\sigma(\mathrm{Alt}(p)) \geq (p-2)! > 60^p$ for a large enough prime not of the form $(q^k - 1)/(q - 1)$. □

Proposition 10. *Let G be a finite group. If V is a complemented normal abelian subgroup of G and $V \cap Z(G) = 1$, then $\sigma(G) < 2|V|$. In particular, if V is a minimal normal subgroup, then $\sigma(G) \leq 1 + q + \cdots + q^n$ where $q = |\operatorname{End}_G(V)|$ and $|V| = q^n$.*

Proof. Let H be a complement of V in G; we shall prove that G is covered by the family of subgroups $\mathcal{A} = \{H^v \mid v \in V\} \cup \{C_H(v)V \mid 1 \neq v \in V\}$. Let $g = hw \in G$, where $h \in H$, $w \in V$. If $h \notin C_H(v)$ for every $v \in V \setminus \{1\}$, then $C_V(h) = 1$ and the cardinality of the set $\{h^v \mid v \in V\}$ is $|V : C_V(h)| = |V|$. Therefore $\{h^v \mid v \in V\} = \{hv \mid v \in V\}$ and $g = hw \in H^v$ for some $v \in V$. Thus $\sigma(G) \leq |\mathcal{A}| \leq |V| + (|V| - 1) < 2|V|$. In particular, if V is H-irreducible, then $\operatorname{End}_G(V) = \operatorname{End}_H(V) = \mathbb{F}$ is a finite field. We may identify H with a subgroup of $\mathrm{GL}(n, q)$, where $|\mathbb{F}| = q$ and $\dim_{\mathbb{F}} V = n$. In this case G is covered by $\mathcal{A} = \{H^v \mid v \in V\} \cup \{C_H(W)V \mid W \leq V,\ \dim_{\mathbb{F}} W = 1\}$, so $\sigma(G) \leq q^n + (1 + \cdots + q^{n-1})$. □

Corollary 11. *Let H be a finite group, V an H-module, $G = V \rtimes H$ the semidirect product of V by H and assume that $C_V(H) = 0$. Then*

1. *if $\mathrm{H}^1(H, V) \neq 0$, then $\sigma(G) = \sigma(H)$;*
2. *if $\sigma(H) \geq 2|V|$, then $\mathrm{H}^1(H, V) = 0$.*

Proof. Assume by contradiction that $\sigma(G) < \sigma(H)$. By Lemma 5, $c + 1 \leq \sigma(G)$ where c is the number of complements of V in G. If $\mathrm{H}^1(H, V) \neq 0$, then there are at least two conjugacy classes of complements for V in G and, since $C_V(H) = 0$, any conjugacy class consists of $|V|$ complements, hence $c \geq 2|V|$ and $\sigma(G) > 2|V|$ against Proposition 10. □

Corollary 12. *Let V the fully deleted module for $\mathrm{Alt}(n)$ over $\mathbb{F}_2$ and let G be the semidirect product of V by $\mathrm{Alt}(n)$.*

1. *If $n = p$ is a large odd prime not of the form $(q^k - 1)/(q - 1)$ where q is a prime power and k an integer, then $\sigma(G) < \sigma(\mathrm{Alt}(n))$.*
2. *If n is even, then $\sigma(G) = \sigma(\mathrm{Alt}(n))$*

Proof. 1) Since $|V| = 2^{p-1}$ (see e.g. [11, Prop. 5.3.5]), Proposition 10 gives that $\sigma(G) < 2|V| < 2^p$. On the other hand, by Theorem [12, 4.4], $\sigma(\mathrm{Alt}(p)) \geq (p - 2)! > 2^p$ for a large enough prime not of the form $(q^k - 1)/(q - 1)$.

2) This follows from Corollary 11 and the fact that $\mathrm{H}^1(\mathrm{Alt}(n), V) \neq 0$ whenever n is even (see e.g. [2, p. 74]). □

Corollary 13. *Let $V \neq W$ be non-Frattini non-central abelian minimal normal subgroups of G. Then*

1. *if $\delta_G(V) > 1$, then $\sigma(G) = \sigma(G/V)$;*

2. $\sigma(G) = \min\{\sigma(G/V), \sigma(G/W)\}$.

Proof. 1) By a result in [3], the number c of complements of V in G is

$$c = |\operatorname{Der}(G/V, V)| = |\operatorname{End}_{G/V}(V)|^{\delta_G(V)-1} |\operatorname{Der}(G/C_G(V), V)|$$

hence $c \geq 2|V|$ whenever $\delta_G(V) > 1$. If $\sigma(G) < \sigma(G/V)$, then by Lemma 5 and Proposition 10, $2|V| < c + 1 \leq \sigma(G) < 2|V|$, a contradiction.

2) If V and W are G-equivalent, then by (1) $\sigma(G) = \sigma(G/V) = \sigma(G/W)$. So assume that V and W are not G-equivalent and, by contradiction, that $\sigma(G) < \min\{\sigma(G/V), \sigma(G/W)\}$. A complement of V in G has at least $|V|$ conjugates and it is a maximal subgroup of G, so we can find at least $|V|$ complements of V. In the same way there are at least $|W|$ distinct complements of $|W|$ in G. Moreover, since V and W are not G-equivalent, V and W cannot have a common complement. Arguing as in Lemma 5 we see that all the complements of V and W have to appear in a minimal cover of G. Therefore $\sigma(G) \geq |V| + |W| \geq \min\{2|V|, 2|W|\}$, against Proposition 10.

□

3 The structure of σ-primitive groups

We collect some known properties of σ-primitive groups and some consequences of the previous section.

Corollary 14. *Let G be a non-abelian σ-primitive group. Then:*

1. $Z(G) = 1$;
2. *the Frattini subgroup of G is trivial;*
3. *if N is a minimal normal subgroup of G, then $\delta_G(N) = 1$;*
4. *there is at most one abelian minimal normal subgroup of G;*
5. *the socle* $\operatorname{soc}(G) = G_1 \times \cdots \times G_n$ *is a direct product of non-G-equivalent minimal normal subgroups and at most one of them is abelian.*
6. *G is a subdirect product of the monolithic primitive groups $X_i = G/R_G(G_i)$ associated to the minimal normal subgroups G_i, $1 \leq i \leq n$.*

Proof. Part (1) is Theorem 4 in [8]. If $\Phi(G)$ is the Frattini subgroup of G and H is a proper subgroup of G, then also $H\,\Phi(G)$ is a proper subgroup of G. Hence we can assume that $\Phi(G)$ is contained in every subgroup of a minimal cover of G so that $\sigma(G) = \sigma(G/\Phi(G))$ and therefore (2) holds. Parts (3) and (4) follows from Corollaries 8 and 13. Then (3) and (4) implies (5). To prove (6) we consider the intersection $R = \bigcap_{i=1}^n R_G(G_i)$. If $R \neq 1$, then R contains a minimal normal subgroup N of G. By (2) and (5), N is non-Frattini and G-equivalent to G_i for some $1 \leq i \leq n$. Hence by Proposition 4 (5), $R_G(G_i)N \neq R_G(G_i)$, in contradiction with $N \leq R \leq R_G(G_i)$. $\square$

Definition 15. *Let X be a primitive monolithic group and let N be its socle. For any non-empty union $\Omega = \bigcup_i \omega_i N$ of cosets of N in X with the property that $\langle \Omega \rangle = X$, define $\sigma_\Omega(X)$ to be the minimum number of supplements of N in G needed to cover Ω. Then we define*

$$\sigma^*(X) = \min \left\{ \sigma_\Omega(X) \mid \Omega = \bigcup_i \omega_i N,\ \langle \Omega \rangle = X \right\}.$$

Proposition 16. *Let G be a non-abelian σ-primitive group, $G_1, \dots, G_n$ the minimal normal subgroups, and $X_1, \dots X_n$ the monolithic primitive groups associated to G_i, $i = 1, \dots n$. Then $\sigma(G) \geq \sum_{i=1}^n \sigma^*(X_i)$.*

Proof. Let $\mathcal{M}$ be a set of $\sigma = \sigma(G)$ maximal subgroups whose union is G. Define $\mathcal{M}_{\neg G_i} = \{M \in \mathcal{M} \mid M \not\geq G_i\}$; note that

- $\mathcal{M}_{\neg G_i} \neq \varnothing$ for each $1 \leq i \leq n$; otherwise every maximal subgroup of $\mathcal{M}$ would contain G_i and the set $\{M/G_i \mid M \in \mathcal{M}\}$ would cover G/G_i with $\sigma(G) < \sigma(G/G_i)$ subgroups.
- $\mathcal{M}_{\neg G_i} \cap \mathcal{M}_{\neg G_j} = \varnothing$ for $i \neq j$; indeed if there exists $M \in \mathcal{M}_{\neg G_i} \cap \mathcal{M}_{\neg G_j}$, then $G_i M_G/M_G$ and $G_j M_G/M_G$ are minimal normal subgroups of the primitive group G/M_G, hence $\delta_G(G_i) \geq 2$, contrary to Corollary 14.

Therefore $\mathcal{M}$ contains the disjoint union of the non-empty sets $\mathcal{M}_{\neg G_i}$, $1 \leq i \leq n$, and we are reduced to prove that $|\mathcal{M}_{\neg G_i}| \geq \sigma^*(X_i)$, for every i. Let us fix an index i and let $\pi : G \mapsto X$ be the projection of G over $X = X_i$. We set $N = \operatorname{soc} X \cong G_i$, $\mathcal{M}_i = \{M \in \mathcal{M} \mid M \geq G_i\} = \mathcal{M} \setminus \mathcal{M}_{\neg G_i}$ and

$$\Omega = \left\{ \pi(g) \mid g \in G \setminus \bigcup_{M \in \mathcal{M}_i} M \right\}.$$

By minimality of the cover $\mathcal{M}$, $G \neq \bigcup_{M \in \mathcal{M}_i} M$ hence $\Omega \neq \varnothing$. Moreover, as $G_i \leq M \in \mathcal{M}_i$ and $\pi(G_i) = \operatorname{soc} X = N$, we get that for every $x \in \Omega$ the coset xN is contained in Ω. If $\langle \Omega \rangle = H \neq X$, then G is covered by the set $\mathcal{M}_i \cup \{\pi^{-1}(H)\}$ and this actually is a minimal cover of G, since $|\mathcal{M}_i| + 1 \leq \sigma$. But then, as $\pi^{-1}(H) \geq G_i$, we would have $\sigma(G/G_i) \leq |\mathcal{M}_i| + 1 = \sigma(G)$, a contradiction. Hence $\langle \Omega \rangle = X$.

