Boletín de la

SOCIEDAD MATEMÁTICA MEXICANA

Tercera Serie

Volumen 19
Número 1
Abril de 2013

Contenido

Continúa/Continued on back cover

SOCIEDAD MATEMÁTICA MEXICANA

A TRIBUTE TO LEOPOLDO GARCÍA-COLÍN SCHERER

On 8th October 2012, Professor Leopoldo García-Colín Scherer, one of the most relevant mexican physicists passed away. His influence in the academic life in Mexico was notorious. During his trajectory he was a Professor in the most important universities and scientific institutes in Mexico City and Puebla.

He was deeply involved in the foundation of the Centro de Investigación y de Estudios Avanzados del IPN. He was also a founder member of the Escuela Superior de Física y Matemáticas del IPN and the Universidad Autónoma Metropolitana, creating an active research group in statistical mechanics, where he was a world leader. Nevertheless, he also ventured in several other branches of physics. He was both the first Distinguished Professor and the first Emeritus Professor of the Universidad Autónoma Metropolitana.

He was elected member of the "Colegio Nacional" in 1977, honored with the 1988 National Prize for Arts and Sciences from Mexico, and awarded an Honoris Causa Doctorate from the Universidad Nacional Autónoma de México in 2007. In this volume, as a tribute to Professor García-Colín, the Boletín de la Sociedad Matemática Mexicana has included some papers in honor of his work and memory.

The Editorial Board

Bol. Soc. Mat. Mexicana (3) Vol. 19, 2013

LEOPOLDO GARCíA-COLíN'S CONTRIBUTIONS IN THE KINETIC THEORY OF MODERATELY DENSE GASES

R. M. VELASCO

ABSTRACT. This paper is written as a tribute to Leopoldo García-Colín Scherer, with the spirit to give emphasis to his scientific contribution in the kinetic theory of moderately dense gases. I have chosen a set of papers representing what I consider his most relevant contribution. It should be mentioned that those papers represent only a small sample of his legacy and, I regret that my limitations, prevent myself to be more ambitious in the review of the great quantity of papers published by him along more than fifty years of his academic life.

Foreword

The 8^{th} of October 2012, a day which at a first sight seemed as any other one, has marked the scientific life in Mexico with the loss of our professor, colleague, mostly friend, Leopoldo García-Colín Scherer. He was the example to be followed by me and I am sure that he was the compass for a lot of my colleagues who undertook their academic life under his guidance. However, not only some of us miss him, but the statistical physics school in Mexico feels a vacuum which will last forever.

1. Introduction

In an effort to put García-Colín's work in perspective, let me recall some important concepts in the kinetic theory of gases according to the periods of development in kinetic theory pointed out by Ernst [22]:

1. Classic Era (1855-1945)
2. Renaissance (1946-1964)
3. Modern Era (1965-1985)
4. Post-modern Era ($>$ 1985)

The *Classic Era* began with Clausius who introduced the mean free path idea in 1858 [12]. Afterwards, Maxwell derived the distribution function of velocities for molecules in a gas in thermal equilibrium [39] up to Boltzmann, who in 1872 formulated his famous equation for the time evolution of the single particle distribution $f(x,t) = f(\mathbf{r},\mathbf{v},t)$ for a dilute gas in equilibrium [3]. Such a long history has been rewieved in an excelent book by Brush, going from the original and main ideas to the equations which allow the transport coefficients calculation [7]. To be specific, let us consider a simple fluid out of the thermodynamic equilibrium state, then one of our goals will be the description of phenomena occuring in such a system. To do that we have different approaches: first, a phenomenological approach which describes the system in a purely macroscopic way. It means that we need a

2010 *Mathematics Subject Classification*: 00A79, 01A70.
Keywords and phrases: kinetic theory of gases.

selection of relevant variables, such as $(n, \mathbf{u}, e)$ where n is the number density, $\mathbf{u}$ the hydrodynamic velocity and e the total specific energy. All these variables are local ones, so they depend on position $\mathbf{r}$ and time t. Once the variables are chosen it is necessary to write the balance equations for them, which are based on the general conservation principles, namely: the continuity equation for the number density (the mass density is obtained by mutiplication of n by the masss m of the particles), a generalization of Newton's second law driving to a balance equation for the momentum density and the equation for the conservation of energy. This procedure drives us to a non closed system of partial differential equations (PDEs) in which it it necessary to introduce the constitutive equations (taken form the corresponding experiments) for the heat flux and the viscous tensor, as well as the local equation of state and the local caloric equation. The last step mentioned above introduces in the balance equations, the properties of the gas under consideration, then the transport coefficients such as the thermal conductivity, shear and bulk viscosities and, some others in more complicated system appear in the theory. Besides this, the complete solution needs the specification of the initial and boundary conditions. The program we have just described is done by the Classical Hydrodynamics, where the behavior of molecules constituting the system does not play any role in the description.

As a second approach we focus our attention on the behavior of the molecules in the system, in order to study some problems that are out of the validity limits of hydrodynamics. In addition, the calculation of transport coefficients from the interaction between particles can be done, at least in principle. It is this approach the one we will develop in what follows.

Let us consider a monatomic gas in which the interaction between particles occurs through molecular collisions, since this one is the most important mechanism in the description of phenomena occuring in the system. To give a qualitative discussion we will relate the molecular collisions with the mean free path, which is the average distance traveled by the molecules without a collision between themselves or with the walls containing the gas. Let σ be the cross section for collision between molecules, if the particles have a typical size a (for example the hard sphere diameter, or the range of the intermolecular potential) the cross section is $\sigma \sim a^2$. If we call n the number of molecules per unit volume in the gas and, L the typical macroscopic length (volume $V = L^3$) the quantity nLa^2 will be the number of collisions occurring within the distance L. Then the mean free path can be estimated as

$$\lambda \sim \frac{L}{na^2L} \sim \frac{1}{na^2}, \tag{1.1}$$

also, λ/L gives us the frequency of collisions and it is known in the literature as the Knudsen number $K_n = \lambda/L$ [44].

The qualitative elements given above allow us to identify three characteristic lengths: (a) The microscopic distance which can be the hard sphere diameter, or the range of the intermolecular potential. (b) The mean free path λ and (c) the macroscopic lenght L. In a dilute gas, those distances are of different order of magnitude: $a \ll \lambda \ll L$, for a typical diluted gas at temperature $T = 273K$ and atmospheric pressure we have $a \sim 10^{-8}cm, \lambda \sim 10^{-5}cm, L \sim 1cm$. Also, this comparison can be made in terms of the characteristic times, which are well separated $t_{micro} \sim 10^{-12}s \ll \tau \sim 10^{-9}s \ll t_{macro} \sim 10^{-4}s$, here τ is the mean free time.

The existence of well separated lenght/time scales gives place to describe the time evolution of our system in terms of well defined stages:

1. **The kinetic stage** which occurs in times $0 < t \leq \tau$, and we will describe the system behavior in term of a single particle distribution function. In the case of a dilute gas, such distribution function satisfies the Boltzmann equation. Then, the particles interact via uncorrelated binary collisions and the molecular chaos hypothesis is valid.
2. **The hydrodynamic stage** describes the system for times $t \gg \tau$ in terms of hydrodynamic variables, those variables correspond to quantities which do not change drastically because of the collisions between molecules, as is the case of conserved variables. For a simple fluid the variables chosen are $n(\mathbf{r},t), \mathbf{u}(\mathbf{r},t), e(\mathbf{r},t)$.
3. **Thermodynamic equilibrium**. After the hydrodynamical stage, the system reaches the thermodynamic equilibrium and all variables become constant.

In a schematic way, the time evolution of the system is described as:

kinetic stage → hydrodynamic stage → thermodynamic equilibrium

$$f(\mathbf{r},\mathbf{v},t) \Longrightarrow \begin{pmatrix} n(\mathbf{r},t) \\ \mathbf{u}(\mathbf{r},t) \\ e(\mathbf{r},t) \end{pmatrix} \Longrightarrow \begin{pmatrix} n = \text{constant} \\ \mathbf{u} = 0 \\ T = \text{constant} \end{pmatrix}. \tag{1.2}$$

In a diluted gas such stages are well defined and it is possible to go from the kinetic stage to the hydrodynamical one and then to equilibrium. The kinetic equation which makes such a task is the Boltzmann equation and its momentum averages drive to the hydrodynamical balance equations. Then the transport coefficients can be caculated from the intermolecular potential. The problem comes when we consider a dense gas, it is clear that in this case the stages are not well defined and it has been shown that the kinetic stage does not exist. It is at this point where the most important contibution of Leopoldo García-Colín was done.

In section 2 we will present the Boltzmann equation, then in section 3 the generalized Boltzmann equation will be discussed. In section 4 we present the treatment done for the kinetic stage in a moderately dense gas, whereas in section 5 we discuss the hydrodynamical stage. In section 6 some discussion is pointed about the study of binary mixtures and finally in section 7 we give some concluding remarks.

2. The Boltzmann equation

To describe a macroscopic system with $N \sim 10^{23}$ particles from their dynamics, it is necessary to introduce statistical concepts. Therefore we consider a one species monatomic gas formed with particles of mass m described through the phase space coordinates $x_i = (\mathbf{r}_i, \mathbf{v}_i)$. The quantity $f(\mathbf{v},\mathbf{r},t)d\mathbf{r}d\mathbf{v}$ represents the number of molecules in $\mathbf{v}+d\mathbf{v}$, $\mathbf{r}+d\mathbf{r}$ at time t. The time evolution of the distribution function $f(\mathbf{v},\mathbf{r},t)$ is described with the equation introduced by Boltzmann in 1872 in a heuristic way; in fact it is a balance equation in which the time variation of the distribution function is given through its drift in phase space and the interaction between molecules,

$$\frac{\partial f}{\partial t}+\mathbf{v}\cdot\nabla f=\mathcal{J}(f,f), \tag{2.1}$$

where it is considered that there is no external force acting on the system and, $\mathcal{J}(f,f)$ represents the so called collision kernel. It contains the effect of the binary collisions. Aside of this, the collision kernel which in principle must contain the distribution function associated to the two colliding particles is factorized, so

$$f^{(2)}(x_1,x_2,t)\to f(x_1,t)f(x_2,t)\sim f(\mathbf{r},\mathbf{v}_1,t)f(\mathbf{r},\mathbf{v}_2,t), \tag{2.2}$$

which is an expression based on the "molecular chaos" hypothesis. It is clear that, the binary collision assumption and the molecular chaos hypothesis are both valid only for dilute gases. The collision kernel can be written as follows,

$$\mathcal{J}(f,f)=\int\int(f'f_1'-ff_1)g\sigma(\phi,g)d\mathbf{v}_1d\hat{e}, \tag{2.3}$$

where $f'=f(\mathbf{v}',\mathbf{r},t)$, $f_1'=f(\mathbf{v}'_1,\mathbf{r},t)$ and $\mathbf{v}',\mathbf{v}'_1$ correspond to the velocities of particles after the collision and the velocities without a prime are the ones before the collision. The quantity $\mathbf{g}=g\hat{\mathbf{e}}$ is the relative velocity between particles and g its magnitude, ϕ is the dispersion angle in the collision, $\sigma(\phi,g)$ the cross section.

It is well known that the Boltzmann equation (2.1) has a solution representing the local equilibrium state, which is the Maxwell distribution function

$$f^{(0)}=n\left(\frac{m}{2\pi kT}\right)^{3/2}exp\left[-\frac{m|\mathbf{v}-\mathbf{u}^2|}{2kT}\right], \tag{2.4}$$

where the macroscopic variables $(n,\mathbf{u},T)$ are functions of $(\mathbf{r},t)$ and in the thermodynamic equilibrium state they become (n = constant, $\mathbf{u}=0,T$ = constant). Also, it has been shown that the functional

$$\mathcal{H}[f]=\int f(\mathbf{v},\mathbf{r},t)(Ln\ f(\mathbf{v},\mathbf{r},t)-1)d\mathbf{v}d\mathbf{r}, \tag{2.5}$$

written in terms of the distribution function f satisfies the $\mathcal{H}-$theorem which is given as

$$\frac{d\mathcal{H}[f]}{dt}\leq 0, \tag{2.6}$$

for $f(\mathbf{v},\mathbf{r},t)$ a solution of the Boltzmann equation (see [45], [9] and references therein, for rigorous proofs and mathematical properties concerning the Boltzmann equation). Some consequences of the $\mathcal{H}-$theorem can be seen immediately: for finite kinetic energy, $\mathcal{H}$ is bounded when $t\to\infty$ and this limit correspond to the case when $d\mathcal{H}/dt=0$. In such a case the distribution function corresponds to the Maxwell distribution function we have just quoted in Eq. (2.4). Besides if we calculate $\mathcal{H}[f^{(0)}]$ it gives us the negative of the ideal gas thermodynamic entropy.

Starting with the Boltzmann equation it is possible to construct the transport equations as was done by Maxwell [10], [25] for any function $\Psi(\mathbf{v},\mathbf{r},t)$ of the molecular velocity, then we define the average of such a function as

$$n(\mathbf{r},t)\overline{\Psi}(\mathbf{r},t)=\int\Psi(\mathbf{v},\mathbf{r},t)f(\mathbf{v},\mathbf{r},t)d\mathbf{v}. \tag{2.7}$$

If we take $\Psi = 1$ we obtain the number density, with $\Psi = \mathbf{v}$ the hidrodynamic velocity, lastly when we take $\Psi = m(\mathbf{v}-\mathbf{u})^2/2$, the energy is obtained

$$n(\mathbf{r},t) = \int f(\mathbf{v},\mathbf{r},t)d\mathbf{v}, \tag{2.8}$$

$$n(\mathbf{r},t)\mathbf{u}(\mathbf{r},t) = \int \mathbf{v}f(\mathbf{v},\mathbf{r},t)d\mathbf{v}, \tag{2.9}$$

$$n(\mathbf{r},t)e(\mathbf{r},t) = \int \frac{1}{2}m(\mathbf{v}-\mathbf{u})^2 f(\mathbf{v},\mathbf{r},t)d\mathbf{v} = \frac{3}{2}n(\mathbf{r},t)kT(\mathbf{r},t). \tag{2.10}$$

Notice that the local temperature $T(\mathbf{r},t)$ is now defined through the local energy, essentially as the average of the kinetic energy, a fact which will be very important in dense gases. Once this procedure is done, a set of definitions must be advanced the pressure tensor is written as

$$\underline{P} = \int m(\mathbf{v}-\mathbf{u})(\mathbf{v}-\mathbf{u})f(\mathbf{v},\mathbf{r},t)d\mathbf{v}, \tag{2.11}$$

and the heat flux

$$\mathbf{q} = \int \frac{1}{2}\left[m(\mathbf{v}-\mathbf{u})^2 - 5kT\right]f(\mathbf{v},\mathbf{r},t)d\mathbf{v}. \tag{2.12}$$

Now, to obtain the general transport equations, we multiply the Boltzmann equation (2.1) by the function Ψ and integrate over the velocity

$$\frac{\partial(n\overline{\Psi})}{\partial t} + \nabla_{\mathbf{r}}\cdot(n\overline{\mathbf{v}\Psi}) - n\left[\overline{\frac{\partial\Psi}{\partial t}} + \overline{\mathbf{v}\cdot\nabla_{\mathbf{r}}\Psi}\right] = \int \Psi(\mathbf{v})\mathcal{J}(f,f)d\mathbf{v}. \tag{2.13}$$

The corresponding balance equations obtained from here, are then written as

$$\frac{\partial n}{\partial t} + \nabla_{\mathbf{r}}\cdot n\mathbf{u} = 0, \tag{2.14}$$

$$\frac{\partial}{\partial t}nm\mathbf{u} + \nabla_{\mathbf{r}}\cdot(nm\mathbf{u}\mathbf{u} + \underline{P}) = 0, \tag{2.15}$$

$$\frac{\partial}{\partial t}nme + \nabla_{\mathbf{r}}(nme\mathbf{u} + \mathbf{q}) = -\underline{P}:\nabla_{\mathbf{r}}\mathbf{u}, \tag{2.16}$$

and they are valid in absence of external forces. Also, it should be noted that the tensor $\underline{P}$ contains the hydrostatic pressure and the viscous tensor. On the other hand, the local Maxwell distribution function when substituted in the balance equations (2.14, 2.15, 2.16) yields to the Euler equations of classical hydrodynamics. This fact means that the description done with the Maxwell distribution function will correspond to a fluid in which the viscosity and thermal conductivity vanish, i. e. a fluid in which there are not any dissipative effects.