Now we shall prove that $|\mathcal{M}_{\neg G_i}| \geq \sigma_\Omega(X) \geq \sigma^*(X)$. By [9, Proposition 11] the kernel $R = R_G(G_i)$ of the projection π_i of G over X has the property that if H is a proper subgroup of G such that $HG_i = G$ then $HR \neq G$. Therefore every maximal subgroup $M \in \mathcal{M}_{\neg G_i}$ contains R, $M = \pi^{-1}(\pi(M))$ and $\pi(M)$ is a a maximal subgroup of X supplementing N. Clearly, as $\bigcup_{M \in \mathcal{M}_{\neg G_i}} M$ covers $G \setminus \bigcup_{M \in \mathcal{M}_i} M$, we have that $\bigcup_{M \in \mathcal{M}_{\neg G_i}} \pi(M)$ covers Ω. Therefore $|\{\pi(M) \mid M \in \mathcal{M}_{\neg G_i}\}| = |\mathcal{M}_{\neg G_i}| \geq \sigma_\Omega(X) \geq \sigma^*(X)$. $\square$

Remark 17. *For every primitive monolithic group X_i, $\sigma^*(X_i) \leq l_{X_i}(\mathrm{soc}(X_i))$, where $l_{X_i}(\mathrm{soc}(X_i))$ is the smallest index of a proper subgroup of X_i supplementing $\mathrm{soc}(X_i)$. Indeed, if a supplement of $N_i = \mathrm{soc}(X_i)$ in X_i has non trivial intersection with a coset gN_i, then $|gN_i \cap M| = |N_i \cap M| = |gN_i|/|G : M|$, and therefore we need at least $l_{X_i}(\mathrm{soc}(X_i))$ supplements to cover gN_i. So in particular the previous proposition implies that $\sigma(G) \geq \sum_{i=1}^n l_{X_i}(N_i)$.*

Lemma 18. *Let N be a normal subgroup of a group X. If a set of subgroups covers a coset yN of N in X, then it also covers every coset $y^\alpha N$ with α prime to $|y|$.*

Proof. Let s be an integer such that $s\alpha \equiv 1 \mod |y|$. As s is prime to $|y|$, by a celebrated result of Dirichlet, there exists infinitely many primes in the arithmetic progression $\{s + |y|n \mid n \in \mathbb{N}\}$; we choose a prime $p > |X|$ in $\{s + |y|n \mid n \in \mathbb{N}\}$. Clearly, $y^p = y^s$. As p is prime to $|X|$, there exists an integer r such that $pr \equiv 1 \mod |X|$. Hence, if $yN \subseteq \cup_{i\in I} M_i$, for every $g \in y^\alpha N$ we have that $g^p \in (y^\alpha)^p N = (y^\alpha)^s N = yN \subseteq \cup_{i\in I} M_i$ and therefore also $g = (g^p)^r$ belongs to $\cup_{i\in I} M_i$. □

Corollary 19. *Let G be a non-abelian σ-primitive group, N a minimal normal subgroup and X the monolithic primitive groups associated to N. Then:*

1. *if $X = N$, then $G = N$;*
2. *if $|X/N|$ is a prime, then $G = X$.*

Proof. Note that if $X = N$, then there is only one coset of N in X hence $\Omega = N$, $\sigma^*(N) = \sigma_N(N) = \sigma(N)$. By Proposition 16, $\sigma^*(N) = \sigma(N) \leq \sigma(G)$. As $N = X$ is a homomorphic image of G, we get $G = N$.

Now let $|X/N|$ be a prime. Let Ω be a non-empty union of cosets of N in X with the property that $\langle\Omega\rangle = X$; then Ω contains a coset yN which is a generator for X/N. By Lemma 18 we have that if $\bigcup_i M_i$ covers Ω, then $\bigcup_i M_i$ covers every coset of N with the exception, at most, of the subgroup N itself. Hence, $\sigma(X) \leq \sigma_\Omega(X) + 1$ that is $\sigma^*(X) \geq \sigma(X) - 1$. By Proposition 16, $\sigma(G) \geq \sum_{i=1}^n \sigma^*(X_i)$. Moreover, by Remark 16, $\sigma^*(X_i) \geq 2$. Therefore, as $\sigma(G) \leq \sigma(X)$, there is no room for another minimal normal subgroup in G. □

Corollary 20. *If $N = \mathrm{Alt}(n)$, $n \neq 6$, is a normal subgroup of G, then either $\sigma(G) = \sigma(G/N)$ or $G \in \{\mathrm{Sym}(n), \mathrm{Alt}(n)\}$.*

Proof. It is sufficient to consider a σ-primitive image of G and then apply Corollary 19. □

Actually, the corollary holds also for $n = 6$, thanks to the following proposition.

Proposition 21. *Let G be a σ-primitive group and let $\mathrm{O}^\infty(G)$ be the smallest normal subgroup of G such that $G/\mathrm{O}^\infty(G)$ is solvable. If G is non solvable, then $G/\mathrm{O}^\infty(G)$ is a cyclic group.*

Proof. By Corollary 14, G is a subdirect product of the monolithic primitive groups X_i associated to the minimal normal subgroups G_i, $1 \leq i \leq n$; call $N_i = \mathrm{soc}(X_i) \cong G_i$. Let $\mathcal{M}$ be a set of $\sigma = \sigma(G)$ maximal subgroups whose union is G and define $\mathcal{M}_{\neg G_i} = \{M \in \mathcal{M} \mid M \not\geq G_i\}$. Let m_i be the minimal index of a supplement of N_i in X_i: by Remark 17, $\sigma(G) \geq \sum_{i=1}^{n} m_i$.

Let $R = \mathrm{O}^{\infty}(G)$ and assume by contradiction that G/R is not cyclic. Then, by Tomkinson's result [14], $\sigma(G/R) = q + 1$ where q is the order of the smallest chief factor $A = H/K$ of G/R having more than a complement in G/R. As G is not solvable, then $\sigma(G) < \sigma(G/R) = q + 1$. Since G is the subdirect product of the X_i's, without loss of generality we can assume that A is a chief factor of $X = X_1$.

If $N = \mathrm{soc}(X)$ is an elementary abelian p-group, then, by Corollary 6 and Corollary 14 (1), $|N| + 1 \leq \sigma(G) < q + 1$. Therefore $|N| < q$ and A is a chief factor, say U/V, of an irreducible linear group $X/N \leq GL(N)$ acting on N. By Clifford Theorem, U is a completely reducible linear group hence $\mathrm{O_p}(U) = 1$. Then, by Theorem 3 in [4], $|U/U'| < |N| < q$, against $|A| = |U/V| = q$.

Assume now that $N = S$ is a simple non-abelian group. Then A is isomorphic to a chief factor of a subgroup of $\mathrm{Out}(S)$ hence $q = |A| \leq |\mathrm{Out}(S)| < m_1$ (see e.g. Lemma 2.7 [4]). But $\sigma(G) \geq \sum_{i=1}^{n} m_i \geq m_1 > q$, against $\sigma(G) < q + 1$.

We are left with the case $N = S^r$ where S is a simple non-abelian group. Then X/N is isomorphic to a subgroup of $\mathrm{Out}(S) \wr \mathrm{Sym}(r)$. If A is isomorphic to a chief factor of a transitive subgroup of $\mathrm{Sym}(r)$, then Theorem 2 in [4] gives that $q = |A| \leq 2^r < (n_1)^r \leq m_1$, where n_1 is the minimal index of a subgroup of S. But this contradicts $m_1 \leq \sigma(G) \leq q$. Therefore A has to be a chief factor of a subgroup of $\mathrm{Out}(S)^r$. Then $q = |A| \leq |\mathrm{Out}(S)|^r \leq n_1^r \leq m_1$ gives the final contradiction. □

Lemma 22. *Let G be a non-solvable transitive permutation group of degree n. Then either $\sigma(G) \leq 4^n$ or every non-abelian composition factor of G is isomorphic to an alternating group of odd degree.*

Proof. Let G be a non-solvable transitive permutation group of degree n. We can embed G into a wreath product of its primitive components, let say $G \leq K_1 \wr K_2 \wr \cdots \wr K_t$ where K_i is a primitive permutation group of degree n_i and $n_1 n_2 \cdots n_t = n$ (see for example [7]). Let K_j be a non-solvable component and assume that K_j is not an alternating or symmetric group of odd degree; then G has an homomorphic image $\overline{G}$ which is embedded in a wreath product $K \wr H$ where $K = K_j$ is a permutation group of degree $a = n_j$ and H has degree b with $ab \leq n$. If K does not contain $\mathrm{Alt}(a)$ then $|K| \leq 4^a$ [13] and $\overline{G}$ has a non-solvable normal subgroup of order at most 4^{ab}. By Proposition 7 this implies that $\sigma(G) \leq \sigma(\overline{G}) \leq 4^{ab} \leq 4^n$. So assume that K contains $\mathrm{Alt}(a)$ where a is even. We identify $\overline{G}$ with its image in $K \wr H$: $\overline{G}$ is a transitive group of degree ab, with a system of imprimitivity $\mathcal{B}$ with blocks of size a and K is the permutation group induced on a block by its stabilizer. Let $\mathcal{M}_1$ be the set of subgroups $\overline{G} \cap M$ where M is a maximal intransitive subgroup of $\mathrm{Sym}(ab)$ and let $\mathcal{M}_2$ be the set of subgroups $\overline{G} \cap (M \wr H)$ where $M \cong \mathrm{Sym}(a/2) \wr \mathrm{Sym}(2)$ is a maximal imprimitive subgroup of $\mathrm{Sym}(a)$; if $T \in M_2$ and $B \in \mathcal{B}$, then the permutation group induced on B by the stabilizer T_B is isomorphic to the imprimitive proper subgroup $\mathrm{Sym}(a/2) \wr \mathrm{Sym}(2)$ of K, hence T is a proper subgroup of $\overline{G}$. Now let $x \in \overline{G}$: if x is not a cycle of length ab then there exists $T \in \mathcal{M}_1$ containing x; otherwise there exists $T \in \mathcal{M}_2$ containing x. Hence the set $\mathcal{M}_1 \cup \mathcal{M}_2$ covers $\overline{G}$ with

$$\sum_{i=1}^{ab/2} \binom{ab}{i} + \frac{1}{2}\binom{a}{a/2} \leq 2^{ab} \leq 2^n$$

proper subgroups. Therefore $\sigma(G) \leq 2^n$. □

Proposition 23. *Let G be a σ-primitive group with a non-abelian minimal normal subgroup N. If $G/NC_G(N)$ is not cyclic, then all the non-abelian composition factors of $G/NC_G(N)$ are alternating groups of odd degree.*