To go further, the Chapman-Enskog method to solve, albeit in an approximate way, the Boltzmann equation is needed [10], [25]. In this method the distribution function is expressed as an expansion around the local equilibrium described by the Maxwell distribution function, so

$$f(\mathbf{v},\mathbf{r},t) = f^{(0)}(\mathbf{v},\mathbf{r},t)\left[1 + \Phi^{(1)}(\mathbf{v},\mathbf{r},t) + \dots\right]. \tag{2.17}$$

In addition an assumption called the "functional hypothesis" is also done, it says that the time and spatial dependence in the distribution function occurs only through the time dependence in the local variables $n, \mathbf{u}, e$, so $f(\mathbf{v},\mathbf{r},t) = f(\mathbf{v}|n(\mathbf{r},t), \mathbf{u}(\mathbf{r},t), e(\mathbf{r},t))$. With this assumption we are going from the kinetic to the hydrodynamic stage, in which the system is described by the variables $(n, \mathbf{u}, e)$. Then the expansion (2.17) and the functional hypothesis allow us to find a solution valid up to first order in the spatial gradients in the macroscopic variables chosen. Once this long procedure is done, the distribution function obtained is susbtituted in the expressions for the pressure tensor and the heat flux,in order to obtain the shear viscosity η_0 and the thermal conductivity λ_0 in the gas, defined as

$$\underline{p}^0 = -2\eta_0(\nabla_{\mathbf{r}}\mathbf{u})^0, \quad \mathbf{q} = -\lambda_0\nabla_{\mathbf{r}}T. \tag{2.18}$$

Obviously, such results will be valid only for dilute gases (the zeroth density limit) and, it must be said that their comparison with experimental data for monatomic gases, taken at the low density regime are very good.

Not withstanding the strong limitations in the Boltzmann equation and its method of solution, the results provided by this theory have been the corner stone in the kinetic theory of dilute gases.

3. The generalized Boltzmann equation

The *Renaissance* in kinetic theory began essentially in 1946 when Bogoliubov [2], [25] and almost simultaneously Born, Green, Kirkwood and Yvon used the now called BBGKY-hierachy, [11]. It is constructed from the Liouville equation for a classical infinite system of $N\to\infty$ particles with mass m in a volume $V\to\infty$ with $N/V = n$ =constant, interacting via an additive pairwise potential $V(|\mathbf{r}_{12}|)$, by means of the integration over the x_{N-s} particle phase space coordinates to obtain the $s-$particle distribution function. With this procedure, the F_s-particle distribution function satisfies an equation coupled to the F_{s+1} and so on. Then the complete infinity hierarchy of equations are equivalent to the Liouville equation. Besides, Bogoliubov's theory assumes: (a) The $s-$particle distribution function is a time independent functional of the single particle distribution function $F_1(x_1,t)$. (b) The long time decay of the initial correlations are expressed as

$$\lim_{t\to\infty}\left[\mathcal{S}_t^{(s)}(F_s|S_t^{(1)}F_1) - \Pi_{i=1}^{s}S_t^{(1)}F_1\right] = 0. \tag{3.1}$$

These two assumptions in Bogoliubov's theory deserve some physical explanation, the first one is consistent with the description made in section (1) about the different stages to reach equilibrium. It means that after some characteristic time, the time evolution of the $s-$particle distribution function is expressed only in terms of the evolution of the single-particle one, consistently with the existence of the kinetic stage. The second assumption tells us that the long time behavior is such that the time evolution in the $s-$particle distribution function is factorized in terms of single particle distribution functions. A fact, which is consistent with the absence of corrrelations between particles after some characteristic time. Both assumptions imply that the system cannot be far from the equilibrium state [17].

The first two hierarchy equations read as

$$\left(\frac{\partial}{\partial t} + \mathbf{v}_1\cdot\nabla_1\right)F_1 = n\int \theta_{12}F_2(x_1,x_2,t)dx_2, \tag{3.2}$$

$$\left(\frac{\partial}{\partial t} + \mathcal{L}_{12}\right)F_2 = n\int [\theta_{13}+\theta_{23}]F_3(x_1,x_2,x_3,t)dx_3, \tag{3.3}$$

where

$$\theta_{12} = \frac{\partial V(r_{12})}{\partial \mathbf{r}_{12}} \cdot \frac{1}{m}\left(\frac{\partial}{\partial \mathbf{v}_1} - \frac{\partial}{\partial \mathbf{v}_2}\right), \tag{3.4}$$

$$\mathcal{L}_{12} = \mathbf{v}_1 \cdot \nabla_1 + \mathbf{v}_2 \cdot \nabla_2 - \theta_{12}, \tag{3.5}$$

here θ_{12} is the weak scattering operator and $\mathcal{L}_{12}$ is the two-particle Liouville operator.

In fact, the so called Generalized Boltzmann Equation (GBE) corresponds to the first hierarchy equation (3.2) which is closed with the second assumption in Bogoliubov's theory, then

$$\frac{\partial f^{(1)}}{\partial t} + \frac{\mathbf{p}_1}{m} \cdot \frac{\partial f^{(1)}}{\partial \mathbf{q}_1} = \Phi(x_1 | f^{(1)}), \tag{3.6}$$

$$\Phi(x_1 | f^{(1)}) = \int \frac{\partial V}{\partial \mathbf{q}_1} \cdot \frac{\partial}{\partial \mathbf{p}_1} f^{(2)}(x_1, x_2 | f^{(1)}) dx_2. \tag{3.7}$$

To establish the conection with the hydrodynamic variables we will use the definitions given in (2.14, 2.15) and the energy must be modified in order to take into account the interaction potential, so

$$e(\mathbf{r}, t) = \int f_1 \frac{(\mathbf{v} - \mathbf{u})^2}{2} d\mathbf{p} + \frac{1}{2} \int V(|\mathbf{r}_{12}|) f_2(x_1, x_2, t) dx_2 d\mathbf{p}. \tag{3.8}$$

Also, the definitions of the pressure tensor and the heat flux become different due to the presence of the interaction potential, then

$$\underline{P}^K = \frac{1}{m} \int f_1 \mathbf{p}\mathbf{p} d\mathbf{p}, \tag{3.9}$$

$$\underline{P}^V = -\frac{1}{2} \int d\mathbf{p} dx_2 \, \frac{V'}{r} \mathbf{r}\mathbf{r} \int_0^1 d\mu \, f_2(\mathbf{r} + \mu \mathbf{r}, (\mu + 1)\mathbf{r}), \tag{3.10}$$

$$\mathbf{q}^K = \int d\mathbf{p} \frac{p^2}{2m} \frac{\mathbf{p}}{m} f_1, \tag{3.11}$$

$$\mathbf{q}^{V_1} = -\frac{1}{4} \int d\mathbf{p} dx_2 \, V' \frac{\mathbf{r}\mathbf{r}}{r} \cdot \left(\frac{\mathbf{p}}{m} + \frac{\mathbf{p}_2}{m}\right) \int_0^1 d\mu f_2(\mu \mathbf{r}, (\mu + 1)\mathbf{r}), \tag{3.12}$$

$$\mathbf{q}^{V_2} = \frac{1}{2} \int d\mathbf{p} dx_2 \, \frac{\mathbf{p}}{m} V(r) f_2, \tag{3.13}$$

where we have written $\mathbf{r} = \mathbf{r}_{12}$ to shorten the notation and, the dependence on momenta in the distribution functions is not written out explicitly, though it is present. The balance equations are the same as the ones valid for a dilute fluid, however the pressure tensor and the heat flux must be calculated according to the definitions given above.

Once the kinetic equation is given we need to obtain a solution and calculate the transport coefficients. In section (2) we mentioned the Chapman-Enskog method to perform such a task. However, it is well known that the properties of dense fluids in the equilibrium state can be expressed as expansions in density, in such a way that the first order in the density represents the contribution of clusters with two particles, the second order corresponds to clusters with three particles and so on. Then it is natural to ask for the corresponding density expansion in the transport coefficients. These considerations can be sinthesized by saying that we need a solution for the GBE which will reproduce the gradient and the density expansions in transport coefficients.

4. The kinetic stage in a moderately dense gas

As a first intent to solve the BBGKY hierarchy, Bogoliubov [2], Choh [11] and Cohen [14] expanded the distribution function in powers of the density

$$F_s(x_1,\dots,x_s|F_1) = F_s^{(0)}(x_1,\dots,x_s|F_1) + nF_s^{(1)}(x_1,\dots,x_s|F_1) + \dots, \tag{4.1}$$

where the superindex indicates the order of approximation in the density. Consequently the GBE now reads as

$$\frac{\partial F_1}{\partial t} + \frac{\mathbf{p}}{m}\cdot\frac{\partial F_1}{\partial \mathbf{r}} = n\int dx_2\, \theta_{12}\left[F_2^{(0)}(x_1,x_2|F_1) + nF_2(x_1,x_2|F_1) + \dots\right], \tag{4.2}$$

which is similar to the density expansions in the study of gases in the equilibrium state. Though this expansion exists in dilute gases at the equilibrium state it is assumed that it also does for the dynamical behavior in dense gases, an assumption which cannot be taken for granted. It should be noticed that the method of solution proposed in Eq. (4.2) does the expansion in powers of the density, before undertaking its solution for the single particle distribution function F_1. Then the Chapman-Enskog method, up to first order in the gradients of the hydrodynamical velocity and the temperature, allowed them to find expressions for the transport coefficients.

On the other hand García-Colín in a series of papers [32], [29], [27], [28] found a solution up to first order in the gradients by means of the Chapman-Enskog method valid to all orders in the density. That is, he and his collaborators considered the solution up to first order in the gradients before the expansion in the density. Though the first order correction in the density coincided with the results calculated by Choh et al. At the same time, several authors were working on the same problem [23], [13], [14], [20] [24], [37] and their results and discussions have shown that the expansion in density diverges after the second order, even though all approaches coincide to the first order in density [15], [16], [20], [19], [24] [43]. Even Prigogine's theory was taken into account [6], [26] and it was shown that the results in this theory agree with the other approaches. Going back to the periods identified in the development of kinetic theory we can see that all these calculations and discussions were done just at the end of the *Renaissance* and the begining of the *Modern Era*.

(4.1) The definition of temperature. At first sight, all the calculations mentioned above must give the same result for transport coefficients, however it was not the case. In particular the bulk viscosity was not the same when calculated with the Choh's assumption as when calculated by García-Colín. Also, Ernst [23] calculated transport coefficients by means of autocorrelation method and the result was the same as García-Colín's. The discrepancy was an important one, and it was necessary to go deeper in both methods to discover that though the expressions for the bulk viscosity were different, the results were completely equivalent. Let us go now into some details about this problem.

First of all, let us recall Eq. (2.10) where the local temperature was defined as the average of the kinetic energy in a diluted gas. Taking such a temperature definition as a valid one in dense gases, some authors took the kinetic energy average as a definition of temperature, simply making an extension from the dilute gas behavior. However in a dense or a moderately dense gas the transport of energy is not purely kinetic, so the definition of temperature must be taken with the complete average of energy. It means that the temperature must contain the contribution of the average potential energy as well as the kinetic contribution. In a very important paper García-Colín and Green [31] have shown that there is a complete equivalence of results in the trasnport coefficients, no matter the temperature definition. In fact, once we adopt a temperature definition the consistent calculation gives a result for the bulk viscosity which is equivalent to the other one, even though the expressions seem to be different. The translation between themselves shows the equivalence.

In fact, there exists a one-to-one relation between the macroscopic variables in both methods in such a way that

$$F_1(x|n,\mathbf{u},T) = F_1(x|n,\mathbf{u},e), \tag{4.3}$$

where T is the local temperature defined only in terms of kinetic energy. The variables $(n,\mathbf{u},e)$ recomend themselves because they are approximate single-valued integrals of motion of the system.

5. The hydrodynamic stage in a moderately dense gas

The study of the hydrodynamical stage in dense gases means that the single-particle distribution function is a functional of the macroscopic variables $(n,\mathbf{u},e)$. In other words, we need the solution of the GBE (3.6) in terms of such variables and their spatial gradients. As we said above, García-Colín [32] proposed to find such a solution making an expansion in the gradients valid to all orders in the density, this means that

$$\Phi(x_1|f_1) = \Phi(x_1|f_1(\mathbf{q})) + \int dx'\Phi'(x_1,x'|f_1(\mathbf{q}))(\mathbf{q}'-\mathbf{q})\cdot\left(\frac{\partial f_1}{\partial \mathbf{q}'}\right)_{\mathbf{q}'=\mathbf{q}}, \tag{5.1}$$

where $f_1 = nF_1$ and, Φ' denotes the functional derivative of Φ. Then an expansion similar to Eq. (2.17) is done with the conditions imposed by Eqs. (2.14, 2.15,2.16). As a result the Maxwell distribution function is found to describe the total equilibrium and, the two-particle distribution function is found to satisfy

$$f_2(x_1,x_2|f_1^{(0)}(\mathbf{q})) = f_1^{(0)}(\mathbf{p}_1)f_1^{(0)}(\mathbf{p}_2)G(|\mathbf{r}_{12}|), \tag{5.2}$$

here the quantity $G(|\mathbf{r}_{12}|)$ is the pair correlation function. The following step is done by means of the application of the Chapman-Enskog method, which gives us a solution expressed in terms of some functions satisfying a pair of integral equations. After a cumbersome but direct algebra, the thermal conductivity and the shear viscosity as defined in eqs. (2.18) can be constructed:

$$\lambda = \lambda^K + \lambda_1^{V_1} + \lambda_1^{V_2} + \lambda_2^{V_1} + \lambda_2^{V_2}, \tag{5.3}$$

$$\eta = \eta^K + \eta^{V_1} + \eta^{V_2}, \tag{5.4}$$

$$\zeta = \zeta^K + \zeta^{V_1} + \zeta^{V_2}, \tag{5.5}$$

where the contributions with superindexes K, V_1, V_2 correspond to the kinetic part and to the potential parts respectively. The complete expressions are given in reference [32]. It should be noticed that such expressions are only formal, that is without the use of any particular intermolecular potential, in such a way that when we take a specific potential the solution of some integral equations as well as all integrations must be performed.

On the other hand, the expansions in powers of the density can also be done and in fact it was shown that all methods give the same thermal conductivity, shear and bulk viscosities only up to first order in the density. Besides, the bulk viscosity is quantity of second order in the density [28].

(5.1) The challenges involved in the transport coefficients calculation. The program we have summarized above sounds almost finished in the sense that the formal expressions are given and all we have to perform is a calculation of the elements involved. However, it was soon discovered that the expansions in the density are divergent [43]. In fact, the expansions in the density take into account the contributions of the dynamics of clusters of particles and it has been shown [20] that they cannot be used for two purposes, the computation of the long-time behavior of F_2 beyond $\mathcal{O}(n)$ and the demonstration of the decay of the initial state beyond $\mathcal{O}(n^2)$. Similar divergences were encountered in the computation of the transport coefficients from time-correlation functions. The nature of those divergences suggest (a) there is no kinetic stage in the approach of a dense gas to equilibrium,in the Bogolyubov sense, (b) a weak logarithmic density dependence of the transport coefficients. The explicit calculation lead to expressions like

$$\eta(n,T) = \eta_0(T) + \eta_1(T)n_* + \eta_2'(T)n_* Ln\, n_* + \eta_2''(T)n_*^2 Ln\, n_* + \dots, \tag{5.6}$$

where all coefficients have been computed for a gas of hard spheres, and n_* is a dimensionless density. Similar behavior appears for the thermal conductivity.

The understanding of such divergences in the density expansion for transport coefficients has taken several years, first it was necessary to be sure that the divergences were present no matter the method employed to solve the GBE. Then, in a second step to identify the divergent contributions and study why such contributions diverge [19]. Essentially, the divergences problem comes from the dynamical aspects of collisions for three or more particles. Let us think in a cluster of three particles and let us name the particles as 1, 2, 3, then there can be simultaneous collisions between all the three particles (genuine 3-particle collision) but also the collision occur in the sequence (21) → (13) → (12) called as a recollision. Also, the sequence (21) → (13) → (32) is possible and it is called as a cyclic collision. In the first case, the phase space available to the genuine collision is finite and is determined by the potential range, however in the other cases and when bigger clusters of particles occur, the available phase space diverges logarithmically.

The solution for this problem has been discussed in the literature [16], [17] where the consideration of the cage and vortex diffusion effects lead to a non divergent calculation. No matter this fact, there are some unsolved problems in the kinetic theory of dense gases.

6. Binary mixtures

So far, we have studied a simple fluid composed by one species of molecules meaning that all molecules have the same mass. An interesting problem is pointed when we consider multicomponent fluids due to the fact that new transport phenomena can take place. Let us concentrate in a binary mixture, then diffusion, Dufour and Soret effects can be present in a neutral mixture. When the mixture is a chemically active one, there can be some chemical reactions in the system and consequently there are more transport phenomena to be described.

On the other hand, let us say that the GBE is not the only approach which has been developed to study dense gases. The Enskog equation was constructed in 1922 by Enskog [8] and it is based on a heuristic generalization of the Boltzmann equation, in contrast with the GBE which comes from the Liouville equation. For a multicomponent system it can be written as follows

$$\frac{\partial f_i}{\partial t} + \mathbf{v}_i \cdot \nabla_{\mathbf{r}} f_i = \sum_{j=1}^{2} \int \left[\chi_{ij}(\mathbf{r}_j + y_{ij}\mathbf{k}) f_j'(\mathbf{r}_i + \sigma_{ij}\mathbf{k}) f_i'(\mathbf{r}_i) - \chi_{ij}(\mathbf{r}_i - y_{ij}\mathbf{k}) f_j(\mathbf{r}_i - \sigma_{ij}\mathbf{k}) f_i(\mathbf{r}_i)\right] \sigma_{ij}^2 (\mathbf{g}_{ji}\cdot\mathbf{k}) d\mathbf{k} d\mathbf{v}_j. \tag{6.1}$$

The Enskog equation contains the following hipotheses: (1) Only binary collisions are taken into account. (2) The molecular chaos assumption is made, i.e., the correlations between positions and velocities of two particles in phase space are neglected. (3) The function χ_{ij} accounts for the shielding and the excluded volume in collisions between molecules of species i and j. This function is evaluated at an arbitrary point located between the centers of the colliding molecules [1], [33], [40], [41].

In a first set of papers García-Colín studied the compatibility of kinetic theory as described by the Enskog equation with hamiltonian dynamics, thermodynamics, hydrodynamics and the thermodynamics of a chemically reacting fluid [34], [35], [36]. In fact, the compatibility was shown by answering several questions, such as: Is it possible to derive the kinetic equation from a more microscopic view of the system considered, for example, from the Hamiltonian dynamics of the particles composing the system? What error we make by replacing the Hamilton dynamics by the kinetic equation? Does there exist an approach, as the time goes to infinity, of solutions of the kinetic equation to a time-independent state as considered in thermodynamics? What is the equation of state implied by the kinetic equation? Is it possible to replace the kinetic equation by the hydrodynamic equations if our interest is focused only on the long-time behavior of solutions to the kinetic equation? What is the error that we make by this replacement? What are the kinetic coefficients in the hydrodynamic equations? All questions were studied for the Enskog equation and some important answers such as the validity of an $\mathcal{H}-$theorem, the conditions that the function χ_{ij} (measures the correlation between particles in a binary collision) must satisfy, the Onsager-Casimir symmetry for transport coefficients among others, were given.

Also, the generalization of the GBE for binary mixtures was studied in some other papers [4], [30], [5] together with problems which were in the literature at the same time [18], [21]. Coming back to the classification in the development in kinetic theory, we can say that we are now in the *Post Modern Era*, though a lot of problems remain to be solved. Also, some other approaches have been developed to study the phenomena in dense gases, as an example of such alternatives we have the so called Kinetic Variational Theory (KVT). In this approach we have two main ingredients to work with, the first one is based on the hierarchy equations as writen in (3.2, 3.3) and a definition of the Shanon information entropy. Then the maximization of the entropy and the equations of motion allow the construction of a set of closed equations in which we have taken all the available information for the system [38], [42], see references therein for a detailed description.

7. Concluding remarks

To summarize, let us give some emphasis to the García-Colín's main contributions by saying that he was a pioneer in the development of the Statistical Physics School in Mexico. His works in the kinetic theory of moderately dense gases were the corner stone in the construction of a very big group of researches interested in phenomena occuring out of the equilibrium. Besides, this interest has been diversified to other specialities related directly or indirectely with the kinetic theory. He was not only the promoter of such themes but an enthusiastic participant in several related fields. No matter his main interest, he was always able to give a suggestion in the solution of problems, going from physics, teaching at all levels, applied physics like pollution or biological problems, up to politics of science. His work as a pioneer as well as his legacy has given an exceptional example of a full scientific life.

Acknowledgements

The author would like to thank Dra. S. M. T. de la Selva for her kindness and careful reading of the manuscript.

Received March 15, 2013

Final version received April 16, 2013

R. M. VELASCO
DEPARTMENT OF PHYSICS,
UNIVERSIDAD AUTÓNOMA METROPOTITANA
IZTAPALAPA, 09340, MÉXICO
rmvb@xanum.uam.mx

REFERENCES

[1] L. BARAJAS, L. S. GARCÍA-COLÍN, AND E. PIÑA, *On the Enskog-Thorne theory for a mixture of dissimilar rigid spheres*, J. Stat. Phys., **7** (2), (1973), 161–183.

[2] N. N. BOGOLIUBOV, *Problems of a dynamical theory in statistical physics*, vol. I of Studies in Statitical Mechanics, North Holland, 1962.

[3] L. BOLTZMANN, *Further studies on the thermal equilibrium of gas molecules*, (1872). Reprinted in [7].

[4] E. BRAUN, A. FLORES, AND L. S. GARCÍA-COLÍN, *On the nonequilibrium statistical mechanics of a binary mixture. I. the distribution functions*, J. Stat. Phys., **8** (3), (1973), 239–248.

[5] ———, *A convergent nonequilibrium statistical mechanical theory for dense gases. iii. transport coefficients to second order in the density*, J. Stat. Phys., **10** (1), (1974), 49–58.

[6] E. BRAUN AND L. S. GARCÍA-COLÍN, *Note on the equivalence of the Bogoliubov and Prigogine theories*, Phys. Lett., **23** (7), (1966), 460–461.

[7] S. G. BRUSH, *The Kinetic Theory of Gases, An Anthology of Classical Papers with Historical Commentary*, vol. 1 of History of Modern Physical Sciences, Imperial College Press, 2003.

[8] ———, *History of the kinetic theory of gases*, vol. 7 of Storia della Scienza, Istituto della Enciclopedia Italiana, 2004, ch. 44.

[9] C. CERCIGNANI, *The Boltzmann Equation and its Applications*, vol. 67 of Applied Mathematical Sciences, Springer-Verlag, 1988.

[10] S. CHAPMAN AND T. G. COWLING, *The Mathematical Theory of Non-uniform Gases*, Cambridge Univ. Press, third ed., 1990.

[11] S. T. CHOH AND G. E. UHLENBECK, *The kinetic theory of phenomena in dense gases*. Navy Theoretical Physics. Report 1224, February 1958.

[12] R. CLAUSIUS, *On the mean lengths of the paths described by separate molecules of gaseous bodies*, Phil. Mag., **17**, (1858), 81–91. This reference corresponds to the English translation.

[13] E. G. D. COHEN, *On the generalization of the Boltzmann equation to general order in the density*, Physica, **28**, (1962), 1025–1044.

[14] ———, *On the kinetic theory of dense gases*, J. Math. Phys., **4**, (1963), 183–189.

[15] ———, *Kinetic theory of non-equilibrium fluids*, Physica A, **118**, (1983), 17–42.

[16] ———, *Fifty years of kinetic theory*, Physica A, **194**, (1993), 229–257.

[17] ———, *Non existence of a density expansion of the transport coefficients in a moderately dense gas*, J. Stat. Phys., **137**, (2009), 878–889.

[18] E. G. D. COHEN AND M. H. ERNST, *Kinetic theory of moderately dense gas mixtures. I.*, Physica, **50**, (1970), 177–205.

[19] J. R. DORFMAN, *Advances and challenges in the kinetic theory of gases*, Physica A, **106**, (1981), 77–101.

[20] J. R. DORFMAN AND E. G. D. COHEN, *Difficulties in the kinetic theory of dense gases*, J. Math. Phys., **8**, (1967), 282–297.

[21] M. H. ERNST, *Kinetic theory of moderately dense gas mixtures II. symmetry relations for transport coefficients*, Physica, **50**, (1970), 477–498.

[22] ———, *Bogoliubov Choh Uhlenbeck theory: cradle of modern kinetic theory*, in Progress in Statistical Physics, W. e. a. Sung, ed., Singapore, 1998, World Scientific Publ. Proceedings of the International Conference on Statitical Physics in Memory of Prof. Soo-Tahk Choh.

[23] M. H. ERNST, J. R. DORFMAN, AND E. G. D. COHEN, *Transport coefficients in dense gases I. the dilute and moderately dense gas*, Physica, **31**, (1965), 493–521.

[24] M. H. ERNST, L. K. HAINES, AND J. R. DORFMAN, *Theory of transpor coefficients for moderately dense gases*, Rev. Mod. Phys., **41** (2), (1969), 296–316.

[25] J. H. FERZIGER AND H. G. KAPER, *Mathematical thoery of transport processes in gases*, North-Holland, 1972.

[26] L. S. GARCÍA-COLÍN AND E. BRAUN, *Equivalence of the Bogoliubov and Prigogine theories*, Il Nuovo Cimento, **50** (2), (1967), 193–198.

[27] L. S. GARCÍA-COLÍN AND A. FLORES, *Note on the transport coefficients of moderately dense gases*, Physica, **32**, (1966), 444–449.

[28] ———, *of Choh-Uhlenbeck method in the kinetic theory of dense gases*, J. Math. Phys., **7**, (1966), 254–259.

[29] ———, *On the transport coefficients of moderately dense gases*, Physica, **32**, (1966), 289–303.

[30] L. S. GARCÍA-COLÍN, A. FLORES, AND E. BRAUN, *On the nonequilibrium statistical mechanics of a binary mixture. II. the transport coefficients*, J. Stat. Phys., **8** (3), (1973), 249–264.

[31] L. S. GARCÍA-COLÍN AND M. S. GREEN, *Definition of temperature in the kinetic theory of dense gases*, Phys. Rev., **150** (1), (1966), 153–158.

[32] L. S. GARCÍA-COLÍN, M. S. GREEN, AND F. CHAOS, *The Chapman-Enskog solution of the generalized Boltzmann equation*, Physica, **32**, (1966), 450–478.

[33] P. GOLDSTEIN AND L. S. GARCÍA-COLÍN, *Irreversible thermodynamics of a binary mixture of hard spheres: The Onsager relations*, Physica A, **242**, (1997), 467–481.

[34] M. GRMELA AND L. S. GARCÍA-COLÍN, *Compatibility of the Enskog kinetic theory with thermodynamics. i.*, Phys. Rev. A, **22** (3), (1980), 1295–1304.

[35] ———, *Compatibility of the Enskog kinetic theory with thermodynamics. II. chemically reacting fluids*, Phys. Rev. A, **22** (3), (1980), 1305–1314.

[36] M. GRMELA, R. ROSEN, AND L. S. GARCÍA-COLÍN, *Compatibility of the Enskog kinetic theory with hydrodynamics.*, J. Chem. Phys., **75**, (1981), 5474–5484.

[37] D. K. HOFFMAN AND C. F. CURTISS, *Kinetic theory of dense gases. III. the genaralized Enskog equation*, Phys. Fluids, **7**, (1964), 1887–1897.

[38] J. KARKHECK AND G. STELL, *Kinetic mean-field theories*, J. Chem. Phys., **75**, (1981), 1475–1487.

[39] J. C. MAXWELL, *On the dynamical theory of gases*, Phil. Trans. Royal Soc. of London, **157**, (1867), 49–88.

[40] E. PIÑA, L. S. GARCÍA-COLÍN, AND P. GOLDSTEIN, *A new generalization of Enskog's equation for mixtures of dense gases*, Physica A, **217**, (1995), 87–106.

[41] J. A. ROBLES-DOMÍNGUEZ AND L. S. GARCÍA-COLÍN, *Kinetic theory of moderately dense gas mixtures. III. transport coefficients for binary mixtures of hard spheres*, Physica, **73**, (1974), 510–526.
[42] L. ROMERO-SALAZAR, M. MAYORGA, AND R. M. VELASCO, *Maximum entropy formalism for a dense gas*, Physica A, **237**, (1997), 150–168.
[43] J. V. SENGERS, D. T. GILLESPIE, AND J. J. PEREZ-ESANDI, *Three-particle collision effects in the transport properties of a gas of hard spheres*, Physica A, **90**, (1978), 365–409.
[44] H. SRUCHTRUP, *Macroscopic Transport Equations for Rarefied Gas Flows*, Springer, 2005.
[45] C. VILLANI, *Entropy Production and Convergence to Equilibrium*, vol. 1916 of Lecture Notes in Mathematics, Springer, 2008, ch. 1.

Bol. Soc. Mat. Mexicana (3) Vol. 19, 2013

CONTRIBUTIONS OF LEOPOLDO GARCÍA-COLÍN SCHERER ON GLASS TRANSITION

Dedicated to the memory of our friend and teacher Leopoldo García-Colín Scherer

PATRICIA GOLDSTEIN AND LEONARDO DAGDUG

ABSTRACT. A review work on some of the most important contributions made by Leopoldo García-Colín Scherer on relaxation phenomena that occur both in fragile and strong glass-formers is presented.

1. Introduction

In the mid 80's, Prof. Leopoldo García-Colín was working on relaxation phenomena in several kinds of systems. Particularly, he had been studying the response function to an external perturbation in viscous and viscoelastic fluids. One evening in 1988, he began to discuss with his co-workers a paper that he had been reading, written by W. H. Saslow [20], on the temperature dependence of the average relaxation time in relaxation processes in a supercooled liquid in the glass transition region. García-Colín became interested in the temperature dependence of the Logarithmic Shift Factor (*LSF*), $\log a_T$, that may be expressed in terms of the average relaxation time τ or the viscosity of the supercooled liquid η,

$$LSF \equiv -\log a_T = \log \frac{\tau(T)}{\tau(T_s)} = \log \frac{\eta(T)}{\eta(T_s)} \tag{1.1}$$

where T_s is a reference temperature.

Ever since, Prof. García-Colín made contributions in this field, both for fragile and strong glassformers [8], for almost twenty years, as it will be presented in the following sections.

2. Theoretical basis for the Vogel-Fulcher-Tammann equation

Back in 1989, García-Colín, del Castillo and Goldstein [3] studied the dependence on temperature of the average relaxation time in supercooled fragile liquids in the vicinity of the glass transition temperature T_g.

The response function to external perturbations in a system, particularly on a supercooled fluid in the glass transition region, is given by the stretched exponential form proposed by Kohlrausch-Williams-Watts [11],

$$\phi(t) = \phi_0 \exp\left\{-(t/\tau)^{\beta}\right\}, \tag{2.1}$$

with $0 < \beta < 1$, where $\phi_0 = \phi(t = 0)$, and τ is the average relaxation time of the relevant processes that occur in the fluid.

2010 *Mathematics Subject Classification*: 01-02, 01A70.
Keywords and phrases: Glass transition, viscosity, transport processes.

An important behavior of a supercooled liquid, approaching the glass transition, is the rapid increase of its viscosity. Theoretically, many efforts had been undertaken to study the temperature dependence of the viscosity and other thermodynamic properties. The dependence of the average relaxation time τ with temperature was first described empirically by Vogel, Fulcher and Tammann, which is well known as the Vogel-Fulcher-Tammann (VFT) equation, given by,

$$\log \frac{\tau(T)}{\tau(T_s)} = A - \frac{B}{T - T_0}, \tag{2.2}$$

with the adjustable parameters A and B, and the temperature T_0 has been interpreted as the isoentropic temperature, fact that will be discussed in the next sections.

A similar empirical relation was proposed by Williams, Landel and Ferry to account the influence of the thermal history on the relaxation time. The so called Williams-Landel-Ferry (WLF) equation is given by

$$\log a_T = \frac{C_1(T - T_s)}{C_2 + T - T_s} \tag{2.3}$$

where T_s is a reference temperature and C_1 and C_2 are constants. In the original work, the WLF considers that the values of C_1 and C_2 are the same for all the substances considered.

Both the WLF and the VFT equations describe the temperature dependence of the relaxation time or the viscosity in the case of supercooled fragile liquids.

García-Colín et al. [11] derived a Vogel-Fulcher-Tammann type equation using the expression proposed by Adam and Gibbs [1] for the relaxation time. In 1965, Adam and Gibbs studied the temperature dependence of the relaxation processes in glass formers and obtained the relationship between the logarithmic shift factor and the configurational entropy of the system, the well known Adam-Gibbs equation. Both authors proposed an expression for the relaxation time in terms of the average transition probability, $W(T)$, as a function of the temperature, given by the form

$$W(T) = D \exp\left[\frac{-K}{TS_c(T)}\right] \tag{2.4}$$

where D is a constant, K is a quantity to be defined in Eq. (2.6), and S_c is the molar configurational entropy of the system. Since the relaxation time $\tau(T)$ is reciprocally proportional to $W(T)$, one may arrive to the expression,

$$\log a_T = -\log \frac{\tau(T)}{\tau(T_s)} = K\left[\frac{1}{T_s S_c(T_s)} - \frac{1}{TS_c(T)}\right] \tag{2.5}$$

T_s is an appropriately chosen reference temperature, and

$$K = 2.303 \frac{\Delta\mu s_c^*}{k_B} \tag{2.6}$$

In this expression, $\Delta\mu$ is the chemical potential, mainly the potential energy involved in the cooperative rearrangement of the segments in their model, s_c^* is a critical configurational entropy, and k_B is Botzmann's constant. The problem of obtaining the explicit form for the relaxation time in terms of the temperature was solved by García-Colín et al. calculating the configurational entropy by means of the well-known thermodynamic equation,

$$S_{c_b} - S_{c_a} = \int_{T_a}^{T_b} \left[\frac{\Delta C_{p_c}}{T} \right] dT \tag{2.7}$$

where ΔC_p is the change of the configurational specific heat in the glass transition.

In order to evaluate the configurational entropy given by Eq. (2.7), García-Colín and co-workers [11] used the expression for ΔC_p obtained by di Marzio and Dowell using the Gibbs-di Marzio microscopic model for the glass transition that evaluates the configurational entropy [12]. The expression for ΔC_p given by di Marzio and Dowell may be written as,

$$\Delta C_p = \frac{A}{T^2} + BT - CT^2 \tag{2.8}$$

The structure and values of the coefficients A, B, and C for different polymers are given in Ref. [11]. Using the form for the configurational entropy, Eq. (2.7), evaluated by means of the configurational specific heat, eq. (2.8), the authors arrive to an expression for the relaxation time in terms of the temperature given by,

$$\tau(T) = \tau_0 \exp\left[\frac{K}{F(T)(T - T_0)} \right] \tag{2.9}$$

that may also be rewritten as,

$$-\log a_T = \log \frac{\tau(T)}{\tau_0} = \left[\frac{K}{F(T)(T - T_0)} \right] \tag{2.10}$$

where T_0 is the temperature where the configurational entropy vanishes, and,

$$F(T) = \frac{A}{2T_0^2}\gamma + BT - \frac{CT^2}{2}\gamma \tag{2.11}$$

where $\gamma = 1 + (T_0/T)$.