Proof. Let $N = S^r$, where S is a non-abelian simple group. By Corollary 6, $5^r + 1 \leq \sigma(G)$. Denote by X the monolithic primitive group associated to the G-group N; then X is a subgroup of $\mathrm{Aut}(S) \wr \mathrm{Sym}(r)$. Let K be the image of X in $\mathrm{Sym}(r)$. If K is solvable, then, by Schreier Conjecture, $X/\mathrm{soc}(X) \cong G/NC_G(N)$ is solvable. By Proposition 21 it follows that $G/NC_G(N)$ is cyclic.

Thus, if $G/NC_G(N)$ is not cyclic, then K is non-solvable. Since $5^r + 1 \leq \sigma(G) \leq \sigma(K)$, the previous lemma implies that every non-abelian composition factor of K is an alternating group of odd degree. Then, by Schreier Conjecture, the same holds for $G/NC_G(N)$. □

Theorem 24. *Let G be a σ-primitive group with no abelian minimal normal subgroups. Then either G is a primitive monolithic group and $G/\mathrm{soc}(G)$ is cyclic, or $G/\mathrm{soc}(G)$ is non-solvable and all the non-abelian composition factors of $G/\mathrm{soc}(G)$ are alternating groups of odd degree.*

Proof. By Corollary 14, G is a subdirect product of the monolithic primitive groups X_i associated to the minimal normal subgroups G_i, $1 \leq i \leq n$. By Proposition 23 and Proposition 21, for every i, $G/G_iC_G(G_i) \cong X_i/\mathrm{soc}(X_i)$ is either cyclic or non-solvable and all of its non-abelian composition factors are alternating groups of odd degree. Therefore either $G/\mathrm{soc}(G)$ is solvable (hence cyclic by Proposition 21) or non-solvable and all of its non-abelian composition factors are alternating groups of odd degree.

We are left to prove that if $G/\mathrm{soc}(G)$ is cyclic then $n = 1$. Assume by contradiction that $n \geq 2$.

Let u_i be the number of distinct prime divisors of the order of the cyclic groups $X_i/\mathrm{soc}(X_i)$ and assume that $u_1 \leq \cdots \leq u_n$.

Step 1. *Let m_i be the minimal index of a supplement of $\mathrm{soc}(X_i)$ in X_i; then $m_i \geq u_i$*

If $\mathrm{soc}(X_i) = S$ is a simple group, then X_i/S is isomorphic to a subgroup of $\mathrm{Out}(S)$, and thus $u_i \leq 2^{u_i} \leq |\,\mathrm{Out}(S)| \leq m_i$ (see e.g. Lemma 2.7 [4]).

If $\mathrm{soc}(X_i) = S^r$ where $r \neq 1$, then $X_i/\mathrm{soc}(X_i)$ is isomorphic to a subgroup Y of $\mathrm{Out}(S) \wr \mathrm{Sym}(r)$. Let K be the intersection of Y with the base subgroup $(\mathrm{Out}(S))^r$ of the wreath product

$\mathrm{Out}(S) \wr \mathrm{Sym}(r)$ and let a be the number of distinct prime divisors of $|K|$; since $|K|$ divides $|\,\mathrm{Out}(S)|^r$, we get that $2^a \leq |\,\mathrm{Out}(S)| \leq n_i$ where n_i is the minimal index of a subgroup of S. Now $b = u_i - a$ is smaller or equal than the number of distinct prime divisors of the order of Y/K which is isomorphic to a non trivial subgroup of $\mathrm{Sym}(r)$, hence $1 \leq b < r$ and thus $u_i = a + b \leq (2^a)^b \leq (2^a)^r \leq (n_i)^r \leq m_i$ whenever $a > 0$. If $a = 0$, then $X_i/\,\mathrm{soc}(X_i)$ is isomorphic to a subgroup of $\mathrm{Sym}(r)$ and thus $u_i < r \leq (n_i)^r \leq m_i$. This proves the first step.

Let π be the projection of G over $X = X_1$ and call $N = \mathrm{soc}\, X$. Note that there exist precisely u_1 maximal subgroups of the cyclic group X/N; let $H_1, \ldots, H_{u_1}$ be the maximal subgroups of G such that their images in X/N give all the maximal subgroups of X/N.

Let $\mathcal{M}$ be a set of $\sigma = \sigma(G)$ maximal subgroups whose union is G and define $\mathcal{A}$ to be the set of maximal subgroups of $\mathcal{M}$ containing G_1, $\mathcal{B} = \mathcal{M} \setminus \mathcal{A}$ and

$$\Omega = \left\{ \pi_1(g) \mid g \in G \setminus \bigcup_{M \in \mathcal{A}} M \right\}.$$

Step 2. *Assume that Ω contains a coset yN such that $\langle yN \rangle = X/N$.*

By Lemma 18, if Ω is covered by $\sigma_\Omega(X)$ maximal subgroups, then the same subgroups cover every coset $y^\alpha N$ with α prime to $|y|$. All the other elements of X are covered by the u_1 maximal subgroups $\pi(H_1), \ldots, \pi(H_{u_1})$, since the images of these elements are not generators of X/N. Then $\sigma(X) \leq \sigma_\Omega(X) + u_1$. On the other hand, by Proposition 16, $\sigma_\Omega(X) + \sum_{i \neq 1} \sigma^*(X_i) \leq \sigma(G) < \sigma(X)$, hence $\sum_{i \neq 1} \sigma^*(X_i) < u_1$. Remark 17 and Step 1 give that $\sum_{i \neq 1} u_i \leq \sum_{i \neq 1} m_i < u_1$, and this contradicts the minimality of u_1.

Step 3. *Assume that Ω does not contain a coset yN such that $\langle yN \rangle = X/N$.*

Then Ω is covered by the images in X of the subgroups $H_1, \ldots, H_{u_1}$ and thus, by definition of Ω, G is covered by the subgroups in $\mathcal{A}$ and $H_1, \ldots, H_{u_1}$. It follows that $|\mathcal{B}| + |\mathcal{A}| = \sigma(G) \leq u_1 + |\mathcal{A}|$, hence, by Step 1, $|\mathcal{B}| \leq u_1 \leq m_1$, against Lemma 3.2 in [14]. This final contradiction implies that G has to be a primitive monolithic group and proves the proposition. □

4 There is no group for which $\sigma(G) = 11$

In this section we will show that $\sigma(G)$ can never be equal to 11. The first trivial observation is that $\sigma(G) \neq 11$ whenever G is solvable, since in this case by Tomkinson's result $\sigma(G) = q + 1$, for a prime power q.

Assume by contradiction that there exists a primitive 11-sum group G. By Corollary 14, $\mathrm{soc}(G)$ is the direct product of n non G-equivalent minimal normal subgroups $G_1, \ldots, G_n$, where at most one of them is abelian.

Lemma 25. *Suppose that G is a primitive 11-sum group. Then G has no abelian minimal normal subgroups.*

Proof. Assume by contradiction that G_1 is abelian. By Corollary 14, G_1 is a complemented non-central factor of G, hence, by Corollary 6, $|G_1|+1 \leq \sigma(G) = 11$. Moreover, by Proposition 10, $11 = \sigma(G) < 2|G_1|$. Hence $|G_1|$ can only be $7, 2^3$ or 3^2. Actually, if $|G_1| = 7$, then the bound in Proposition 10 gives $\sigma(G) \leq 1+7$, against $\sigma(G) = 11$.

Note that, by Proposition 16, $\sigma(G) = 11 \geq \sum_{i=1}^{n} \sigma^*(X_i)$ where X_i are the monolithic groups associated to the G_i's; since G_1 is the only abelian subgroup and $\sigma^*(X_i) \geq 5$ if G_i is non-abelian, then G_1 is the unique minimal normal subgroup of G and $G \leq G_1 \rtimes \mathrm{Aut}(G_1)$.

If $|G_1| = 9$, then $G \leq \mathbb{F}_3^2 \rtimes \mathrm{GL}(2,3)$; hence G is solvable, a contradiction.