García-Colín et al. found that the function $F(T)$ behaves practically as a constant within experimental error, thus Eq. (2.9) has the form of a VFT equation for the polymers under study.

3. The Williams-Landel-Ferry equation and the determination of the isoentropic temperature

In the case of polymers, the *LSF* may be described in terms Williams-Landel-Ferry equation [27],

$$\log a_T = \frac{C_1(T - T_s)}{C_2 + T - T_s} \tag{3.1}$$

where T_s is a reference temperature and C_1 and C_2 are constants.

(3.1) Determination of the isoentropic temperature in polymeric liquids in the glass transition. In 1993, using the results of their previous work [11], Goldstein, del Castillo and García-Colín [14] obtained an expression for the relaxation time that was compared with the WLF equation, and found the value of the isoentropic temperature.

Using Eqs. (2.5) and (2.7) , the LSF may be rewritten as,

$$\log a_T = K\left[\frac{1}{F'(T_s)(T_s-T_0)} - \frac{1}{G(T)}\frac{1}{T^2}\right] \tag{3.2}$$

with,

$$F'(T) = \left(1+\frac{T_0}{T}\right)\left(\frac{A}{2T_0^2} - \frac{CT^2}{2}\right) + BT \tag{3.3}$$

and,

$$G(T) = \left(1-\frac{T_0^2}{T^2}\right)\left(\frac{A}{2T_0^2 T} - \frac{CT}{2}\right) + B\left(1-\frac{T_0}{T}\right) \tag{3.4}$$

where K is given by Eq. (2.6), and A, B and C are the coefficients of ΔC_p given by Eq. (2.8).

On the other hand, the WLF equation may be rewritten as,

$$\log a_T = \frac{C_1(T-T_s)(T+T_s-C_2)}{1-\frac{(T_s-C_2)^2}{T^2}}\frac{1}{T^2} \tag{3.5}$$

The form for the LSF given by Eq. (3.2) depends strongly on the isoentropic temperature T_0. The authors compared both Eqs. (3.2) and (3.5), and, from this comparison, they found the isoentropic temperature for five polymers. It is important to point out that these values are similar to those reported in other works, Eq. (3.5) may be represented as a linear relation of $(T_0/T)^2$, that is,

$$LSF = m\left(\frac{T_0}{T}\right)^2 + b(T_s) \tag{3.6}$$

where the intercept b depends on T_s, the reference temperature, and the slope m is independent of this temperature.

(3.2) On the generalization of the Williams-Landel Ferry equation. In 1998, Dagdug and García-Colín [4] extended the results of Ref. [14] and calculated the isoentropic temperature T_0 for non polymeric glass formers using the generalization of the WLF equation proposed by Adam and Gibbs. This equation may be rewritten as

$$\log a_T = \frac{a_1'(T-T_s)}{a_2'+(T-T_s)}, \tag{3.7}$$

where

$$a_1' = \frac{(2.303)^{-1}}{T_s S_c(T_s)}\left(\frac{\Delta\mu s_c^*}{k_B}\right) \tag{3.8}$$

and

$$a_2'(T) = \frac{T_s S_c(T_s)}{S_c(T_s) + \frac{T(S_c(T)-S_c(T_s))}{T-T_s}}. \tag{3.9}$$

Dagdug and García-Colín compute T_0 setting $T_s = T_g$, $T = T_0$, and using the fact that the coefficient $a_2^{'}$ must be equal to C_2, where $C_2 = T_s - T_0$. Thus,

$$\frac{T_g}{T_0} = 1 + \frac{C_2}{T_g - C_2}. \tag{3.10}$$

Furthermore, comparing $a_1^{'}$ with the coefficient C_1 of the original WLF equation, $a_1^{'} = C_1 = 8.86$ [6], they obtain the relation,

$$\frac{\Delta\mu s_c^*}{k_B} = 2.303 C_1 T_g S_c(T_g). \tag{3.11}$$

The authors find, on one hand, the value of the critical configurational entropy s_c^* defined by Adam and Gibbs, given in Eq. (2.6) in terms of the configurational entropy evaluated in T_g, namely,

$$s_c^* = \frac{T_0}{C_2} S_c(T_g). \tag{3.12}$$

On the other hand, they were able to write an expression to evaluate the free energy $\Delta\mu$ given by

$$\Delta\mu = 2.303 C_1 C_2 k_B \frac{T_g}{T_0}. \tag{3.13}$$

They report both the values for s_c^* and $\Delta\mu$ for six polymeric glassformers.

4. On the Stokes-Einstein relation in glass forming liquids

In 2000, Goldstein, García-Colín and del Castillo [15] studied the relationship between the diffusion coefficient of a tracer and the viscosity of a fragile glass-former liquid where the diffusion process takes place. The relaxation phenomena described by the VFT or WLF equations correspond to the very slow α-relaxation processes. One may find, however, that fast relaxation processes occur in the vicinity of T_g, namely the β-relaxation processes. Relaxation and diffusion mechanisms present drastic changes around a cross-over temperature T_c which lies within the interval $[1.15T_g, 1.28T_g]$. There are two important aspects that characterize this cross-over region. Both the VFT and the WLF empirical equations do no longer describe the experimental results for the viscosity below T_c, and, furthermore, the diffusion mechanisms undergo changes.

One of the most significant features of a supercooled liquid approaching the glass transition is the rapid increase of the viscosity.

As it has been discussed in the previous sections, one of the most important empirical equations, that deals with the behavior of the viscosity as the system approaches T_g, is the VFT equation, eq. (2.2).

The Stokes-Einstein (SE) equation establishes that the diffusion coefficient of a sphere of radius a moving in a fluid whose viscosity is η, is given by

$$D = \frac{k_B T}{6\pi a \eta} \tag{4.1}$$

where k_B is Boltzmann's constant. In the case of a glassformer at temperatures above T_c, the SE equation works. However, for temperatures below T_c, the SE equation breaks down, and the diffusion process is enhanced. In this region, the influence of the viscous relaxation upon the diffusion coefficient may be expressed in terms of the relation

$$D \sim \eta^{-\xi} \tag{4.2}$$

where $0 < \xi < 1$.

In fact, both experimental and theoretical results have indicated that as a supercooled glass forming liquid is cooled towards T_g, its dynamics becomes increasingly heterogeneous presenting magnified diffusion mechanisms.

In this work, they presented a form for the viscosity for temperatures below T_c, and obtained the value of the exponent ξ for temperatures below and above the cross-over temperature.

They examined the experimental results for the viscosity of three fragile glass-former liquids, salol, phenolphthaleine-dimethyl-ether (PDE) and orthoterphenil (OTP), as a function of temperature. They used the method proposed by Stickel et al. [23, 24], in terms of the time derivative analysis for the quantities,

$$x = \{f/Hz, \sigma_{dc}\varepsilon_0/s^{-1}, \text{poise}/\eta\} \tag{4.3}$$

where f is the frequency of the peak of the loss function, the imaginary part of the complex dielectric function, σ_{dc} is the dc conductivity, and η the viscosity. The method consists on the evaluation of three derivatives of a given empirical form for $\log x$, namely,

$$\left[\frac{d\log x}{dT}\right]^{-1/2}, \tag{4.4}$$

$$\frac{d}{dT}\left[\left(\frac{d\log x}{dT}\right)^{-1/2}\right], \tag{4.5}$$

$$\Theta = \frac{\frac{d\log x}{dT}}{\frac{d^2\log x}{dT^2}}. \tag{4.6}$$

In the case that $\log\eta$ is given by the VFT equation, the derivatives are given by

$$\left[\frac{d\log\eta_{VFT}}{dT}\right]^{-1/2} = B^{-1/2}(T - T_0), \tag{4.7}$$

$$\frac{d}{dT}\left[\left(\frac{d\log\eta_{VFT}}{dT}\right)^{-1/2}\right] = B^{-1/2}, \tag{4.8}$$

$$\Theta_{VFT} = -\frac{T - T_0}{2}. \tag{4.9}$$

Through the analysis of the experimental values for the viscosity, they found that for temperatures above T_c, the derivatives (4.7) and (4.8) are valid. Nevertheless, for temperatures below T_c, the values of these derivatives are not satisfied. Goldstein et al. propose a form for the viscosity, whose derivatives satisfy the experimental values given by,

$$\log\frac{\eta(T)}{\eta(T_s)} = A'(T_a - T)^2 + B'. \tag{4.10}$$

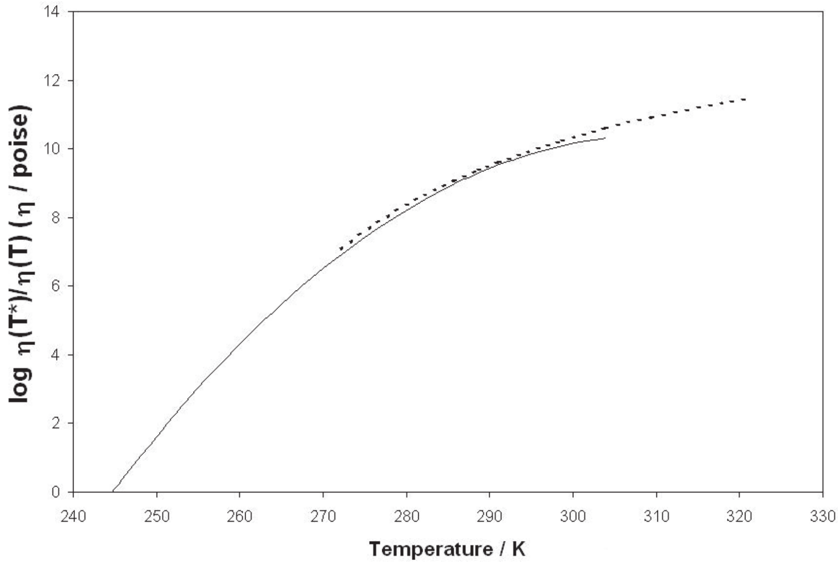

Figure 1. The viscosity of OTP [15]. The dashed line corresponds to the expression for the viscosity above T_c given by Eq. (4.11), namely the VFT equation. The full line represents the viscosity obtained using eq. (4.10) below T_c.

For temperatures above the crossover temperature, the viscosity is well described by a VFT form,

$$\log \frac{\eta(T)}{\eta(T_s)} = C' - \frac{D'}{T - T_0} \tag{4.11}$$

where the values of A' ,B' C',D' and T_a, T_0 are reported in Ref. [15].

On the other hand, the value of T_c may be evaluated using the derivative Θ, given in eq. (4.9), evaluated for the expressions given by eqs. (4.10) and (4.11). Both values of Θ intersect each other at a temperature T_c'

$$\Theta_{T<T_c}(T_c') = \Theta_{T>T_c}(T_c')$$

that is in a very good agreement with the experimental values of T_c.

In Figure 1 one may see the dependence on temperature of the viscosity of OTP, where for $T > T_c$, eq. (4.11) is valid, while for $T < T_c$, the viscosity is given by eq. (4.10).

Finally, they studied the dependence of the diffusion coefficient in terms of the viscosity and found that the Stokes-Einstein relationship broke down for $T < T_c$, eq. (4.2). In the case of temperature above T_c, the exponent ξ is nearly one. On the other side, for temperatures below T_c, ξ is less than one. As an example, in Figure 2, one may see different tracers in OTP and the values of ξ, which are the values of the slopes of each straight line.

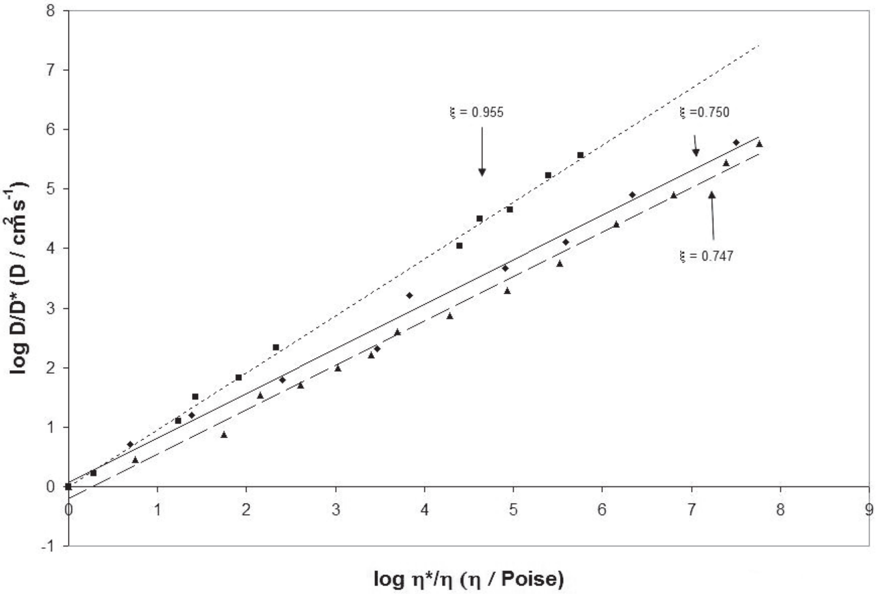

Figure 2. Diffusion coefficients for tracers in OTP in terms of the viscosity. Experimental values for TTI (♦), rubrene (■) and tetracene (▲) [15]. The full, dotted and dashed lines represent respectively the linear fits for TTI, rubrene and tetracene, respectively, where ξ is the slope in each case.

5. Strong glasses, the relaxation time or viscosity

Between 1998 and 2000 García-Colín and Dagdug derived a theoretical VFT equation for the viscosity of strong glasses inspired in a previous works by R. A. Barrio, R. Kerner, M. Micoulat and G. G. Naumis [16, 2]. In this section we shall discuss this results using as example the strong glass forming B_2O_3, [5, 6], and for covalent networks [7]. The main idea of this work is to take the average relaxation time (or viscosity) as inversely proportional to the average transition probability using a Markov chain.

(5.1) Boroxol. Through a Markov chain it is possible to describe the growth process of a solid [16]. With this method such process can be described by a matrix acting on a vector. The matrix components are the probabilities of finding a given site at the border (rim) of a glass cluster of a certain size. The vector components represent the probabilities of finding a given site on the rim of such cluster. The matrix transforms this vector onto a new one after adding one new unit to the cluster. The transformation of the rim depends on the site on which the new unit sticks to. Each sticking process has a certain probability of occurring, in this aim the matrix elements contain the probabilities of transforming each kind of site into others. The probability factors must include two contributions: (i) The statistical weight for each process, that is the number of ways leading to the same final result, and (ii) the Boltzmann factor taking into account the energy barrier necessary to form a certain kind of bond.

The elementary unit, dictated by the bond chemistry, is a triangle $B(O_{1/2})_3$, a *singlet*. Two singlets can be connected only using one bond to form a *doublet*. The energy cost to form this bonding es E_1. After a doublet is produced, two situations can occur if a new singlet is added: the newly arriving singlet forms a longer chain (a *triplet*) or it can close a ring with an energetic cost E_2. The agglomeration process occurs at a given temperature T, at which the individual bonds reach equilibrium. With this idea in mind one can write the matrix $\mathbf{M}$, modeling the growth of clusters by a successive application on an arbitrary initial vector $\mathbf{v}_0$. Thus the evolution of the probabilities on the rim after j steps is given by $\mathbf{v}_j = \mathbf{M}^j \mathbf{v}_0$. One also can derive an expression for the probability of forming a ring, P'_B, before many steps, obtained by counting the proportion of rings that were formed during the precess.

The final configuration depends only on the eigenvectors of the stochastic matrix. It is easy to prove that a matrix with all the columns normalized to one has at least one eigenvalue equal to one, while the others can be real, complex or imaginary, depending on the values of the parameters involved. Only eigenvectors with norm one remain after many successive applications of the stochastic matrix. If one assumes that $\mathbf{M}$ has only one such eigenvalue ($\lambda = 1$), with eigenvector $\mathbf{e}_1$, then, in the limit of large j, $\mathbf{v}_j$ converges to this eigenvector, independently of the initial condition.

As a consequence, the evolution of the rim attains a stable statistical regime after successive steps of growth and this regime is governed solely by the statistics represented by the eigenvector with eigenvalue one. Barrio *et al* found for the B_2O_3 that the eigenvector is given by [2],

$$e_1 = \frac{1}{84\xi^2 + 107\xi + 25} \begin{pmatrix} 1 + 4\xi \\ 24\xi^2 + 34\xi + 9 \\ 24\xi^2 + 34\xi + 10 \\ 12\xi + 15 \\ 3\xi(4\xi + 3) \\ 2\xi(12\xi + 7) \end{pmatrix}, \tag{5.1}$$

where,

$$\xi = \exp\left[\frac{\triangle E}{k_B(T - T')}\right], \tag{5.2}$$

and $\triangle E = E_1 - E_2$. In this model the only free parameter is ξ, the excess free energy used when closing a ring.