Thus $|G_1| = 8$ and $G = \mathbb{F}_2^3 \rtimes \mathrm{GL}(3,2)$, since every proper subgroup of $\mathrm{GL}(3,2)$ is solvable. Let $\mathcal{M} = \{M_1, \cdots, M_{11}\}$ be a set of 11 maximal subgroups covering G. In [6] it is proved that $\sigma(\mathrm{GL}(3,2)) = 15$ and, in particular, that one needs at least 7 subgroups to cover the seven point stabilizers of $\mathrm{GL}(3,2)$. It follows that all the 8 complement of G_1 in G occur in $\mathcal{M}$, let say they are $M_1, \ldots, M_8$. As in the proof of Proposition 10, for every point stabilizer $g \in \mathrm{GL}(3,2)$ there exists an element $v_g \in G_1$ such that gv_g does not belong to any complement of G_1 in G. Hence the remaining subgroups M_9, M_{10}, M_{11} of $\mathcal{M}$ have to cover all the elements gv_v where g is a point stabilizer. Since M_9, M_{10} and M_{11} contain G_1, this would imply that we can cover the seven point stabilizers of $\mathrm{GL}(3,2)$ with only three subgroups, a contradiction. □

Theorem 26. *There is no group G with $\sigma(G) = 11$.*

Proof. Suppose that G is a primitive 11-sum group and let $G_1, \ldots, G_n$ be its minimal normal subgroups. By the previous lemma every G_i is non-abelian. If $G_i = \mathrm{Alt}(5)$ for some i, then, by Corollary 20, $G = \mathrm{Alt}(5)$ or $\mathrm{Sym}(5)$. Otherwise, $\sigma^*(X_i) \geq l_{X_i}(G_i) > 5$ for every i and Proposition 16 implies that there is at most one minimal normal subgroup in G. By the same argument, if $G_1 = S^r$, where S is a simple non-abelian group, since $l_{X_1}(G_1) \geq 5^r$ and, by Lemma 5, $5^r+1 \leq \sigma(G) = 11$, we have that $G_1 = S$ and $l_{X_1}(G_1)+1 \leq 11$. Therefore G is an almost-simple group with socle S and $l_G(S) \leq 10$, in particular

$$S \in \{\mathrm{Alt}(n) \mid 5 \leq n \leq 10\} \cup \{\mathrm{Sym}(n) \mid 5 \leq n \leq 10\} \cup \{\mathrm{PSL}(2,q) \mid 7 \leq q \leq 8\}.$$

Thanks to the works of Maroti [12] and Bryce et al. [6], we can exclude most of these cases: indeed $\sigma(\mathrm{Alt}(n)) \geq 2^{n-2}$ if $n \neq 7, 9$, $\sigma(\mathrm{Alt}(5)) = 10$, $\sigma(\mathrm{Alt}(9)) \geq 80$, $\sigma(\mathrm{Sym}(n) = 2^{n-1}$ if n is odd and $n \neq 9$, $\sigma(\mathrm{Sym}(9)) \geq 172$, $\sigma(\mathrm{PSL}(2,7)) = 15$, $\sigma(\mathrm{PGL}(2,7)) = 29$, $\sigma(\mathrm{PSL}(2,8)) = 36$. Moreover, $\sigma(\mathrm{Aut}(\mathrm{Alt}(6))) \leq \sigma(C_2 \times C_2) = 3$ and $\sigma(\mathrm{Sym}(6)) = 13$ (see e.g. [1]). The remaining cases are $G = \mathrm{Alt}(7), \mathrm{Sym}(8), \mathrm{Sym}(10), \mathrm{M}_{10}, \mathrm{PGL}(2,9)$ and $\mathrm{Aut}(\mathrm{PSL}(2,8))$.

• $G \neq \mathrm{Alt}(7)$. Assume by contradiction $\sigma(\mathrm{Alt}(7)) = 11$. There are seven maximal subgroups of $\mathrm{Alt}(7)$ isomorphic to $\mathrm{Alt}(6)$; since $\sigma(\mathrm{Alt}(6)) = 16 > 11$, each of them has to appear in a minimal cover of G. Moreover, there are two conjugacy classes with 15 maximal subgroups isomorphic to

PSL(3, 2) and since $\sigma(\mathrm{PSL}(3,2)) = \sigma(\mathrm{PSL}(2,7)) = 15 > 11$ we have that $\sigma(\mathrm{Alt}(7))$ is at least $7 + 15 + 15$.

• $G \neq \mathrm{Sym}(8)$. If $\sigma(\mathrm{Sym}(8)) \leq 11$ then, since $\sigma(\mathrm{Sym}(7)) = 2^6$ and $\sigma(\mathrm{Alt}(8)) \geq 2^6$, arguing as in the previous case we get that a minimal cover $\mathcal{M}$ of Sym(8) contains the 8 point stabilizers and Alt(8). Let $g_1 = (1, 2, 3, 4, 5, 6, 7, 8)$, $g_2 = (1, 2, 3, 7, 4, 5, 6, 8)$ and $g_3 = (1, 2, 3, 5, 4, 6, 7, 8)$; any couple of them generate Sym(8) so that we need at least 3 more subgroups in $\mathcal{M}$, and thus $\sigma(\mathrm{Sym}(8)) > 11$.

• $G \neq \mathrm{Sym}(10)$. If $\sigma(\mathrm{Sym}(10)) \leq 11$, then, as $\sigma(\mathrm{Sym}(9)) = 2^8$ and $\sigma(\mathrm{Alt}(10)) \geq 2^8$, a minimal cover $\mathcal{M}$ of Sym(10) contains 10 point stabilizers and Alt(10). But these subgroups do not cover the 10-cycles. Thus $\sigma(\mathrm{Sym}(10)) > 11$.

• $G \neq \mathrm{M}_{10}$. In M_{10} there are 180 elements of order 8. The only maximal subgroups containing elements of order 8 are the Sylow 2-subgroups and each of them contains 4 of these elements; thus we need at least $180/4 = 45$ subgroups to cover the elements of order 8.

• $G \neq \mathrm{PGL}(2, 9)$. In PGL(2, 9) there are 144 elements of order 10. The only maximal subgroups containing elements of order 10 are the normalizers of the Sylow 5-subgroups and each of them contains 4 of these elements; thus we need at least $144/4 = 36$ subgroups to cover the elements of order 10.

• $G \neq \mathrm{Aut}(\mathrm{PSL}(2, 8))$. In $\mathrm{Aut}(\mathrm{PSL}(2, 8)) \setminus \mathrm{PSL}(2, 8)$ there are 336 elements of order 9. The only maximal subgroups containing elements of this kind are the normalizers of the Sylow 3-subgroups; each of them contains 12 of these elements thus we need at least $336/12 = 28$ subgroups to cover Aut(PSL(2, 8)). □

5 Direct products

Proposition 27. *Let $G = H_1 \times H_2$ be the direct product of two subgroups. Let N_i be the smallest normal subgroup of H_i such that H_i/N_i is a direct product of simple groups. If H_1/N_1 and H_2/N_2 have at most one non-abelian simple group S in common and the multiplicity of S in H_1/N_1 is at most one, then either $\sigma(G) = \min\{\sigma(H_1), \sigma(H_2)\}$, or the cyclic group C_p is an epimorphic image of both H_1 and H_2 and $\sigma(G) = p + 1$.*

Proof. Let G be a counterexample with minimal order. We first prove that G is a σ-primitive group. As $\Phi(G) = \Phi(H_1) \times \Phi(H_2)$, we have $\Phi(G) = 1$. Let N be a minimal normal subgroup of G and assume by contradiction that $\sigma(G) = \sigma(G/N)$. If $N \leq H_1$, then, by minimality of $|G|$, we have that either $\sigma(G/N) = \sigma(H_1/N \times H_2) = \min\{\sigma(H_1/N), \sigma(H_2)\} \geq \min\{\sigma(H_1), \sigma(H_2)\} \geq \sigma(G)$, and so $\sigma(G) = \min\{\sigma(H_1), \sigma(H_2)\}$, or C_p is a common factor of H_1/NN_1 and H_2/N_2, and $\sigma(G/N) = p+1$; in this case $\sigma(G) = \sigma(G/N) = p + 1$. Now assume that N is not contained in H_1 or H_2. Then N is a central minimal normal subgroup of G, $N = C_p \cong N_1N/N_1 \cong N_2N/N_2$ and G has a factor group isomorphic to $C_p \times C_p$; therefore $\sigma(G) \leq p + 1$. On the other hand, $\overline{N} = NH_2 \cap H_1 \cong N$ is a central minimal normal subgroup of G contained in H_1; by the previous case, $\sigma(G) < \sigma(G/\overline{N})$. Since $\delta_G(\overline{N}) \geq 2$, $\overline{N}$ has at least $|\overline{N}| = p$ complements; hence, by Lemma 5, $\sigma(G) \geq p+1$ and therefore $\sigma(G) = p+1$. Thus a counterexample G with minimal order is a σ-primitive group.

If G is solvable, then either $G \cong C_p^2$ and $\sigma(G) = p+1$ or G is monolithic: the second possibility cannot occur as G is the direct product of two non trivial normal subgroups. So from now on we may assume that G is non solvable, and in particular, by Proposition 21, that H_1/N_1 and H_2/N_2 have no common abelian factor.

Now observe that if M is a maximal subgroup of G and M does not contain H_1 and H_2, then G/M_G is a primitive group with nontrivial normal subgroups H_1M_G/M_G and H_2M_G/M_G. If $H_1M_G/M_G = H_2M_G/M_G$, then $G/M_G = H_1M_G/M_G = H_2M_G/M_G$ is a central factor of G/M_G and H_1/N_1 and H_2/N_2 have a common abelian factor, a contradiction. Thus $H_1M_G/M_G \neq H_2M_G/M_G$, and since $G/M_G = H_1M_G/M_G \times H_2M_G/M_G$ is a primitive group, H_1M_G/M_G and H_2M_G/M_G are isomorphic simple groups. Therefore, if H_1/N_1 and H_2/N_2 have no simple groups in common, then every maximal subgroup M of G contains either H_1 or H_2, and we obtain the result arguing as in Lemma 4 of [8].

So, we assume that H_1/N_1 and H_2/N_2 have precisely one non-abelian simple group S in common and the multiplicity of S in H_1/N_1 is one: let $K_i \geq N_i$ be the normal subgroups of H_i such that $H_1/K_1 = S$ and $H_2/K_2 = S^n$, being n the multiplicity of S in H_2/N_2 , and set $K = K_1 \times K_2$.