Up to now, the probability of sticking a new unit in the bulk at any temperature T is taken to be proportional to $\exp(-E_i/k_B T)$, where E_i is the energy cost of sticking a unit in the i form at temperature T. To generalize this results we use the fact that below T_g the glass system is unable to displace any unit to stick in to the bulk, and because of that a temperature T' is introduced such that the probability to stick a unit to the bulk is equal to zero. Thus, the probability of sticking anew unit in the bulk may be generalized to $\exp(-E_i/k_B(T - T'))$. To identify T' as a physical property of the system, the relaxation time (or viscosity) for the growth of the system is calculated as inversely proportional to the probability of forming a ring, P'_B, before many steps, obtained by counting the proportion of rings that were formed during the process.

$$\tau \propto 1/P(\xi). \tag{5.3}$$

Particularly for the B_2O_3 the probability of forming a ring when passing form the jth layer to the $(j+1)$th one, is simply given by counting the proportion of rings that were formed between the step j and the step $j+1$. If it is calculated for a large number of steps of growth, P'_B can be replaced by its limiting value which according to Barrio *et al* for B_2O_3, from equation (5.1), is given by [2],

$$P'_B = \frac{24\xi^2 + 16\xi}{84\xi^2 + 107\xi + 25}. \tag{5.4}$$

As our main hypothesis we take the transition probability as the probability of forming a ring for the B_2O_3.

Taking the derivatives $d^n \log x/dx^n$, and $d \log x/d(1/T)$, for $n = 1,2$ of Eq. (5.4) where $x = \tau$, we have,

$$\left[\frac{d \ln P'_B}{dT}\right]^{-\frac{1}{2}} = (T - T_0)\left[\frac{E_2 - E_1}{k_B}\right]^{-\frac{1}{2}} L_{B_2O_3}^{-\frac{1}{2}}, \tag{5.5}$$

$$\left[\frac{d \ln P'_B}{dT^{-1}}\right]^{-\frac{1}{2}} = \left(1 - \frac{T}{T_0}\right)\left[\frac{E_2 - E_1}{k_B}\right]^{-\frac{1}{2}} L_{B_2O_3}^{-\frac{1}{2}}, \tag{5.6}$$

where,

$$L_{B_2O_3} \equiv \frac{48\xi^2 + 16\xi}{24\xi^2 + 16\xi} - \frac{168\xi^2 + 107\xi}{84\xi^2 + 107\xi + 25}. \tag{5.7}$$

Taking typical values fort he activation energy as quoted in reference [24], we know that for $\tau \ll 1$ and $L_{B_2O_3} \approx 1$, the temperature dependence of $L_{B_2O_3}$ can be neglected so that Eqs. (5.5) and (5.6) can be written as:

$$\left[\frac{d \ln P'_B}{dT}\right]^{-\frac{1}{2}} = (T - T_0)\left[\frac{E_2 - E_1}{k_B}\right]^{-\frac{1}{2}}, \tag{5.8}$$

$$\left[\frac{d \ln P'_B}{dT^{-1}}\right]^{-\frac{1}{2}} = \left(1 - \frac{T}{T_0}\right)\left[\frac{E_2 - E_1}{k_B}\right]^{-\frac{1}{2}}. \tag{5.9}$$

From equation (5.4) we can also calculate $\Theta \equiv [d\ln(x)/dT][d^2\ln(x)/dT^2]$ and if $L_{B_2O_3}$ is a constant we get that,

$$\Theta = -(T - T_0)/2. \tag{5.10}$$

If equations (5.8) - (5.10) are integrated, the method of the temperature derivative, a theoretical VFT-like equation is obtained, namely,

$$\tau = \tau_0 \exp\left[\frac{E_2 - E_1}{k_B(T - T_0)}\right] = \tau_0 \exp\left[\frac{DT_0}{k_B(T - T_0)}\right]. \tag{5.11}$$

Here T' can be identified as T_0, because when T_0 goes to zero, Eq. (5.11) reproduce the Arrhenius equation. It is important to note that in this theoretical context, T_0 is interpreted not only as the temperature that yields an infinite relaxation time, but it also is the temperature at which the probability of sticking a unit into the bulk of the glass system is zero. The constant τ_0 is the preexponential factor and D is a constant equal to $(E_2 - E_1)/k_B T_0$ which can be determined comparing it with the experimental VFT equation. The experimental values are given by $D \approx 35$, $T_0 \approx T_g/2$ and $\tau(T_g) = 10^{13}P$ [21]. Using this values the activation energy turns out to be $E_2 - E_1 = 18.207$ kcalmol^{-1}, which is in excellent agreement with the experimental values [3].

(5.2) Covalent Network. In this subsection we shall follow the same steps as the previous subsection to obtain the VFT-like equation for covalent network glass systems. To this end we shall proceed as follows. First we identify the average transition probability with the average transition probability of forming some kind of preferential link, the link that gives the largest probability of occurring in the glass transition. In fact, only the weakest bonds are broken or rearranged initially in the glass transition region [9]. Then we take the relaxation time (or viscosity) as inversely proportional to the form of the relaxation time (or viscosity) as inversely proportional to the transition probability. Finally we apply the method of the temperature derivative to obtain the form of the relaxation (or viscosity).

To illustrate the method, with out loss of generality, we take as an example a covalent network system from atoms of valences two and three: the As_xSe_{1-x} system for example. The main property derived from the experimental relation between the composition and the transition temperature is: if the Se content increases, the transition temperature decreases [22, 10] and, adding the fact that x-ray and neutron scattering measurements as well as studies of the infrared and Raman spectra have shown that the short-range order in the glass-forming system As_xSe_{1-x} is given by chain-like connected Se atoms and the structural units $AsSe_{3/2}$ in the given range of composition [18], we consider the transition probability to be directly proportional to form SeSe bonds. Further, we are also assuming that the Se chains determine T_g [18]-[26].

The probability of forming the weakest bond when passing from the jth layer to the $(j+1)$th one is simply given by counting the proportion of Se atoms formed between the step j and the step $j+1$ linked by another Se. If for a large number of steps of growth we calculate the probability of forming the weakest bond, we find that it is given by [7],

$$P_{Se-Se} = \frac{8(1-x)^2}{4(1-x)[2(1-x)+3x\zeta]+9x[2(1-x)+3x\mu]}, \tag{5.12}$$

where $\zeta = \exp((E_1 - E_2)/k_B T)$ and $\mu = \exp((E_1 - E_3)/k_B T)$ and E_1, E_2, and E_3 the corresponding energy barriers for the energetic unions of Se-Se, Se-As, and As-As, respectively.

Moreover, since the viscosity is inversely proportional to the average transition probability, $\eta \propto 1/P_{Se-Se}$, taking the derivative $d\ln\eta^{-1}/dT$, we find that,

$$\left[\frac{d\ln P_{Se-Se}}{dT^{-1}}\right]^{-\frac{1}{2}} = (T - T_0)\left[\frac{3x}{k_B}\frac{4(1-x)(E_2-E_1)\zeta + 9(E_3-E_1)\mu}{4(1-x)[2(1-x)+3x\zeta]+9x[2(1-x)+3x\mu}\right]^{-\frac{1}{2}}. \tag{5.13}$$

Equation (5.13) is a non trivial equation that depends on the activation energies, and remarkable, predicts theoretically that the viscosity should be a function of the concentration. Experimentally it is observed that energy differences involved in equation (5.13) are nearly zero or at most ~ 10 kcal mol [19]. This result gives us four situations: $|E_2 - E_1| \gg |E_3 - E_1|$, $|E_2 - E_1| \ll |E_3 - E_1|$, $|E_2 - E_1| \approx |E_3 - E_1| \ll 1$, and $|E_2 - E_1| \approx |E_3 - E_1| \gg 1$. We analyze each case individually.

5.2.1. *Case I,* $|E_2 - E_1| \gg |E_3 - E_1|$. With this approximation, $\zeta \to 0$ and $\mu \to 1$, and equation (5.13) can be written as,

$$\left[\frac{d \ln P_{Se-Se}}{dT^{-1}}\right]^{-\frac{1}{2}} = (T - T_0)\left[\frac{(E_3 - E_1)}{k_B} \frac{27x^2}{8(1-x)^2 + 9x(x+2)}\right]^{-\frac{1}{2}}. \tag{5.14}$$

If equation (5.14) is integrated, a theoretical VFT-like equation is obtained, namely,

$$\eta = \eta_0 \exp\left[\frac{(E_3 - E_1)}{k_B(T - T_0)} \frac{27x^2}{8(1-x)^2 + 9x(x+2)}\right] = \exp\left[\frac{D^* T_0}{(T - T_0)}\right]. \tag{5.15}$$

where the constant η_0 is the pre-exponential factor and may be obtained from a plot of η against $1/T$. In this equation, $D^*(x)$ is not a constant but depends on the concentration:

$$D^* = \frac{27x^2}{8(1-x)^2 + 9x(x+2)} D. \tag{5.16}$$

Equation (5.16) predicts that for these systems, fragility (inversely proportional to D) decreases when the concentration x increases, see Figure 3.

5.2.2. *Case II,* $|E_2 - E_1| \ll |E_3 - E_1|$. A typical material with these characteristics of its activation energies is As_xSe_{1-x} , and in this approximation one has $\zeta \to 1$ and $\mu \to 0$, so equation (5.13) can be written as,

$$\left[\frac{d \ln \eta^{-1}}{dT}\right]^{-\frac{1}{2}} = (T - T_0)\left[\frac{(E_3 - E_1)}{k_B} \frac{12x}{22x + 8}\right]^{-\frac{1}{2}}. \tag{5.17}$$

If equation (5.17) is integrated, a theoretical VFT-like equation is obtained, as in the preceding cases,

$$\eta = \eta_0 \exp\left[\frac{(E_3 - E_1)}{k_B(T - T_0)} \frac{3x}{x+2}\right] = \exp\left[\frac{D^* T_0}{(T - T_0)}\right] \tag{5.18}$$

and

$$D^* = \frac{3x}{x+8} D. \tag{5.19}$$

Equation (5.19) predicts that the largest value for D^* occurs at x = 1, see Figure 3.

5.2.3. *Case III,* $|E_2 - E_1| \approx |E_3 - E_1| \ll 1$. With this approximation for the energies, $\zeta \to 1$ and $\mu \to 1$, equation (5.13) can be written as,

$$\left[\frac{d \ln \eta^{-1}}{dT}\right]^{-\frac{1}{2}} = (T - T_0)\left[\frac{(E_3 - E_1)}{k_B} \frac{3x}{x+2}\right]^{-\frac{1}{2}}. \tag{5.20}$$

If equation (5.17) is integrated, a theoretical VFT-like equation is obtained,

$$\eta = \eta_0 \exp\left[\frac{(E_3 - E_1)}{k_B(T - T_0)} \frac{12x}{22x+8}\right] = \exp\left[\frac{D^* T_0}{(T - T_0)}\right]. \tag{5.21}$$

and as in the last case this equation depends on the concentration, and now,

$$D^* = \frac{12x}{22x+8} D. \tag{5.22}$$

Equation (5.22) predicts that if x increases, the fragility decreases, and as in equations (5.16) and (5.19) the largest value for D^* occurs at x = 1, see Figure 3. Also in Figure 3 one can see that cases I and III are more fragile than case II.

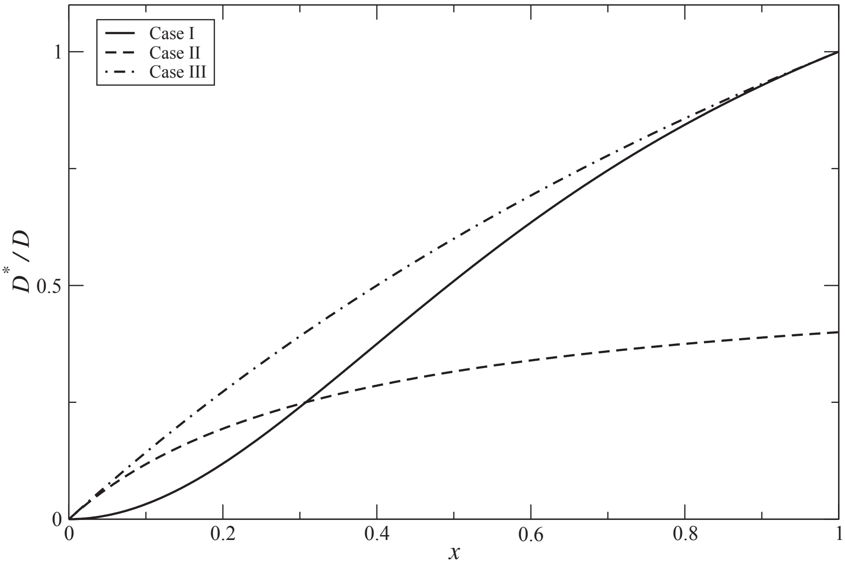

Figure 3. D^* as a function of x from equations (5.16).

5.2.4. *Case IV,* $|E_2 - E_1| \approx |E_3 - E_1| \gg 1$. When $\zeta \to 0$ and $\mu \to 0$ equation (5.17) is equal to zero and predicts that the viscosity is independent of temperature, which it is well known is not a physical solution for vitreous systems. In fact, our model predicts that it is not possible to form an amorphous system with this physical properties.

6. Contribution of floppy modes to configurational and excess entropy

In this section we shall show how García-Colín, Goldstein and Dagdug, following Naumi's ideas included the floppy modes as a free energy in order to obtain the configurationl and excess entropy as well as the jump of the heat capacity of the chalcogenide glasses as function of the coordination number $\langle r \rangle$.

As we discussed in the introduction one of the most successful theoretical efforts to deal with the description os structural relaxation processes in supercooled liquids was developed by Adams and Gibbs [1]. In their theory the mean structural viscosity η depends both on the temperature and on the configurational entropy S_c of the liquid or glass, namely,

$$\eta = \eta_0 \exp\left[\frac{K}{TS_c}\right] \tag{6.1}$$

where η_0 and K are constants, the last is related with the activation energy given by $K = \triangle\mu s_c^*/k_B$. Here $\triangle\mu$ is the energy barrier per particle opposing a cooperative rearrangement of the liquid/glass structure, and s_c^* is the critical configurational entropy.

In Eq. (6.1), it has often been assumed that the configurational entropy, S_c, can be approximated by the difference in entropy between liquid and crystal, the excess entropy (S_{ex}), but it has been known since Goldstein'a analysis of entropy change in glasses between T_g and 0 K that this is a rather poor approximation in many cases [13]. Experimentalist have tested Eq. (6.1) successfully for temperatures not too fa from T_g, using the excess entropy of Kauzmannńs plot in place of S_c. Even the two quantities are not the same, namely S_{ex} and S_c, Matínez and Angell pointed out that Eq. (6.1) would be valid for both if a proportionality constant between S_{ex} and S_c exist [17].

In this work accomplished between 2005 and 2006 we show theoretically that the proportionality constant, $A = S_{ex}/Sc$, exist and ranges from 1.5 for strong liquids, $\langle r \rangle = 2$, to 2 for fragile ones, $\langle r \rangle = 2.4$. To this end, we include the floppy modes as a free energy to calculate the internal energy. The knowledge of the internal energy allow us to calculate S_c. Moreover, using S_c in the Adam-Gibbs equation in the Adam-Gibbs equation we obtain a VFT-like equation. This VFT-like equation let us to understand why D ia larger for strong glasses than for fragile ones.

(6.1) The internal energy. We start by evaluating the internal energy of the liquid melt as the sum of various contributions,

$$U_{conf} = U_{nl} + U_{vib} + U_{trans} + U_f. \tag{6.2}$$

Using the equipartition theorem

$$U_{trans} = \frac{3}{2} N k_B T. \tag{6.3}$$

The non-linear, U_{nl}, contribution is due to small deviations from linearity in a polyatomic molecules, which may generate rotations as well as slight shifts of the atoms from their equilibrium position and may be express as,

$$U_{nl} = N k_B T. \tag{6.4}$$

Following Naumis, floppy energy is taken to be proportional to the floppy modes [1], where these modes are the number of possible independent deformations in the network, the zero frequency modes. According to the equipartition theorem the floppy energy is given by,

$$U_f = 3 f N k_B T \tag{6.5}$$

where $f = 2 - 5\langle r \rangle/6$ is the fraction of floppy modes available in the network.

Because in a glass system the translational degrees of freedom are frozen, using Eqs. (6.3)-(6.3) into Eq. (6.2) we can see that the configurational energy is given by,

$$U_{conf} = \left(7 - \frac{5\langle r \rangle}{2} N k_B T\right). \tag{6.6}$$

From equation (6.6) we can obtain the configurational specific heat and entropy, namely,

$$\triangle C p_c = \left(7 - \frac{5\langle r \rangle}{2}\right) N k_B \tag{6.7}$$

and,

$$S_c = \left(7 - \frac{5\langle r \rangle}{2}\right) N k_B \ln \frac{T}{T_K} \tag{6.8}$$

where $S_c = S_{ex} - S_{vib}$.