Let $\mathcal{M}$ be a minimal cover of G given by $\sigma(G)$ maximal subgroups of G. We set:

$$\begin{aligned}
\mathcal{M}_1 &= \{L \in \mathcal{M} \mid L \geq H_1\} = \{H_1 \times M \mid M \text{ a maximal subgroup of } H_2\},\\
\mathcal{M}_2 &= \{L \in \mathcal{M} \mid L \geq H_2\} = \{M \times H_2 \mid M \text{ a maximal subgroup of } H_1\},\\
\mathcal{M}_3 &= \mathcal{M} \setminus (\mathcal{M}_1 \cup \mathcal{M}_2).
\end{aligned}$$

Then we define the two sets

$$\Omega_1 = H_1 \setminus \bigcup_{M \times H_2 \in \mathcal{M}_2} M, \quad \Omega_2 = H_2 \setminus \bigcup_{H_1 \times M \in \mathcal{M}_1} M,$$

and their images under the projection π_{K_i} of H_i over H_i/K_i

$$\overline{\Omega}_i = \{\pi_{K_i}(w) \mid w \in \Omega_i\}.$$

As H_1/K_1 is not cyclic, we can cover $\overline{\Omega}_1$ with $|\overline{\Omega}_1|$ subgroups. Hence we can cover $H_1 = \{\bigcup_{M \times H_2 \in \mathcal{M}_2} M\} \cup \Omega_1$ with the images of the maximal subgroups in $\mathcal{M}_2$ plus $|\overline{\Omega}_1|$ maximal subgroups, and thus $\sigma(H_1) \leq |\mathcal{M}_2| + |\overline{\Omega}_1|$. On the other hand, $|\mathcal{M}_2| + |\mathcal{M}_3| \leq \sigma(G) < \sigma(H_1)$, and we obtain that

$$|\overline{\Omega}_1| > |\mathcal{M}_3|.$$

Now observe that the elements of the set $\Omega_1 \times \Omega_2$ can not belong to any of the subgroup of $\mathcal{M}_1$ or $\mathcal{M}_2$, thus the set $\Omega_1 \times \Omega_2$ has to be covered by the subgroups of $\mathcal{M}_3$. If $M \in \mathcal{M}_3$, then G/M_G is a primitive group and $G/M_G = H_1M_G/M_G \times H_2M_G/M_G = S \times S$; in particular $M \geq K$ and M/K is a maximal subgroup of diagonal type of G/K. This means that there exists an automorphism α of S and an index $i \in \{1, \dots n\}$, such that the set $(M/K) \cap (\overline{\Omega}_1 \times \overline{\Omega}_2)$ is given by elements of the type $(x, y_1, y_2, \dots, y_n)$ where $x \in \overline{\Omega}_1$, $(y_1, y_2, \dots, y_n) \in \overline{\Omega}_2$ and $y_i = x^\alpha$. For every $y \in S$ we denote by s_y the number of vectors $(y_1, y_2, \dots, y_n)$ such that $(y_1, y_2, \dots, y_n) \in \overline{\Omega}_2$ and $y_i = y$: note that

$$\sum_{y \in S} s_y = |\overline{\Omega}_2| = |\overline{\Omega}_1 \times \overline{\Omega}_2|/|\overline{\Omega}_1|.$$

On the other hand

$$|(M/K) \cap (\overline{\Omega}_1 \times \overline{\Omega}_2)| \leq \sum_{y \in S} s_y = |\overline{\Omega}_1 \times \overline{\Omega}_2|/|\overline{\Omega}_1| < |\overline{\Omega}_1 \times \overline{\Omega}_2|/|\mathcal{M}_3|,$$

since $|\overline{\Omega}_1| > |\mathcal{M}_3|$. This implies that we can not cover $\Omega_1 \times \Omega_2$ with the $|\mathcal{M}_3|$ subgroups of $\mathcal{M}_3$, a contradiction. □

Theorem 28. *Let $G = H_1 \times H_2$ be the direct product of two subgroups. If no alternating group* $\mathrm{Alt}(n)$ *with n odd is a homomorphic image of both H_1 and H_2, then either $\sigma(G) = \min\{\sigma(H_1), \sigma(H_2)\}$ or $\sigma(G) = p + 1$ and $S = C_p$ is a homomorphic image of both H_1 and H_2.*

Proof. Let G be a counterexample with minimal order. Let N_i be the minimal normal subgroup of H_i such that H_i/N_i is a direct product of simple groups. As in the proof of Proposition 27, it is easy to see that G is a σ-primitive group, H_1/N_1 and H_2/N_2 have at least one simple group S in common and S is non-abelian.

By Corollary 14, G has at most one abelian minimal normal subgroup, so we can assume that every minimal normal subgroup of H_1 is non-abelian.

Let K be a normal subgroup of G with $G/K \cong S$. Note that $\delta_G(G/K) \geq 2$, indeed $\delta_G(G/K)$ coincides with the multiplicity of S in $G/(N_1 \times N_2)$. Hence, by Corollary 14 (3), no minimal normal subgroup of G is G-equivalent to G/K. This implies in particular that S is an epimorphic image of $H_1/\operatorname{soc}(H_1)$, and consequently S is an homomorphic image of X/N where X is a monolithic primitive group associated to a minimal normal subgroup N of H_1. By the remark above N is non-abelian, so $N = T^r$ with T a non-abelian simple group. Since X is a subgroup of $\mathrm{Aut}(T) \wr \mathrm{Sym}(r)$ and S is non-abelian, S is an homomorphic image of a transitive group Y of degree r. Then Y satisfies the assumption of Lemma 22 and, since S is not an alternating group of odd degree, we get $\sigma(Y) \leq 4^r$. Since, by Corollary 6, $5^r + 1 \leq \sigma(G) \leq \sigma(Y)$, we get a contradiction. □

Received: July 2008. Revised: August 2008.

References

[1] A. Abdollahi, F. Ashraf and S.M. Shaker, *The symmetric group of degree six can be covered by 13 and no fewer proper subgroups*, Bull. Malays. Math. Sci. Soc., **30**(2) (2007), no. 1, 57–58.

[2] M. Aschbacher, *Finite group theory*, Second edition, Cambridge Studies in Advanced Mathematics 10, Cambridge University Press, Cambridge, 2000.

[3] M. Aschbacher and R. Guralnick, *Some applications of the first cohomology group*, J. Algebra, **90**(1984), 446–460.

[4] M. Aschbacher and R. Guralnick, *On abelian quotients of primitive groups*, Proc. Amer. Math. Soc., **107**(1989), 89–95.

[5] R.A. Bryce and L. Serena, *A note on minimal coverings of groups by subgroups*, Special issue on group theory, J. Aust. Math. Soc., **71**(2001), no. 2, 159–168

[6] R.A. Bryce, V. Fedri and L. Serena, *Subgroup coverings of some linear groups*, Bull. Austral. Math. Soc., **60**(1999), 227–238.

[7] P.J. Cameron, *Permutation groups*, London Mathematical Society Student Texts 45. Cambridge University Press, 1999.

[8] J.H.E. Cohn, *On n-sum groups*, Math. Scand., **75**(1) (1994), 44–58.

[9] E. Detomi and A. Lucchini, *Crowns and factorization of the probabilistic zeta function of a finite group*, J. Algebra, **265**(2003), 651–668.

[10] P. Jiménez-Seral and J. Lafuente, *On complemented nonabelian chief factors of a finite group*, Israel J. Math., **106**(1998), 177–188.

[11] P. Kleidman and M. Liebeck, *The subgroup structure of the finite classical groups*, London Mathematical Society Lecture Note Series, 129 Cambridge University Press, 1990.

[12] A. Maróti, *Covering the symmetric groups with proper subgroups*, J. Combin. Theory Ser. A, **110**(1) (2005), 97–111.

[13] C.E. Praeger and J. Saxl, *On the orders of primitive permutation groups*, Bull. London Math. Soc., **12**(1980), 303–307.

[14] M.J. Tomkinson, *Groups as the union of proper subgroups*, Math. Scand., **81**(2) (1997), 191–198.

[15] Y.M. Wang, *Finite groups admitting a fixed-point-free automorphism group*, Northeast. Math. J., **9**(4) (1993), 516–520.

CUBO A Mathematical Journal
Vol.10, N^o 03, (211–222). October 2008

On Two-Sided Centralizers of Rings and Algebras*

Joso Vukman
Department of Mathematics and Computer Sciences,
Faculty of Natural Sciences and Mathematics, University of Maribor,
Koroška 160, 2000 Maribor, Slovenia
email: joso.vukman@uni-mb.si

and

Irena Kosi-Ulbl
Faculty of Education, University of Maribor,
Koroška 160, 2000 Maribor, Slovenia
email: irena.kosi@uni-mb.si

ABSTRACT

In this paper we prove the following result. Let A be a semisimple H^*-algebra and let $T : A \to A$ be an additive mapping satisfying the relation $(n+1)T(x^{nm+1}) = T(x)x^{nm} + x^m T(x)x^{(n-1)m} + \cdots + x^{nm}T(x)$, for all $x \in A$ and some fixed integers $m \geq 1$, $n \geq 1$. In this case T is a two-sided centralizer.

RESUMEN

En este artículo probamos el siguiente resultado. Sea A una H^*-algebra semi-simple y $T : A \to A$ una aplicación aditiva satisfaziendo la relación $(n+1)T(x^{nm+1}) =$

*This research has been supported by the Research Council of Slovenia.

$T(x)x^{nm}+x^mT(x)x^{(n-1)m}+\cdots+x^{nm}T(x)$, para todo $x\in A$ y ciertos $m\geq 1$, y $n\geq 1$ enteros fixados. En este caso T es un centralizador "two-sided".