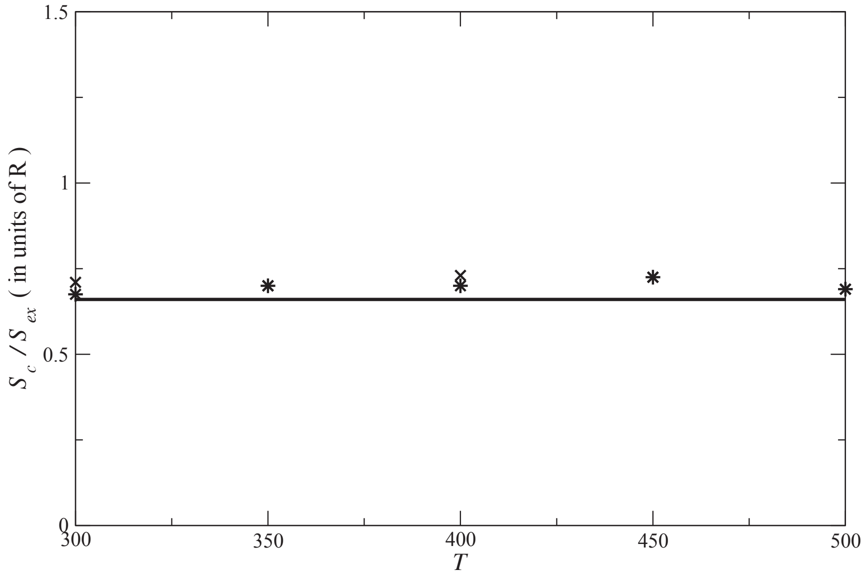

Figure 4. S_c/S_{ex} (continues line) calculated by Eqs.(6.8) and (6.12) as a function of T. Experimental data (circles) taken from Ref. [25]

(6.2) The viscocity. Introducing the configurational entropy given in Eq. (6.8) into the Adam-Gibbs equation, Eq. (6.1), we are able to write an expression for the viscosity. To this end we may expand the logarithm in Eq. (6.8) taking into account that $T/T_K = x > 1/2$,

$$\ln x = \frac{x-1}{x} + \frac{1}{2}\left(\frac{x-1}{x}\right)^2 + \frac{1}{3}\left(\frac{x-1}{x}\right)^3 + \cdots \tag{6.9}$$

Keeping only the first term in the expansion and replacing the Kauzmann temperature T_K by T_0 (equality well known from the experiment [25]), a VFT-like equation can be obtained,

$$\eta = \eta_0 \exp\left[\frac{DT_0}{T - T_0}\right], \tag{6.10}$$

where,

$$D = \frac{C}{\left(7 - \frac{5\langle r\rangle}{2} N k_B T_0\right)}. \tag{6.11}$$

In this last equation, the viscosity depends explicitly on the average coordination number $\langle r\rangle$ and predicts that when $\langle r\rangle$ increase D decrease. as it is well known from the experiment [25].

(6.3) The excess entropy. In order to calculate the excess entropy we simply add the vibrational contribution to the configurational entropy. We shall perform this calculation for selenium, one of the best known chalcogenide glass formers. For Se the experimentalist found that the vibrational contribution is one third of the excess entropy, $S_{vib} = (1/3)S_{ex}$ [26]. Also for this element, $\langle r\rangle = 2$, using this

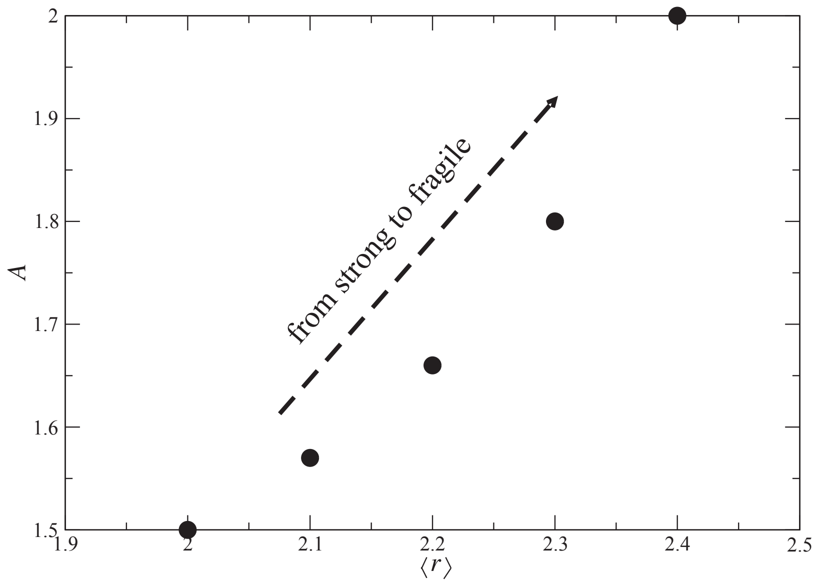

Figure 5. A as a function of $\langle r\rangle$ from equation (6.13).

into Eq. (6.8), we found that $S_c = 2k_B \ln(T/T_0)$. Therefore, the vibrational term can be approximated by $S_{vib} = Nk_B \ln(T/T_0)$. Thus, the excess entropy reads,

$$S_{ex} = \left(8 - \frac{5\langle r\rangle}{2}\right) Nk_B \ln \frac{T}{T_K}. \tag{6.12}$$

The predicted ration S_c/S_{ex}, given by Eqs. (6.8) and (6.12), compared with the experimental data are shown in Figure 4.

From Eqs. (6.8) and (6.12) we find that the ratio A is given by,

$$A = \frac{16 - 5\langle r\rangle}{14 - 5\langle r\rangle}. \tag{6.13}$$

In Figure 5 is shown the dependence of A on the average coordination number.

7. Challenges and open questions in the glass transition

The glassy state is ubiquitous in nature and technology and has been considered as one of the deepest and most interesting challenges in physics, understanding quantitatively the extraordinary viscous slow-down that accompanies supercooling and glass formation is a major scientific interrogation. It is crucial in the processing of foods, protein function and naturalization, the commercial stabilization of labile biochemicals, and the preservation of insect life under extremes of cold or dehydration. In medicine, polymer-based materials are commonly used as excipients of poorly water-soluble drugs, the success of the encapsulation, as well as the physicochemical stability of the products, is often reflected on their glass transition temperature. Dry products obtained from most of the common drying processes are predominantly in a glassy amorphous form. Optical fibres are made of very pure amorphous silica deliberately doped. Window glass, composed mostly of sand, lime and soda, is the best-known example of an engineered amorphous solid. Most engineering plastics are amorphoussolids, as are some metallic glasses and alloys of interest because of their soft magnetism and corrosion resistance. Moreover, it is possible that most water in the Universe may be glassy.

Most of the available information on the experimental data for a supercooled liquid, that ultimately becomes a glass, still remains in a phenomenological framework. Calorimetric and transport properties are reported for a large number of glass formers, and physicists try to do their best to reproduce the behavior of these liquids by means of statistical physics models or computer simulations. Existing theories seem incapable of explaining the most basic questions. A great challenge appears ahead in the description of the glass transition: the contribution of mathematical models to describe a supercooled liquid, with such a large viscosity, that it turns out to become a "frozen" liquid characterized by a topological disorder and a total absence of translational symmetry. The authors invite mathematicians to work in this field to try to find together an insight of a formal theory, both physically and mathematically valid; to arrive to a new vision to solve this problem that has been present in nature since the first glass was produced in ancient civilizations.

8. Complete works of Leopoldo García-Colín Scherer on glass transition

1. L. F. Del Castillo and L. S. García-Colín, *Thermodynamic basis for dielectric relaxation in complex materials*, Phys. Rev. B, **33** (1988), 4944–4951.
2. L. F. Del Castillo and L. S. García-Colín, *Dielectric relaxation thermodynamics: Comparison with experiment*, Phys. Rev. B, **47** (1988), 448–453.
3. L. S. García-Colín, L. F. Del Castillo, and P. Goldstein, *Dependencia del relajamiento en materiales amorfos con la temperatura*, Proceedings IV Congreso Nal. de Polímeros, (1988), 123–129.
4. L. S. García-Colín, *La transición vítrea*, Contactos, **4** (1989), 39–47.
5. L. S. García-Colín, L. F. Del Castillo, and P. Goldstein, *A Theoretical basis for the Vogel-Fulcher-Tammann equation*, Phys. Rev. B, **40** (1989), 7040–7044.
6. L. S. García-Colín and Rosalío Rodríguez Zepeda, Líquidos exóticos, Colección la ciencias desde México 104, Capítulo IV, México, (1991), 72–82.
7. P. Goldstein, L. F. Del Castillo, and L. S. García-Colín, *Determination of the isoentropic temperature in the glass transition*, Macromolecules, **26** (1993), 655–658.
8. L. S. García-Colín, *Avances recientes en la transición vítrea*, Memoria del El Colegio Nacional, (1997), 93–113.
9. L. Dagdug and L. S. García-Colín, *Generalization of the Williams-Landel-Ferry equation*, Physica A, **250** (1998), 133–141.
10. L. S. García-Colín, *Remarks on the glass transition*, Revista Mexicana de Física, **45** sup. I, (1999), 11–17.
11. L. Dagdug and L. S. García-Colín, *Theoretical framework for the Arrhenius equation in strong glasses*, J. Phys.: Condens. Matter, **11** (1999), 2193–2198.
12. L. Dagdug and L. S. García-Colín, *Generalization of the stochastic matrix method and a theoretical framework for the VFT equation in strong glasses*, J. Phys.: Condens. Matter, **11** (1999), 4575–4582.

13. P. Goldstein, L. F. Del Castillo, and L. S. García-Colín, *On the Stokes-Einstein relations in glass forming liquids*, Physica A, **275** (2000), 325–335.
14. L. S. García-Colín and P. Goldstein, La física de los procesos irreversibles: tomo 2, El Colegio Nacional, Capítulo XIX, México, (2003), 539-581.
15. L. Dagdug, L. S. García-Colín, and P. Goldstein, *An entropy based theory for the viscosity of strong glasses*, Revista Mexicana de Física, **51** (2005), 144–148.
16. L. Dagdug, L. S. García-Colín, and P. Goldstein, *Contribution of floppy modes to configurational and excess entropy in chalcogenide glasses*, J. of Non-Crystaline Solids, **352** (2006), 5399–5402.

Received January 6, 2013

Final version received February 23, 2013

PATRICIA GOLDSTEIN
DEPARTAMENTO DE FÍSICA, FACULTAD DE CIENCIAS,
UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO,
CIUDAD UNIVERSITARIA, 04510,
MÉXICO, D. F. MÉXICO
patricia.goldstein@ciencias.unam.mx

LEONARDO DAGDUG
DEPARTAMENTO DE FÍSICA,
UNIVERSIDAD AUTÓNOMA METROPOLITANA-IZTAPALAPA,
09340, MÉXICO, D. F. MÉXICO
dll@xanum.uam.mx

REFERENCES

[1] G. ADAM AND J. H. GIBBS, *On the Temperature Dependence of Cooperative Relaxation Properties in Glass-Forming Liquids*, J. Chem. Phys. **43** (1965), 139–146.
[2] R. A. BARRIO, R. KERNER, M. MICOULAT, AND G. G. NAUMIS, *Evaluation of the concentration of boroxol rings in vitreous B_2O_3 by the stochastic matrix method*, J. Phys.: Condens. Matter **9** (1997), 9219–9234.
[3] N. P. BANSAL AND R. H. DOREMUS, Handbook of glass properties, Academic Press, 1986.
[4] L. DAGDUG AND L. S. GARCÍA-COLÍN, *Generalization of the Williams-Landel-Ferry equation*, Physica A **250** (1998), 133–141.
[5] L. DAGDUG AND L. S. GARCÍA-COLÍN, *Theoretical framework for the Arrhenius equation in strong glasses*, J. Phys.: Condens. Matter **11** (1999), 2193–2198.
[6] L. DAGDUG AND L. S. GARCÍA-COLÍN, *Generalization of the stochastic matrix method and a theoretical framework for the VFT equation in strong glasses*, J. Phys.: Condens. Matter **11** (1999), 4575–4582.
[7] L. DAGDUG, *A theoretical framework for the Vogel-Fulcher-Tammann equation for covalent network glasses derived by the stochastic matrix method*, J. Phys.: Condens. Matter **12** (2000), 9573–9589.
[8] S. R. ELLIOTT, Physics of amorphous materials (2nd edition), Longman, Scientific & Technical, New York, 1990.
[9] A. FELTZ, Amorphous Inorganic Materials and Glasess, New York VCH, 1993.
[10] A. FELTZ, H. AUST, AND A. BLAYER, *Glass formation and properties of chalcogenide systems XXVI: Permittivity and the structure of glasses As_xSe_1-x and Ge_xSe1_1-x*, J. Non-Cryst. Solids **55** (1983), 179–190.
[11] L. S. GARCÍA-COLÍN, L. F. DEL CASTILLO, AND P. GOLDSTEIN, *A theoretical basis for the Vogel-Fulcher-Tammann equation*, Phys. Rev. B **40** (1989), 7040–7044.

[12] J. H. GIBBS AND E. A. DI MARZIO, *Nature of the glass transition and the glassy state*, J. Chem. Phys. **28** (1958), 373–383.

[13] M. GOLDSTEIN, *Liquid water as a lone-pair amorphous semiconductor*, J. Chem. Phys. **64** (1976), 1549–1554.

[14] P. GOLDSTEIN, L. F. DEL CASTILLO, AND L. S. GARCÍA-COLÍN, *Determination of the isoentropic temperature in the glass transition*, Macromolecules **26** (1993), 655–658.

[15] P. GOLDSTEIN, L. F. DEL CASTILLO, AND L. S. GARCÍA-COLÍN, *On the Stokes-Einstein relation in glass forming liquids*, Physica A **275** (2000), 325–335.

[16] R. KERNER, *Two simple rules for covalent binary glasses*, Physica B **215** (1996), 267–272.

[17] L. M. MARTÍNEZ AND C. A. ANGELL, *A thermodynamic connection to the fragility of glass-forming liquids*, Nature **410** (2001), 663–667.

[18] G. G. NAUMIS, *Stochastic matrix description of glass transition in ternary chalcogenide systems*, J. Non-Cryst. Solids **231** (1998), 111–119.

[19] G. G. NAUMIS AND R. KERNER, *Modelling of growth and agglomeration processes leading to various non-crystalline materials*, J. Non-Cryst. Solids **232-234** (1998), 600–606.

[20] W. M. SASLOW, *Scenario for the Vogel-Fulcher law*, Phys. Rev. B **37** (1988), 676–678.

[21] D. SIDEBOTTOM, R. BERGMAN, L. BÖRJESSON, AND L. M. TORELL, *Two-Step Relaxation Decay in a Strong Glassformer*, Phys. Rev. Lett. **71** (1993), 2260–2263.

[22] F. STICKEL, E. W. FISCHER, AND R. RICHTER, *Dynamics of glass-forming liquids. I. Temperature-derivative analysis of dielectric relaxation data*, J. Chem. Phys. **102** (1995), 6251–6257.

[23] F. STICKEL, E. W. FISCHER, AND R. RICHERT, *Dynamics of glass forming liquids. I. Time derivative analysis of dielectric relaxation data*, J. Chem. Phys. **102** (1995), 6251–6257.

[24] F. STICKEL, E. W. FISCHER, AND R. RICHTER, *Dynamics of glass-forming liquids. II. Detailed comparison of dielectric relaxation, dc-conductivity, and viscosity data*, J. Chem. Phys. **104** (1996), 2043–2055.

[25] M. TATSUMISAGO, B. L. HALIPAP, J. L. GREEN, S. M. LINDSAY, AND C. A. ANGELL, *Fragility of Ge-As-Se glass-forming liquids in relation to rigidity percolation, and the Kauzmann paradox*, Phys. Rev. Lett. **64** (1990), 1549–1542.

[26] L. TICHY AND H. TICHÁ, *Covalent bond approach to the glass-transition temperature of chalcogenide glasses*, J. Non-Cryst. Solids **189** (1995), 141–146.

[27] M. L. WILLIAMS, R. F. LANDEL, AND J. D. FERRY, *The temperature dependence of relaxation mechanisms in amorphous polymers and other glass-forming liquids*, J. Am. Cer. Soc. **77** (1955), 3701–3707.

Bol. Soc. Mat. Mexicana (3) Vol. 19, 2013

HAMILTONIAN TENSORIAL SPECIAL RELATIVISTIC MECHANICS OF INTERACTING POINT PARTICLES

In honor of Prof. Leopoldo García-Colín Scherer

E. PIÑA

ABSTRACT. This paper starts with the presentation of a Hamiltonian dynamics in terms of 4-vectors in Minkowski space which was developed to describe the dynamics of interacting relativistic point particles through a field. The independent parameter (replacing time in the classical theory) is associated to a family of surfaces corresponding to the Dirac-form of the dynamics. By a canonical parameter dependent Lorentz–Poincaré transformation, this Hamiltonian formalism treats the Lie–Dirac generators for any form of relativistic dynamics as coefficients of a first degree polynomial in the ten translation and rotation velocities of the Lorentz–Poincaré transformation. The Currie's world line conditions are generalized to any Dirac-form of the dynamics. The formalism is illustrated with the electromagnetic interaction. The explicit relation between the covariant field variables and the more usual 3-dimensional Fourier variables associated to the instant form of the dynamics is exhibited.

Introduction

In the first part of this paper, I solve the mathematical problem of the construction of a Hamiltonian formulation of the relativistic dynamics of interacting particles in terms of a tensorial notation that is evidently invariant with respect to Lorentz transformations. A first answer to this problem was published in the *Bulletin de l'Académie Royale de Belgique* [1]. The purpose here is to present those ideas in a more formal way, stressing the challenge solved, although using traditional physics notation.

The discovery of the Lorentz group and its physical meaning happened with the birth of the Twentieth Century on the shoulders of Lorentz, Poincaré, Einstein, Minkowski, etc. [2]. The time (multiplied by the velocity of light, c, which, in this paper, is taken to be the unit of velocity: $c = 1$) is a fourth coordinate, forming, together with the three coordinates of space, Minkowski's space-time. The coordinates of this space are x^μ in $\mathbb{R}^4$, Greek indices running over $0,1,2,3$.

The basic philosophy at first was to find a 4-vector covariant Hamiltonian providing the equations of motion for particles and fields. For the particles, the non-interacting case obtains an interaction by means of the replacement of the non-interacting 4-momentum of particle j, p_j^μ, by the difference $p_j^\mu - e_j A_j^\mu(x_j)$ of the previous 4-vector and the potential 4-vector for the field.

2010 *Mathematics Subject Classification*: Primary: 70H40; Secondary: 83A05.

Keywords and phrases: covariant Hamiltonian, Dirac's Lie-Poincaré generators, Currie's wordline conditions.

The author was invited to write down this work as a tribute to Professor Leopoldo García-Colín Scherer.

The inspiration for describing Maxwell's equations for the electromagnetic field as a Hamiltonian dynamics was born from considering the Fourier transformation of the space (not time) coordinates and thinking of the transformed Maxwell equations in terms of the Fourier transformation of the 4-vector potential. There comes about an infinite set of harmonic oscillators coupled to the Fourier transformation of the 4-vector current density as a function of the particles' positions and velocities. This is easily set as a Hamiltonian formalism, which however has lost its full 4-vector notation because of the asymmetric roles of position and time. To recover the symmetry of the four coordinates of the particles, one introduces an arbitrary parameter for describing the space-time position of any particle, and one looks for a formal solution of the field equations with some similarity to the harmonic oscillator model.

Two coordinate systems, one moving with respect to the other at a constant velocity, transform the four coordinates of a particle by means of a linear Lorentz transformation that preserves in both frames Maxwell's electromagnetic theory [3, 4, 5]. The equations of motion of the particles should be invariant with respect to the same group.

Denote $\hat{x}^\mu$ in $\mathbb{R}^4$, the transformed coordinates ($x^\mu \to \hat{x}^\mu$) by the Lorentz transformation represented by the matrix $\mathcal{L}^\mu_{\ \lambda}$ in $\mathbb{R}^{4\times 4}$,

$$\hat{x}^\mu = \mathcal{L}^\mu_{\ \lambda}\, x^\lambda \tag{0.1}$$

(the Greek indices when repeated are summed from 0 to 3). A Lorentz transformation is characterized by leaving invariant the metric matrix $\eta_{\alpha\beta}$

$$\eta_{\alpha\beta}\mathcal{L}^\alpha_{\ \mu}\mathcal{L}^\beta_{\ \nu} = \eta_{\mu\nu}\,, \tag{0.2}$$

where $\eta_{\alpha\beta}$ is the metric tensor, equal to its inverse, denoted by $\eta^{\alpha\beta}$

$$\eta_{\alpha\beta} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} = \eta^{\alpha\beta}\,. \tag{0.3}$$

To prove the invariance of the Maxwell theory in a trivial way, it is expressed in terms of coordinates and 4-vectors (that transform as the coordinates) [5].

$$\Box A^\mu(x) = 4\pi J^\mu(x)\,. \tag{0.4}$$

$\Box$ is the D'Alembertian

$$\Box = \eta^{\alpha\beta}\frac{\partial}{\partial x^\alpha}\frac{\partial}{\partial x^\beta}\,, \tag{0.5}$$

$A^\beta(x)$ in $\mathbb{R}^4$ is the β-component of the 4-vector electromagnetic potential: this is a 4-vector field at the point x^μ in space-time. J^μ is the 4-vector current density

$$J^\mu(x) = \sum_{j=1}^{N} e_j \int_{-\infty}^{\infty} d\tau\, \delta^4(x - x_j(\tau))\frac{dx^\mu_j(\tau)}{d\tau}\,; \tag{0.6}$$

where x^μ_j in $[(\mathbb{R})^4]^{\mathrm{N}}$ is the μ-component of the position of the particle j; e_j in $\mathbb{R}^{\mathrm{N}}$ is the charge of the same particle; and τ in $\mathbb{R}$ is an arbitrary parameter to define the trajectories of the particles, such that the 0 component of particle j, x^0_j, representing the time of this particle, is a monotonic increasing function of τ. The dot denotes the derivative with respect to the parameter τ.

The differential equations of motion for the particle j ($1 \leq j \leq$ N) [5] are

$$m_j \eta_{\alpha\beta} \frac{d}{d\tau} \frac{\dot{x}_j^\beta}{[\eta_{\mu\nu}\dot{x}_j^\mu \dot{x}_j^\nu]^{1/2}} = e_j \dot{x}_j^\beta \left[\frac{\partial A_\beta(x_j)}{\partial x^\alpha} - \frac{\partial A_\alpha(x_j)}{\partial x^\beta} \right], \tag{0.7}$$

where m_j, positive, is the mass of the particle j.

The equations for the particles have a Lorentz force in the right hand side which is a function of the electromagnetic potential. Maxwell's equations for the field have as a source term the current density 4-vector that is a function of the positions of the particles.

However, for Quantum Theory and/or Statistical Mechanics, it is convenient to derive these equations of motion from a Hamiltonian theory. The first Hamiltonian dynamics for relativistic particles [6] was formulated assuming time t as the independent parameter, instead of the general parameter τ. The 0 component of a position in space-time and for all the particles is the time

$$x^0 = t, \quad x_j^0 = t \quad (1 \leq j \leq N). \tag{0.8}$$

This Hamiltonian dynamics is formulated in a particular system of coordinates. The Lorentz invariance is not evident *a priori*, although it is implicit because the frame is particular, but otherwise arbitrary. No preferential frame has been chosen.

One introduces the Fourier expression in 3-space for the 4-vector potential

$$A^\beta(\mathbf{r}) = \int \frac{d^3k}{|\mathbf{k}|} \{A_k^\beta \exp[-i\mathbf{k}\cdot\mathbf{r}] + A_k^{\dagger\beta} \exp[i\mathbf{k}\cdot\mathbf{r}]\}, \tag{0.9}$$

where

$$\mathbf{r} = (x^1, x^2, x^3) \tag{0.10}$$

are the components of the 3-space coordinates, and

$$\mathbf{k} = (k^1, k^2, k^3) \tag{0.11}$$

are the components of the plane wave vector.

There results a Fourier transformation of the (0.5) form of the field equations.

The field coordinates A_k^β and $A_k^{\dagger\beta}$ satisfy a system of non-homogeneous harmonic oscillator equations, which allows a Hamiltonian formulation with the Hamiltonian

$$H = \sum_{j=1}^{N} \left\{ e_j \mathcal{A}^0(\mathrm{q}_j) + \sqrt{[\mathbf{p}_j - e_j \mathbf{A}(\mathrm{q}_j)]^2 + m_j^2} \right\} - 4\pi^2 \int d^3k A_k^\lambda A_{k\lambda}^\dagger, \tag{0.12}$$

where we use a different notation for the 4-vector potential to stress that it is now considered a functional of the field Fourier variables

$$\mathcal{A}^\beta(\mathbf{q}_j, [A_k^\beta], [A_k^{\dagger\beta}]) = \int \frac{d^3k}{|\mathbf{k}|} \{A_k^\beta \exp[-i\mathbf{k}\cdot\mathbf{q}_j] + A_k^{\dagger\beta} \exp[i\mathbf{k}\cdot\mathbf{q}_j]\}, \tag{0.13}$$

where the 3-coordinates for the particles are

$$\mathbf{q}_j = (\mathrm{q}_j^1, \mathrm{q}_j^2, \mathrm{q}_j^3) \tag{0.14}$$

with its canonical conjugated coordinates

$$\mathbf{p}_j = (\mathfrak{p}_j^1, \mathfrak{p}_j^2, \mathfrak{p}_j^3) \tag{0.15}$$

and where A_k^β and $A_k^{\dagger\beta}$ are the instant form field variables. The Poisson brackets of these variables are given by

$$[A_k^\mu, A_{k'}^{\dagger\nu}] = -\frac{i}{4\pi^2}\eta^{\mu\nu}\delta^3(\mathbf{k}-\mathbf{k}'), \tag{0.16}$$

$$[A_k^\mu, A_{k'}^\nu] = [A_k^{\dagger\mu}, A_{k'}^{\dagger\nu}] = 0, \tag{0.17}$$

and

$$\mathbf{A}(\mathfrak{q}_j) = (\mathcal{A}^1(\mathfrak{q}_j), \mathcal{A}^2(\mathfrak{q}_j), \mathcal{A}^3(\mathfrak{q}_j)). \tag{0.18}$$

The Hamiltonian is one of the ten generators of the Lie algebra of the Lorentz–Poincaré group. The other nine generators are

$$\mathbf{P} = (P^1, P^2, P^3) = \sum_{j=1}^N \mathbf{p}_j - 4\pi^2 \int \frac{d^3k}{|\mathbf{k}|}\mathbf{k} A_k^\lambda A_{k\lambda}^\dagger, \tag{0.19}$$

$$\begin{aligned}\mathbf{J} = (M^{23}, M^{31}, M^{12}) = {} & \sum_{j=1}^N \mathbf{q}_j \times \mathbf{p}_j \\ & - 4\pi^2 \int \frac{d^3k}{|\mathbf{k}|}\left\{A_k^{\dagger\mu}\left(\mathbf{k}\times\frac{\partial}{\partial\mathbf{k}}\right)A_k^\mu + \mathbf{A}_k \times \mathbf{A}_k^\dagger\right\},\end{aligned} \tag{0.20}$$

and

$$\begin{aligned}\mathbf{K} = (M^{10}, M^{20}, M^{30}) = {} & \sum_{j=1}^N \mathbf{q}_j\left\{e_j\mathcal{A}^0(\mathfrak{q}_j) + \sqrt{[\mathbf{p}_j - e_j\mathbf{A}(\mathfrak{q}_j)]^2 + m_j^2}\right\} + \\ & + 4\pi^2 i \int \frac{d^3k}{|\mathbf{k}|}\left\{A_{k\mu}^\dagger |\mathbf{k}| \frac{\partial}{\partial\mathbf{q}} A_k^\mu - \mathbf{A}_k A_k^{\dagger 0} + \mathbf{A}_k^\dagger A_k^0\right\},\end{aligned} \tag{0.21}$$

where

$$\mathbf{A}_k = (A_k^1, A_k^2, A_k^3), \quad \mathbf{A}_k^\dagger = (A_k^{\dagger 1}, A_k^{\dagger 2}, A_k^{\dagger 3}). \tag{0.22}$$

A useful reference for this Hamiltonian formalism, with many additional details and remarks, is Balescu, Kotera, and Piña [7, 8].

In the first part of the present paper we prove two propositions:

PROPOSITION (0.23). *It is possible to construct a Hamiltonian function, as generalized by Dirac* [9, 10]*, using only tensorial quantities, whose Hamiltonian equations of motion provide the equations of motion for the particles and for the field, making evident the covariant character of this Hamiltonian with respect to Lorentz transformations.*

We have used the electromagnetic field only for simplicity, but it is possible to generalize this theory to the Van Dam–Wigner interaction [11]. The equations of motion for this system are Lorentz-like equations of motion for the particles with an action-at-distance Van Dam–Wigner interaction [11]. A particular case of this interaction is the case of electric charges in the presence of an external field obeying the Maxwell equations, which are easily written in Minkowski space with a tensorial notation [5] showing immediately their invariant character for all coordinate systems related by Lorentz transformations. As is well known, these last equations were formulated in terms of action-at-distance equations of motion in classical papers by Wheeler and Feynman [12, 13]. The Lorentz invariance of the action-at-distance interaction is equally evident.

It has been assumed that the Van Dam–Wigner equations cannot be written in terms of a Hamiltonian formalism, and that action-at-a-distance would exclude a field theory [14]. Both assumptions are shown to be *non sequiturs* in this paper. The second proposition is

PROPOSITION (0.24). *It is possible to reformulate the Van Dam–Wigner action-at-a-distance equations of motion* [11] *as a Hamiltonian theory of particles interacting through a field.*

The proof is at the end of the second section.

In the third section is presented a smooth transition from the previous tensorial notation to the canonical formalism where the Lie algebra of the Lorentz–Poincaré group emerges, establishing the connection between them.

It is possible to take into account the covariant aspect of the relativistic theory by working with a set of canonical generators for the Lorentz group [7, 8]. The origin for this other formalism is also due to Dirac [15], who constructed several sets of generators, relating each set to a constraint on the particle coordinates. He calls each set "a form" (of dynamics).

This point of view has been adopted by several authors to study various aspects of relativistic mechanics.

With this formalism, many authors, beginning with Bakamjian–Thomas [16] and Foldy [17], have constructed generators, depending only on the canonical coordinates for the particles. Currie [18], and Currie, Jordan and Sudarshan [19] showed that these generators may be consistent with invariant trajectories only for the non-interacting case.

Balescu, Kotera and Piña [7, 8] worked with Dirac's instant form of dynamics, introducing canonical variables for the field. The non-interaction Currie theorem does not apply for this case, and they developed an interesting formalism for relativistic statistical mechanics on this basis.

They showed that the Currie world line condition is an essential requirement to prove that the macroscopic 4-vector current density transforms as a 4-vector field *vis a vis* the Lorentz transformation generated by the canonical generators of the Lorentz–Poincaré group. Their proof uses the fact that the 4-vector current density obeys the characteristic system of differential equations of 4-vectors with respect to the Lie parameters, a system which is trivially integrated.

In order to avoid Currie's non-interaction theorem, we have always used canonical variables for the field. This selection requires renormalization techniques in order to suppress the singularities in the field. But these difficulties are not fundamental ones [20].

These field variables are very useful from the physical point of view, when one wishes to express in a simple form many of the electromagnetic phenomena such as radiation, absorption, dispersion, etc.

In the third section by means of a parameter-dependent canonical transformation one proves the theorem

THEOREM (0.25). *The Lie algebra of the Lorentz group corresponding to a particular "form" of the dynamics is obtained as the compatibility between the Hamiltonian and the Routhian formalisms associated to a parameter dependent Lorentz transformation.*

As a consequence, we prove the proposition

PROPOSITION (0.26). *It is possible to generalize Currie's world-line condition to an arbitrary "form" of the dynamics.*

In the last section, we reconsider the "instant form" to transform explicitly the tensorial formalism developed in this paper to the more usual formalism presented in this introductory section.

1. Dirac's Hamiltonian formulation

This section is devoted to the proof of the first proposition and to the description of the proof of the second. For this purpose we use the generalized Hamiltonian dynamics developed by Dirac [9, 10].

Our system will be a collection of particles interacting through a field.

We look for a formal solution of eq. (0.4). Let us introduce the Fourier transformation q^μ_κ of the 4-vector potential

$$A^\mu(x) = \int d^4\kappa \frac{1}{\sqrt{4\pi^3\kappa_\alpha\kappa^\alpha}} q^\mu_\kappa \exp[i\kappa_\beta x^\beta]. \tag{1.1}$$

The formal solution of eq. (0.4) is

$$q^\mu_\kappa = -\frac{1}{\sqrt{4\pi^3\kappa_\alpha\kappa^\alpha}} \sum_{j=1}^{N} e_j \int_{-\infty}^{\infty} d\tau \dot{x}^\mu_j(\tau) \exp[-i\kappa_\beta x^\beta_j(\tau)]. \tag{1.2}$$

We define the canonical variables for the field, the coordinates

$$y^\mu_\kappa = -\frac{1}{\sqrt{4\pi^3\kappa_\alpha\kappa^\alpha}} \sum_{j=1}^{N} e_j \int_{-\infty}^{\tau} d\tau \dot{x}^\mu_j(\tau) \exp[-i\kappa_\beta x^\beta_j(\tau)] \tag{1.3}$$

and the canonical momenta

$$p^\mu_\kappa = -\frac{1}{\sqrt{4\pi^3\kappa_\alpha\kappa^\alpha}} \sum_{j=1}^{N} e_j \int_{\tau}^{\infty} d\tau \dot{x}^\mu_j(\tau) \exp[i\kappa_\beta x^\beta_j(\tau)]. \tag{1.4}$$

They are related to the Fourier transformation of the 4-vector potential by the expression

$$q^\mu_\kappa = y^\mu_\kappa + p^\mu_{-\kappa}. \tag{1.5}$$

From definitions (1.3) and (1.4) we obtain the differential equations of motion for the field variables

$$\dot{y}^\mu_\kappa = -\frac{1}{\sqrt{4\pi^3\kappa_\alpha\kappa^\alpha}} \sum_{j=1}^{N} e_j \dot{x}^\mu_j(\tau) \exp[-i\kappa_\beta x^\beta_j(\tau)] \tag{1.6}$$

and

$$\dot{p}^\mu_\kappa = \frac{1}{\sqrt{4\pi^3\kappa_\alpha\kappa^\alpha}} \sum_{j=1}^{N} e_j \dot{x}^\mu_j(\tau) \exp[i\kappa_\beta x^\beta_j(\tau)]. \tag{1.7}$$

$$A^{\mu}(x) = \int d^4\kappa \frac{1}{\sqrt{4\pi^3\kappa_\alpha\kappa^\alpha}}\{y^\mu_\kappa \exp[i\kappa_\beta x^\beta] + p^\mu_\kappa \exp[-i\kappa_\beta x^\beta]\}. \tag{1.8}$$

We are going to consider Dirac's formalism [9, 10] introduced to generalize the Hamiltonian formulation of dynamics; it is particularly useful for the case where the Lagrangian is a first order homogeneous function of the velocities.