Key words and phrases: *Prime ring, semiprime ring, Banach space, standard operator algebra, H^*-algebra, left (right) centralizer, left (right) Jordan centralizer, two-sided centralizer.*

Math. Subj. Class.: *16W10, 46K15, 39B05.*

Introduction

Throughout, R will represent an associative ring with center $Z(R)$. Given an integer $n\geq 2$, a ring R is said to be $n-$torsion free, if for $x\in R$, $nx=0$ implies $x=0$. As usual the commutator $xy-yx$ will be denoted by $[x,y]$.Let us recall that a ring R is prime if for $a,b\in R$, $aRb=(0)$ implies that either $a=0$ or $b=0$, and is semiprime in case $aRa=(0)$ implies $a=0$. An additive mapping $x\longmapsto x^*$ on a ring R is called involution in case $(xy)^*=y^*x^*$ and $x^{**}=x$ hold for all pairs $x,y\in R$. A ring equipped with an involution is called a ring with involution or $^*-$ring. We denote by Q_r and C Martindale right ring of quotients and extended centroid of a semiprime ring R. For the explanation of Q_r and C we refer to [3]. An additive mapping $T:R\to R$, where R is an arbitrary ring, is called a left centralizer in case $T(xy)=T(x)y$ holds for all pairs $x,y\in R$. The concept appears naturally in C^*-algebras. In ring theory it is more common to work with module homomorphisms. Ring theorists would write $T:R_R\to R_R$ of a right ring module R into itself. For a semiprime ring R all such homomorphisms are of the form $T(x)=qx$, for all $x\in R$, where q is some fixed element of Q_r (see Chapter 2 in [3]). In case R has the identity element $T:R\to R$ is a left centralizer iff T is of the form $T(x)=ax$, for all $x\in R$, where a is some fixed element of R.An additive mapping $T:R\to R$ is called a left Jordan centralizer in case $T(x^2)=T(x)x$ holds for all $x\in R$.The definitions of right centralizer and right Jordan centralizer are self-explanatory. We call $T:R\to R$ a two-sided centralizer in case T is both a left and a right centralizer. In case $T:R\to R$ is a two-sided centralizer, where R is a semiprime ring with extended centroid C, then there exists an element $\lambda\in C$ such that $T(x)=\lambda x$, for all $x\in R$ (see Theorem 2.3.2 in [3]). One of the initial papers using the concept of centralizers (also called multipliers) is due to Wendel [33] for group algebras. Helgason [9] introduced centralizers for Banach algebras. Wang [32] studied centralizers of commutative Banach algebras. Johnson [11] introduced the concept of centralizers for rings. We refer to Busby [7] for a study of so-called double centralizers in the extension of C^*-algebras. Akemann, Pedersen and Tomiyama [1] have studied centralizers of C^*-algebras. Several authors have also studied spectral properties of centralizers on Banach algebras (see [15, 16]). Johnson [12] has studied centralizers on some topological algebras. Johnson [13] has studied the continuity of centralizers on Banach algebras (see also [11]). Husain [10] has also investigated centralizers on topological algebras with particular reference to complete metrizable locally convex algebras and topological algebras with orthogonal bases. Khan, Mohammad and Thaheem [14] have studied centralizers and double centralizers on certain topological algebras. Centralizers have also appeared in a variety, among which we mention representation theory of Banach algebras, the study of Banach modules, Hopf algebras (see [18, 19]), the theory of singular integrals, interpolation theory, stohastic processes, the theory of semigroups of operators, partial differential equations and the study of approximation problems (see Larsen [16] for more details). Zalar [34] has proved that any left (right) Jordan centralizer on a 2-torsion free semiprime ring is a left (right) centralizer. Molnár [17] has proved that in case we have an additive mapping $T : A \to A$, where A is a semisimple H^*-algebra, satisfying the relation $T(x^3) = T(x)x^2$ $(T(x^3) = x^2T(x))$ for all $x \in A$, then T is a left (right) centralizer. Let us recall that a semisimple H^*-algebra is a semisimple Banach *-algebra A whose norm is a Hilbert space norm such that $(x, yz^*) = (xz, y) = (z, x^*y)$ is fulfilled for all $x, y, z \in A$ (see [2]). Benkovič and Eremita [4] have proved that in case there exists an additive mapping $T : R \to R$,where R is a prime ring with suitable characteristic restrictions, satisfying the relation $T(x^n) = T(x)x^{n-1}$, for all $x \in R$ and some fixed integer $n > 1$, then T is a left centralizer. Vukman and Kosi-Ulbl [26] have proved that any additive mapping T , which maps a semisimple H^*-algebra A into itself and satisfies the relation $2T(x^{n+1}) = T(x)x^n + x^nT(x)$, for all $x \in A$ and some fixed integer $n \geq 1$, is a two-sided centralizer (see also [5]). A result of Vukman and Kosi-Ulbl [27] states that in case there exists an additive mapping $T : R \to R$, where R is a $2-$torsion free semiprime $^*-$ring, satisfying the relation $T(xx^*) = T(x)x^*$ $(T(x^*x) = x^*T(x))$, for all $x \in R$,then T is a left (right) centralizer. For results concerning centralizers on prime and semiprime rings, operator algebras and H^*-algebras we refer to $[8, 20-31]$. Let X be a real or complex Banach space and let $L(X)$ and $F(X)$ denote the algebra of all bounded linear operators on X and the ideal of all finite rank operators in $L(X)$, respectively. An algebra $A(X) \subset L(X)$ is said to be standard in case $F(X) \subset A(X)$. Let us point out that any standard algebra is prime, which is a consequence of Hahn-Banach theorem. We denote by X^* the dual space of a Banach space X and by I the identity operator on X.

Vukman [20] has proved the following result.

THEOREM A. Let R be a 2-torsion free semiprime ring and let $T : R \to R$ be an additive mapping. Suppose that

$$2T(x^2) = T(x)x + xT(x)$$

holds for all $x \in R$. In this case T is a two-sided centralizer.

Vukman and Kosi-Ulbl [23] have proved the result below.

THEOREM B. Let R be a 2-torsion free semiprime ring and let $T : R \to R$ be an additive mapping. Suppose that

$$3T(xyx) = T(x)yx + xT(y)x + xyT(x)$$

holds for all pairs $x, y \in R$. In this case T is of the form $T(x) = \lambda x$, for all $x \in R$ and some fixed element λ from the extended centroid C of R.

Motivated by Theorem A and Theorem B Fošner and Vukman [8] have proved the following theorem.

THEOREM C. Let R be a prime ring and let $T : R \to R$ be an additive mapping satisfying the relation

$$nT(x^{n+1}) = T(x)x^{n-1} + xT(x)x^{n-2} + ... + x^{n-1}T(x),$$

for all $x \in R$, where $n \geq 2$ is some fixed integer. If $char(R) = 0$, then T is of the form $T(x) = \lambda x$, for all $x \in R$ and some fixed element λ from the extended centroid C of R.

In the proof of Theorem C Fošner and Vukman used as the main tool the theory of functional identities (Beidar-Brešar-Chebotar theory). The theory of functional identities considers set-theoretic maps on rings that satisfy some identical relations. When threatening such relations one usually concludes that the form of the maps involved can be described, unless the ring is very special (see[6]).

It this paper we consider the following more general relation

$$(n+1)T(x^{nm+1}) = T(x)x^{nm} + x^mT(x)x^{(n-1)m} + ... + x^{nm}T(x), \tag{1}$$

where $m \geq 1$, $n \geq 1$ are some fixed integers. One can notice that the expression (1) for $n = m = 1$ is the same as hypothesis of Theorem A. Obviously, any two-sided centralizer on arbitrary ring satisfies the above relation. We proceed with the following conjecture.

CONJECTURE. Let R be a semiprime ring with suitable torsion restrictions and let $T : R \to R$ be an additive mapping satisfying the relation (1) for all $x \in R$ and some fixed integers $m \geq 1$, $n \geq 1$. In this case T is a two-sided centralizer.

It is our aim in this paper to prove the above conjecture in semisimple H^*-algebras and in semiprime rings with the identity element. Our methods differ from those used in [8].

THEOREM 1. Let A be a semisimple H^*-algebra. Suppose $T : A \to A$ is an additive mapping satisfying the relation (1) for all $x \in A$ and some fixed integers $m \geq 1$, $n \geq 1$. In this case T is a two-sided centralizer.

For the proof of Theorem 1 we need the theorem below which is of independent interest.

THEOREM 2. Let X be a Banach space over the real or complex field $\mathcal{F}$, let $A(X) \subset L(X)$ be a standard operator algebra. Suppose $T : A(X) \to L(X)$ is an additive mapping satisfying the relation

$$(n+1)T(A^{nm+1}) = T(A)A^{nm} + A^mT(A)A^{(n-1)m} + ... + A^{nm}T(A),$$

for all $A \in A(X)$ and some fixed integers $m \geq 1$, $n \geq 1$. In this case T is of the form $T(A) = \lambda A$, for all $A \in A(X)$ and some fixed $\lambda \in \mathcal{F}$.In particular, T is continuous.

Proof. We have the relation

$$(n+1)T(A^{nm+1}) = T(A)A^{nm} + A^m T(A) A^{(n-1)m} + ... + A^{nm}T(A). \tag{2}$$

Let us first consider the restriction of T on $F(X)$. Let A be from $F(X)$ and let $P \in F(X)$, be a projection with $AP = PA = A$. From the above relation one obtains $T(P) = PT(P)P$, which gives

$$T(P)P = PT(P). \tag{3}$$

Putting $A + P$ for A in the relation (2), we obtain

$$\begin{gathered}(n+1)\sum_{i=0}^{nm+1}\binom{nm+1}{i}T\left(A^{nm+1-i}P^i\right) = \left(T\left(A\right)+B\right)\left(\sum_{i=0}^{nm}\binom{nm}{i}A^{nm-i}P^i\right) + \\ \left(\sum_{i=0}^{m}\binom{m}{i}A^{m-i}P^i\right)\left(T\left(A\right)+B\right)\left(\sum_{i=0}^{(n-1)m}\binom{(n-1)\,m}{i}A^{(n-1)m-i}P^i\right) + ... + \\ \left(\sum_{i=0}^{nm}\binom{nm}{i}A^{nm-i}P^i\right)\left(T\left(A\right)+B\right),\end{gathered} \tag{4}$$

where B stands for $T\left(P\right)$. Using (2) and rearranging the equation (4) in sense of collecting together terms involving equal number of factors of P we obtain

$$\sum_{i=1}^{nm} f_i\left(A, P\right) = 0, \tag{5}$$

where $f_i\left(A,P\right)$ stands for the expression of terms involving i factors of P. Replacing A by $A + 2P$, $A + 3P$, $\ldots$, $A + nmP$ in turn in the equation (1), and expressing the resulting system of nm homogeneous equations of variables $f_i\left(A,P\right)$, $i = 1, 2, ..., nm$, we see that the coefficient matrix of the system is a van der Monde matrix

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ 2 & 2^2 & \cdots & 2^{nm} \\ \vdots & \vdots & \vdots & \vdots \\ nm & (nm)^2 & \cdots & (nm)^{nm} \end{bmatrix}.$$

Since the determinant of the matrix is different from zero, it follows that the system has only the trivial solution.

In particular,

$$f_{nm-1}(A,P) = (n+1)\binom{nm+1}{nm-1}T\left(A^2\right) - \binom{nm}{nm-1}T(A)A - \binom{nm}{nm-2}BA^2 -$$
$$\binom{m}{m-2}\binom{(n-1)m}{(n-1)m}A^2B - \binom{m}{m-1}\binom{(n-1)m}{(n-1)m}AT(A)P - \binom{m}{m-1}\binom{(n-1)m}{(n-1)m-1}ABA -$$
$$\binom{m}{m}\binom{(n-1)m}{(n-1)m-1}PT(A)A - \binom{m}{m}\binom{(n-1)m}{(n-1)m-2}BA^2 - \cdots -$$
$$\binom{m}{m}\binom{(n-1)m}{(n-1)m-2}A^2B - \binom{m}{m}\binom{(n-1)m}{(n-1)m-1}AT(A)P - \binom{m}{m-1}\binom{(n-1)m}{(n-1)m-1}ABA -$$
$$\binom{m}{m-1}\binom{(n-1)m}{(n-1)m}PT(A)A - \binom{m}{m-2}\binom{(n-1)m}{(n-1)m}BA^2 -$$
$$\binom{nm}{nm-2}A^2B - \binom{nm}{nm-1}AT(A) = 0$$

and

$$f_{nm}(A,P) = (n+1)\binom{nm+1}{nm}T(A) - \binom{nm}{nm}T(A)P - \binom{nm}{nm-1}BA -$$
$$\binom{m}{m-1}\binom{(n-1)m}{(n-1)m}AB - \binom{m}{m}\binom{(n-1)m}{(n-1)m}PT(A)P - \binom{m}{m}\binom{(n-1)m}{(n-1)m-1}BA - \cdots -$$
$$\binom{m}{m}\binom{(n-1)m}{(n-1)m-1}AB - \binom{m}{m}\binom{(n-1)m}{(n-1)m}PT(A)P - \binom{m}{m-1}\binom{(n-1)m}{(n-1)m}BA -$$
$$\binom{nm}{nm-1}AB - \binom{nm}{nm}PT(A) = 0.$$

The above equations reduce to

$$6(n+1)(nm+1)T\left(A^2\right) = 12(T(A)A + AT(A)) + 6(n-1)(AT(A)P + PT(A)A) +$$
$$(n+1)((2n+1)m-3)\left(A^2B + BA^2\right) + 2m(n-1)(n+1)ABA, \tag{6}$$

and

$$2(n+1)(nm+1)T(A) = 2(T(A)P + PT(A)) +$$
$$n(n+1)m(AB+BA) + 2(n-1)PT(A)P. \tag{7}$$

Right multiplication of the relation (7) by P gives

$$2(n+1)(nm+1)T(A)P = 2(T(A)P + PT(A)) +$$
$$n(n+1)m(AB+BA) + 2(n-1)PT(A)P. \tag{8}$$

Similarly one obtains

$$2(n+1)(nm+1)PT(A) = 2(T(A)P + PT(A)) +$$

$$n\,(n+1)\,m\,(AB+BA)+2\,(n-1)\,PT\,(A)\,P. \tag{9}$$

Combining (8) with (9) we arrive at

$$T(A)P = PT(A),$$

which reduces the relation (6) to

$$6\,(mn+1)\,T\left(A^2\right) = 6\,(T(A)A + AT(A)) +$$

$$((2n+1)\,m-3)\left(A^2B+BA^2\right)+2m\,(n-1)\,ABA, \tag{10}$$

and the relation (7) to

$$2\,(mn+1)\,T\,(A) = 2T\,(A)\,P + mn\,(AB+BA)\,. \tag{11}$$

Right multiplication of the above relation by P and combining the relation so obtained with (11) gives

$$T(A) = T(A)P.$$

According to the above relation the relation (11) reduces to

$$2T\,(A) = AB + BA. \tag{12}$$

From the above relation one obtains

$$2T\left(A^2\right) = A^2B + BA^2. \tag{13}$$

Right and then left multiplication of the relation (12) by A gives

$$2T\,(A)\,A = ABA + BA^2 \tag{14}$$

and

$$2AT\,(A) = A^2B + ABA, \tag{15}$$

respectively. Using the relations (13), (14) and (15) in the relation (10) gives after some calculation

$$A\,(m,n)\,BA^2 + A\,(m,n)\,A^2B - 2A\,(m,n)\,ABA = 0,$$

where $A\,(m,n)$ stands for $mn-m+3$. The above relation reduces to

$$A^2B + BA^2 - 2ABA = 0. \tag{16}$$

Applying the relations (13) and (16) in the relation (10) one obtains

$$2T\left(A^2\right) = T(A)A + AT(A). \tag{17}$$

From the relation (12) one can conclude that T maps $F(X)$ into itself. We have therefore an additive mapping $T : F(X) \to F(X)$ satisfying the relation (17) for all $A \in F(X)$. Since $F(X)$

is prime one can apply Theorem A and conclude that T is a two-sided centralizer of $F(X)$. We intend to prove that there exists an operator $C \in L(X)$, such that

$$T(A) = CA, \quad \text{for all} \quad A \in F(X). \tag{18}$$

For any fixed $x \in X$ and $f \in X^*$ we denote by $x \otimes f$ an operator from $F(X)$ defined by $(x \otimes f)y = f(y)x$, for all $y \in X$. For any $A \in L(X)$ we have $A(x \otimes f) = ((Ax) \otimes f)$.Let us choose f and y such that $f(y) = 1$ and define $Cx = T(x \otimes f)y$. Obviously, C is linear. Using the fact that T is left centralizer on $F(X)$ we obtain

$$(CA)x = C(Ax) = T((Ax) \otimes f)y = T(A(x \otimes f))y = T(A)(x \otimes f)y = T(A)x, \quad x \in X.$$

We have therefore $T(A) = CA$ for any $A \in F(X)$. Since T right centralizer on $F(X)$ we obtain $C(AP) = T(AP) = AT(P) = ACP$, where $A \in F(X)$ and P is arbitrary one-dimensional projection. We have therefore $[A, C]P = 0$. Since P is arbitrary one-dimensional projection it follows that $[A, C] = 0$ for any $A \in F(X)$. Using closed graph theorem one can easily prove that C is continuous. Since C commutes with all operators from $F(X)$ one can conclude that $Cx = \lambda x$ holds for any $x \in X$ and some $\lambda \in \mathcal{F}$, which gives together with the relation (17) that T is of the form

$$T(A) = \lambda A \tag{19}$$

for any $A \in F(X)$ and some $\lambda \in F$. It remains to prove that the above relation holds on $A(X)$ as well. Let us introduce $T_1 : A(X) \to L(X)$ by $T_1(A) = \lambda A$ and consider $T_0 = T - T_1$. The mapping T_0 is, obviously, additive and satisfies the relation (2). Besides, T_0 vanishes on $F(X)$. It is our aim to prove that T_0 vanishes on $A(X)$ as well. Let $A \in A(X)$, let P be an one-dimensional projection and $S = A + PAP - (AP + PA)$. Note that S can be written in the form $S = (I - P)A(I - P)$, where I denotes the identity operator on X, Since, obviously, $S - A \in F(X)$, we have $T_0(S) = T_0(A)$. Besides, $SP = PS = 0$. We have therefore the relation

$$(n+1)T_0(A^{nm+1}) = T_0(A)A^{nm} + A^m T_0(A)A^{(n-1)m} + ... + A^{nm}T_0(A), \tag{20}$$

for all $A \in A(X)$. Applying the above relation we obtain

$$\begin{gathered}
T_0(S)S^{nm} + S^m T_0(S)S^{(n-1)m} + ... + S^{nm}T_0(S) = (n+1)T_0(S^{nm+1}) = \\
(n+1)T_0(S^{nm} + P) = (n+1)T_0((S+P)^{nm+1}) = \\
T_0(S+P)(S+P)^{nm} + (S+P)^m T_0(S+P)(S+P)^{(n-1)m} + ...+ \\
(S+P)^{(n-1)m}T_0(S)(S+P)^m + (S+P)^{nm}T_0(S+P) = T_0(S)S^{nm} + \\
S^m T_0(S)S^{(n-1)m} + ... + S^{nm}T_0(S) + T_0(S)P + S^m T_0(S)P + \\
PT_0(S)S^{(n-1)m} + ... + S^{(n-1)m}T_0(S)P + \\
PT_0(S)S^m + PT_0(S) + (n-1)PT_0(S)P.
\end{gathered}$$

We have therefore

$$T_0(A)P + S^m T_0(A)P + PT_0(A)S^{(n-1)m} + ... + S^{(n-1)m}T_0(A)P+$$
$$PT_0(A)S^m + PT_0(A) + (n-1)PT_0(A)P = 0. \tag{21}$$

Multiplying the above relation from both sides by P we obtain

$$PT_0(A)P = 0. \tag{22}$$

Now right multiplication of the relation (21) by P gives because of (22)

$$T_0(A)P + S^m T_0(A)P + ... + S^{(n-1)m}T_0(A)P = 0. \tag{23}$$

Replacing A by $2A$, $3A$, $\ldots$, nA in turn in the equation (23), and expressing the resulting system of n homogeneous equations of variables $T_0(A)P$, $S^{im}T_0(A)P$, $i = 1, 2, ..., n-1$, we see that the coefficient matrix of the system is a matrix of the form

$$\begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2^m & \cdots & 2^{(n-1)m} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & n^m & \cdots & n^{(n-1)m} \end{bmatrix}.$$

Since the determinant of the matrix is different from zero, it follows that the system has only the trivial solution. We have therefore $T_0(A)P = 0$. Since P is an arbitrary one-dimensional projection, one can conclude that $T_0(A) = 0$, for any $A \in A(X)$, which completes the proof of the theorem.

It should be mentioned that in the proof of Theorem 2 we used some ideas similar to those used by Molnár in [17]. Let us point out that in Theorem 2 we obtain as a result the continuity of T under purely algebraic assumptions concerning T, which means that Theorem 2 might be of some interest from the automatic continuity point of view.

Proof of Theorem 1. The proof goes through using the same arguments as in the proof of Theorem in [17] with the exception that one has to use Theorem 2 instead of Lemma in [17].

We are ready for our last result.

THEOREM 3. Let $n \geq 1$, $m \geq 1$ be integers and let R be a 2, m, n, $n+1$ and $((n-1)\,m+3)-$torsion free semiprime ring with the identity element. Suppose that we have an additive mapping $T : R \to R$ satisfying the relation (1) for all $x \in R$. In this case T is of the form $T(x) = ax$, for all $x \in R$ and some fixed element $a \in Z(R)$.

Proof. We have the relation (1). Using similar approach as in the proof of Theorem 2, with the exception that we use the identity element e instead of a projection, we obtain from the above relation

$$6(nm+1)T\left(x^2\right) = 6\left(T(x)x + xT(x)\right) + \\ \left(\left(2n+1\right)m-3\right)\left(x^2a+ax^2\right) + 2\left(n-1\right)mxax, \quad x \in R \tag{24}$$

and

$$2T\left(x\right) = xa + ax, \quad x \in R, \tag{25}$$

where a stands for $T\left(e\right)$. In the procedure mentioned above we used the fact that R is m, n and $n+1$-torsion free.

The substitution x^2 for x in (25)gives

$$2T\left(x^2\right) = x^2a + ax^2, \quad x \in R. \tag{26}$$

Multiplying the relation (25) first from the right side then from the left side by x we obtain

$$2T\left(x\right)x = xax + ax^2, \quad x \in R \tag{27}$$

and

$$2xT\left(x\right) = x^2a + xax, \quad x \in R. \tag{28}$$

Using (26), (27) and (28) in the relation (24)and applying the fact that R is $(n-1)m+3-$torsion free we obtain after some calculation

$$x^2a + ax^2 - 2xax = 0, \quad x \in R,$$

which can be written in the form

$$\left[\left[a,x\right],x\right] = 0, \quad x \in R. \tag{29}$$

Putting $x+y$ for x in the above relation we obtain

$$\left[\left[a,x\right],y\right] + \left[\left[a,y\right],x\right] = 0, \quad x,y \in R. \tag{30}$$

The substitution xy for y in relation (30) gives because of (29) and (30)

$$0 = \left[\left[a,x\right],xy\right] + \left[\left[a,xy\right],x\right] = \\ = \left[\left[a,x\right],x\right]y + x\left[\left[a,x\right],y\right] + \left[\left[a,x\right]y + x\left[a,y\right],x\right] = \\ = x\left[\left[a,x\right],y\right] + \left[\left[a,x\right],x\right]y + \left[a,x\right]\left[y,x\right] + x\left[\left[a,y\right],x\right] = \left[a,x\right]\left[y,x\right], \quad x,y \in R.$$

Thus we have

$$\left[a,x\right]\left[y,x\right] = 0, \quad x,y \in R.$$

The substitution ya for y in the above relation gives $[a, x]\, y\, [a, x] = 0$, for all pairs $x, y \in R$. Let us point out that so far we have not used the assumption that R is semiprime. Since R is semiprime, it follows from the last relation that $[a, x] = 0$, for all $x \in R$. In other words, $a \in Z\,(R)$, which reduces the relation (25) to $T\,(x) = ax$, $x \in R$, since R is $2-$torsion free. The proof of the theorem is complete.

Received: May 2008. Revised: August 2008.

References

[1] C.A. Akemann, G.K. Pedersen and J. Tomiyama, *Multipliers of C^*-algebras*, J. Funct. Anal., **13**(1973), 277–301.

[2] W. Ambrose, *Structure theorems for a special class of Banach algebras*, Trans. Amer. Math. Soc., **57**(1945), 364–386.

[3] K.I. Beidar, W.S. Martindale III and A.V. Mikhalev, *Rings with generalized identities*, Marcel Dekker, Inc., New York, 1996.

[4] D. Benkovič and D. Eremita, *Characterizing left centralizers by their action on a polynomial*, Publ. Math. (Debr.), **64**(2004), 343–351.

[5] D. Benkovič, D. Eremita and J. Vukman, *A characterization of the centroid of a prime rings*, Studia Sci. Math. Hung., to appear.

[6] M. Brešar, M. Chebotar and W.S. Martindale 3rd, *Functional identities*, Birkhäuser Verlag, Basel, Boston, Berlin, 2007.

[7] R.C. Busby, *Double centralizers and extension of C^*-algebras*, Trans. Amer. Math. Soc., **132**(1968), 79–99.

[8] M. Fošner and J. Vukman, *An equation related to two-sided centralizers in prime rings*, Houston J. Math., to appear.

[9] S. Helgason, *Multipliers of Banach algebras*, Ann. of Math., **64**(1956), 240–254.

[10] T. Husain, *Multipliers of topological algebras*, Dissertations Math. (Rozprawy Mat.), **285**(1989), 40pp.

[11] B.E. Johnson, *An introduction to the theory of centralizers*, Proc. London Math. Soc., **14**(1964), 299–320.

[12] B.E. Johnson, *Centralizers on certain topological algebras*, J. London Math. Soc., **39**(1964), 603–614.

[13] B.E. Johnson, *Continuity of centralizers on Banach algebras*, J. London Math. Soc., **41**(1966), 639–640.

[14] L.A. Khan, N. Mohammad and A.B. Thaheem, *Double multipliers on topological algebras*, Internat. J. Math.& Math, Sci., **22**(1999), 629–636.

[15] R. Larsen, *An introduction to the Theory of Multipliers*, Springer-Verlag, Berlin, 1971.

[16] K.B. Larsen, *Mulptipliers and local spectral theory, Functional Analysis and Operator Theory*, Banach Center Publications, Institute of Mathematics, Polish Academy of Sciences, Warszawa, 223–236, 1994.

[17] L. Molnár, *On centralizers of an H^*-algebra*, Publ. Math. Debrecen, **46**(1995), 1–2, 89–95.

[18] A. Van Daele, *Multiplier Hopf algebras*, Trans. Amer. Math. Soc., **342**(1994), 917–932.

[19] A. Van Daele and Y. Zhang, *A Survey on multiplier Hopf algebras*, Hopf algebras and quantum groups (Brussels, 1998), 269–306, Lecture Notes in Pure and Appl. Mat, **209**, Dekker, New York, 2000.

[20] J. Vukman, *An identity related to centralizers in semiprime rings*, Comment. Math. Univ. Carol., **40**(3) (1999), 447–456.

[21] J. Vukman, *Centralizers of semiprime rings*, Comment. Math. Univ. Carol., **42**(2) (2001), 237–245.

[22] J. Vukman and I. Kosi Ulbl, *On centralizers of semiprime rings*, Aequationes Math., **66**(2003), 277–283.

[23] J. Vukman and I. Kosi-Ulbl, *An equation related to centralizers in semiprime rings*, Glasnik Mat. Vol., **38**(2003), 253–261.

[24] J. Vukman and I. Kosi-Ulbl, *On certain equations satisfied by centralizers in rings*, Intern. Math. Journal., **5**(2004), 437–456.

[25] J. Vukman, *An equation on operator algebras and semisimple H^*-algebras*, Glasnik Mat., **40**(60) (2005), 201–206.

[26] J. Vukman and I. Kosi-Ulbl, *Centralizers on rings and algebras*, Bull. Austral. Math. Soc., Vol. 71 (2005), 225–234.

[27] J. Vukman and I. Kosi-Ulbl, *On centralizers of semiprime rings with involution*, Studia Sci. Math. Hungar., **43**(1) (2006), 77–83.

[28] J. Vukman and I. Kosi-Ulbl, *A remark on a paper of L. Molnár*, Publ. Math. Debrecen, **67**(3-4) (2005), 419–427.

[29] J. Vukman and I. Kosi-Ulbl, *On centralizers of standard operator algebras and semisimple H^*-algebras*, Acta Math. Hungar., **110**(3) (2006), 217–223.

[30] J. Vukman and M. Fošner, *A characterization of two-sided centralizers on prime rings*, Taiwanese J. Math., vol. 11, No. 5 (2007), 1431–1441.

[31] J. Vukman and I. Kosi-Ulbl, *On centralizers of semisimple H^*-algebras*, Taiwanese J. Math., Vol. 11, No. 4 (2007), 1063-1074.

[32] J.K. Wang, *Multipliers of commutative Banach algebras*, Pacific. J. Math., **11**(1961), 1131–1149.

[33] J.G. Wendel, *Left centralizers and isomorphisms of group algebras*, Pacific J. Math., **2**(1952), 251–266.

[34] B. Zalar, *On centralizers of semiprime rings*, Comment. Math. Univ. Carol., **32**(1991), 609–614.

CUBO

A Mathematical Journal

Submitting: By Submitting a paper to this journal, authors certify that the manuscript has not been previously published nor is it under consideration for publication by another journal or similar publication. Work submitted to **CUBO** will be refereed by specialist appointed by the Editorial Board of the journal. The acceptance (or not) of the paper will be communicated to the corresponding author by the Editor in Chief.

Manuscript: Manuscript should be written in English or French and submitted in duplicate to any one of the editors or to the Editorial Office. Authors should send two copies in their final form, printed on one side of the paper with generous margins. The first page should contain a short descriptive title, the name(s) of the author(s), the institutional affiliation and complete address (including e-mail) of each author. Paper should be accompanied by a short abstract, key words and phrases and Mathematical Subject Classification. References are indicated in the text by consecutive arabic numeral enclosed in square bracket. The full list should be collected and typed at the end of the paper in numerical order.

Press requirement: The abstract, no longer than 100 words, must be send in two languages, being english one of these. The **CUBO** Journal strongly encourages the use of LaTeX(or one of the TeXvariants: AMSTeX or AMSLaTeX) for manuscript preparation. References should be headed numerically, alphabetically organized and complete in all its details. Authors' initial should precede their name; journal title abbreviations should follow the style of Mathematical Reviews Index.

Reprint: Electronic reprints(PDF) will be supplied free of charge to the corresponding author.

Exchanges: The **CUBO** Journal accepts exchanges with other mathematical publications. More information can be obtained from the Editor.

For any technical question about **CUBO**, please send a e-mail to: Oscar Neto(oscar@dmat.ufpe.br).