For this situation, the ordinary Hamiltonian is identically equal to zero and one finds a family of constraints among the canonical variables

$$\phi_n(x,p) = 0 \quad (n = 1,2,\dots). \tag{1.9}$$

Dirac then introduces the generalized Hamiltonian

$$H = \sum_n v_n \phi_n(x,p), \tag{1.10}$$

where the v_n's are Lagrange multipliers.

We will work out a formalism of this type for the simple situation in which the functions ϕ_n that determine the constraints of the system satisfy the restriction that the Poisson bracket between any two of them is equal to zero:

$$[\phi_n, \phi_m] = 0. \tag{1.11}$$

Let us introduce the constraints (one for each particle)

$$\phi_j = [p^\alpha_j - e_j A^\alpha(x_j)]\eta_{\alpha\beta}[p^\beta_j - e_j A^\beta(x_j)] - m_j^2 = 0. \tag{1.12}$$

Let us prove the first proposition: Dirac's Hamiltonian now becomes

$$H = \sum_{j=1}^{N} v_j\{p^\alpha_j - e_j A^\alpha(x_j)]\eta_{\alpha\beta}[p^\beta_j - e_j A^\beta(x_j)] - m_j^2\}, \tag{1.13}$$

where $A^\mu(x)$ is given explicitly in terms of the canonical field variables as in (1.8). From Hamilton's equations associated to the Hamiltonian (1.13), making use of the constraints (1.12), it is possible to determine the v_j as follows.

$$v_j = \frac{1}{2m_j}[\eta_{\alpha\beta}\dot{x}^\alpha_j \dot{x}^\beta_j]^{1/2}. \tag{1.14}$$

This means that the Lagrange multipliers v_j are proportional to the "velocity" along the world line of the respective particle, measured in τ-units.

Substituting (1.14) into Hamilton's equations for the particles and field variables, we recover the equations of motion (0.7), (1.6) and (1.7) for the particles and field.

In order to arrive at the field equation (0.4) we must use the boundary conditions

$$\lim_{\tau\to-\infty} y^\mu_\kappa(\tau) = 0 \tag{1.15}$$

and

$$\lim_{\tau\to\infty} p^\mu_\kappa(\tau) = 0. \tag{1.16}$$

□

It is interesting to note that the equations of motion are invariant with respect to a change of parameter τ

$$\tau \longrightarrow F(\tau). \tag{1.17}$$

This property implies an undetermined character of the equations of motion, as long as the τ parameter is not fixed by additional restrictions. In the next section we will introduce a different dynamical approach to give a determinate aspect to the equations of motion.

The following gives a sketch of the proof of the second proposition. It is based on the particular case of the electromagnetic field. This is formulated as explained before in terms of field variables that are Fourier transformations of the 4-vector potential.

As shown by Wheeler and Feynmann [12, 13] the field is replaced by an action-at-a-distance theory by a formal solution of the Maxwell equations. The Wheeler and Feynman equations of motion are a particular case of the Van Dam–Wigner equations [11]. In the place where Van Dam and Wigner introduced an arbitrary function of $\eta_{\alpha\beta}(x^\alpha - y^\alpha)(x^\beta - y^\beta)$, Wheeler and Feynmann have the Green's function corresponding to the D'Alembertian operator $\square$ defined in (0.5).

On the other hand, in the expressions (1.1) for the 4-vector potential and (1.2) and (1.3) one finds the square root of the Fourier transform of the Green's function of the same operator. To generalise the previous Hamiltonian with field variables to the Van Dam–Wigner interaction, it is sufficient to replace the square root of the Fourier transformation of the Green's function of the Maxwell equations by the square root of the Fourier transformation of the arbitrary Van Dam–Wigner function. $\square$

2. Dirac's canonical formulation

In this section we want to relate the previous formulation to another one associated to Dirac's ideas.

Dirac set up [15] a canonical representation for the Lorentz–Poincaré group

$$[P_\mu, P_\nu] = 0\,, \tag{2.1}$$

$$[M_{\mu\nu}, P_\lambda] = \eta_{\lambda\nu} P_\mu - \eta_{\lambda\mu} P_\nu\,, \tag{2.2}$$

$$[M_{\alpha\beta}, M_{\mu\nu}] = \eta_{\alpha\nu} M_{\beta\mu} + \eta_{\beta\mu} M_{\alpha\nu} + \eta_{\alpha\mu} M_{\nu\beta} + \eta_{\beta\nu} M_{\mu\alpha}\,. \tag{2.3}$$

Starting from a trivial (geometrical) representation for this group, he proposed constructing a new one, where the ten generators P_ν, $M_{\alpha\beta}$, must obey the Lie algebra restrictions of this group, using as Lie bracket the canonical Poisson bracket. He also requires the ten generators to have zero Poisson bracket with a particular function of the coordinates which specifies the "form" of the dynamics. He gave some solutions, but he did not introduce any specific technique for finding these solutions. For instance, Dirac considers various forms, some of which are listed below.

The instant form,

$$\mathrm{q}^0 = 0\,. \tag{2.4}$$

The light cone form,

$$\mathrm{q}^\alpha \mathrm{q}_\alpha = 0\,. \tag{2.5}$$

The hyperboloid form

$$\mathrm{q}^\alpha \mathrm{q}_\alpha - A^2 = 0\,. \tag{2.6}$$

Etc.

Our main aim will be to give a clearer physical or geometrical meaning to this formalism, to obtain it systematically from the Hamiltonian formalism developed in the previous section, and to give a synthetic method for obtaining the solution to the Dirac's problem of constructing a canonical representation of the Lorentz–Poincaré group, consistent with any "form" of the dynamics.

We consider the τ-dependent Lorentz inhomogeneous transformation obtained by canonical transformation of the Hamiltonian problem presented in last section. This canonical transformation will be generated by the function [21]

$$F_2 = \sum_{j=1}^{N} \mathfrak{p}_{j\nu}[\mathcal{L}^{\nu}_{\ \mu} x^{\mu}_{j} - z^{\nu}] + \int d^4k \mathcal{P}_{k\beta} \mathcal{L}^{\beta}_{\ \mu} y^{\mu}_{\kappa} \exp[ik_{\gamma} z^{\gamma}], \tag{2.7}$$

where $\mathcal{L}^{\alpha}_{\ \mu}$ are the components of the Lorentz transformation matrix; z^{ν} is a 4-vector translation; $\mathfrak{p}_{j\nu}$ are the new four momenta of the particles; $\mathcal{P}_{k\beta}$ are the new canonical momenta of the field, and k^{α} is a new wave vector related to the old one κ^{α} by the same Lorentz transformation

$$\kappa_{\alpha} = \mathcal{L}^{\nu}_{\ \alpha} k_{\nu}. \tag{2.8}$$

Both quantities $\mathcal{L}^{\alpha}_{\ \mu}$ and z^{ν} will be assumed to be explicit functions of the parameter τ used to describe the motion. This gives a τ-dependence for the F_2 generating function.

Let us make use of F_2 to generate the canonical transformation.

The new particle coordinates are

$$\mathfrak{q}^{\nu}_{j} = \mathcal{L}^{\nu}_{\ \mu} x^{\mu}_{j} - z^{\nu} = \frac{\partial F_2}{\partial \mathfrak{p}_{j\nu}}, \tag{2.9}$$

which is a τ-dependent Lorentz–Poincaré transformation of the coordinates of the particles.

Analogously, we find the old momenta

$$p_{j\mu} = \mathfrak{p}_{j\nu} \mathcal{L}^{\nu}_{\ \mu} = \frac{\partial F_2}{\partial x^{\mu}_{j}}. \tag{2.10}$$

The new coordinates for the field are

$$\mathcal{Q}^{\beta}_{k} = \mathcal{L}^{\beta}_{\ \mu} y^{\mu}_{\kappa} \exp[ik_{\gamma} z^{\gamma}] = \frac{\delta F_2}{\delta \mathcal{P}_{k\beta}} \tag{2.11}$$

and the old momenta for the field are given by

$$p_{\kappa\mu} = \mathcal{P}_{k\beta} \mathcal{L}^{\beta}_{\ \mu} \exp[ik_{\gamma} z^{\gamma}] = \frac{\delta F_2}{\delta y^{\mu}_{\kappa}}. \tag{2.12}$$

The new Hamiltonian is found by the prescription [21]

$$\mathcal{H} = H + \frac{\partial F_2}{\partial \tau}. \tag{2.13}$$

In order to calculate this expression we need the result that follows from (2.11)

$$\frac{\partial \mathcal{Q}^{\beta}_{k}}{\partial k_{\nu}} = i z^{\nu} \mathcal{Q}^{\beta}_{k} + \mathcal{L}^{\beta}_{\ \mu} \frac{\partial y^{\mu}_{\kappa}}{\partial \kappa^{\alpha}} \mathcal{L}^{\nu}_{\ \alpha} \exp[ik_{\gamma} z^{\gamma}]. \tag{2.14}$$

The τ-derivative of the Lorentz tensor $\mathcal{L}^{\alpha}_{\ \mu}$ will be expressed in terms of an antisymmetric tensor as, similarly, in the theory of a rigid rotating body [22]

$$\dot{\mathcal{L}}^{\beta}_{\ \mu} = \omega^{\beta}_{\ \gamma}\mathcal{L}^{\gamma}_{\ \mu}, \tag{2.15}$$

where $\omega_{\alpha\beta}$ is an antisymmetric angular velocity tensor.

Let us now calculate the derivative $\partial F_2/\partial\tau$ and afterwards transform it to the new variables by using equations (2.15), (2.9), (2.11) and (2.14), and the antisymmetric character of the tensor $\omega_{\alpha\beta}$

$$\begin{aligned}\frac{\partial F_2}{\partial \tau} = -\dot{z}^{\nu}\sum_{j=1}^{N}\mathfrak{p}_{j\nu} + \dot{z}^{\nu} i\int d^4k\, k_{\nu}\mathcal{P}_{k\beta}\mathcal{Q}^{\beta}_{k} + \\ \frac{1}{2}\omega_{\alpha\beta}\int d^4k\left[(\mathcal{P}^{\alpha}_{k}\mathcal{Q}^{\beta}_{k} - \mathcal{P}^{\beta}_{k}\mathcal{Q}^{\alpha}_{k}) + \mathcal{P}_{k\gamma}\left(k^{\alpha}\frac{\partial}{\partial k_{\beta}} - k^{\beta}\frac{\partial}{\partial k_{\alpha}}\right)\mathcal{Q}^{\gamma}_{k}\right] + \\ \frac{1}{2}\omega_{\alpha\beta} i\int d^4k\,\mathcal{P}_{k\gamma}(k^{\alpha}z^{\beta} - k^{\beta}z^{\alpha})\mathcal{Q}^{\gamma}_{k}.\end{aligned} \tag{2.16}$$

In order to get the new Hamiltonian as a function of the new variables, we transform the 4-vector potential at the position of particle j

$$\mathcal{A}^{\beta}(\mathfrak{q}_j) = \int d^4k\frac{1}{\sqrt{4\pi^3 k_{\nu}k^{\nu}}}\{\mathcal{Q}^{\beta}_{k}\exp[ik_{\gamma}\mathfrak{q}^{\gamma}_{j}] + \mathcal{P}^{\beta}_{k}\exp[-ik_{\gamma}\mathfrak{q}\gamma_{j}]\}. \tag{2.17}$$

There follows

$$\mathcal{A}^{\beta}(\mathfrak{q}_j) = \mathcal{L}^{\beta}_{\ \mu}A^{\mu}(x_j). \tag{2.18}$$

With this result, the old Hamiltonian in the new variables has the same formal aspect as in the previous formulation

$$H = \sum_{j=1}^{N} v_j\{[\mathfrak{p}^{\beta}_{j} - e_j\mathcal{A}^{\beta}(\mathfrak{q}_j)]\eta_{\beta\gamma}[\mathfrak{p}^{\gamma}_{j} - e_j\mathcal{A}^{\gamma}(\mathfrak{q}_j)] - m^2_j\}. \tag{2.19}$$

and the new Hamiltonian is found by adding (2.16) and (2.19) according to (2.13).

The Hamiltonian formulation is completed by taking into account the transformed constraints

$$[\mathfrak{p}^{\alpha}_{j} - e_j\mathcal{A}^{\alpha}(\mathfrak{q}_j)]\eta_{\alpha\beta}[\mathfrak{p}^{\beta}_{j} - e_j\mathcal{A}^{\beta}(\mathfrak{q}_j)] - m^2_j = 0. \tag{2.20}$$

We are now going to determine the τ-parametrization by imposing a new constraint for each particle; these constraints fix the “form” of the dynamics

$$g(\mathfrak{q}^{\alpha}_{j}) = 0, \tag{2.21}$$

where g is a point function in the $\mathfrak{q}$ coordinate space.

This constraint in terms of the old coordinates is

$$g(\mathcal{L}^{\alpha}_{\ \mu}x^{\mu}_{j} - z^{\alpha}) = 0, \tag{2.22}$$

which shows more clearly the physical meaning: this constraint fixes the parametrization of the particles by the intersection of the world line of each particle with the τ-dependent family of surfaces

$$g(\mathcal{L}^{\alpha}_{\ \mu}(\tau)x^{\mu} - z^{\alpha}(\tau)) = 0. \tag{2.23}$$

Dirac’s examples [15] are now interpreted as follows. In the instant form,

$$g(\mathfrak{q}^{\alpha}_{j}) \equiv \mathfrak{q}^{0}_{j} = 0, \tag{2.24}$$

the particles are parametrized by a family of hyperplanes in the original space.

In the light-cone form,

$$g(\mathfrak{q}^\alpha_j) \equiv \mathfrak{q}^\alpha_j \eta_{\alpha\beta} \mathfrak{q}\beta_j = 0\,, \tag{2.25}$$

the particles are parametrized by a family of light-cones.

Returning to a general constraint, we look now to the preservation, following the motion, of the constraint

$$0 = \frac{dg}{d\tau} = \frac{\partial g}{\partial \mathfrak{q}^\alpha_j} \dot{\mathfrak{q}}^\alpha_j = [g, \mathcal{H}]\,. \tag{2.26}$$

This equation determines the Lagrange multipliers v_j in $\mathcal{H}$ as follows

$$v_j = \frac{\dot z^\beta \frac{\partial g}{\partial \mathfrak{q}^\beta_j} + \frac{1}{2}\omega_{\beta\gamma}[(\mathfrak{q}^\beta_j + z^\beta)\eta^{\gamma\mu} - (\mathfrak{q}^\gamma_j + z^\gamma)\eta^{\beta\mu}]\frac{\partial g}{\partial \mathfrak{q}^\mu_j}}{2[\mathfrak{p}^\alpha_j - e_j \mathcal{A}^\alpha(\mathfrak{q}_j)]\frac{\partial g}{\partial \mathfrak{q}^\alpha_j}}\,. \tag{2.27}$$

The Hamiltonian $\mathcal{H}$ is now written in Dirac's form as [20]

$$\mathcal{H} = -\dot z^\nu_R P_\nu - \frac{1}{2}\omega^{\alpha\beta} M_{\alpha\beta}\,, \tag{2.28}$$

where

$$\dot z^\nu_R = \dot z^\alpha - \omega^\alpha_{\ \beta} z^\beta = \mathcal{L}^\alpha_\beta \frac{d}{d\tau}(\mathcal{L}^\beta_\gamma z^\gamma)\,, \tag{2.29}$$

$$\begin{aligned} P_\nu = \sum_{j=i}^N \mathfrak{p}_{j\nu} - i \int d^4k\, k_\nu \mathcal{P}_{k\beta} \mathcal{Q}^\beta_k \\ - \sum_{j=1}^N \frac{\frac{\partial g}{\partial \mathfrak{q}^\nu_j}}{2[\mathfrak{p}^\alpha_j - e_j \mathcal{A}^\alpha(\mathfrak{q}_j)]\frac{\partial g}{\partial \mathfrak{q}^\alpha_j}} \{[\mathfrak{p}^\beta_j - e_j \mathcal{A}^\beta(\mathfrak{q}_j)]\eta_{\beta\gamma}[\mathfrak{p}^\gamma_j - e_j \mathcal{A}^\gamma(\mathfrak{q}_j)] - m^2_j\}\,, \end{aligned} \tag{2.30}$$

and where $M^{\alpha\beta}$ is the antisymmetric tensor

$$\begin{aligned} M^{\alpha\beta} = \sum_{j=1}^N (\mathfrak{q}^\alpha_j \mathfrak{p}^\beta_j - \mathfrak{q}^\beta_j \mathfrak{p}^\alpha_j) + \\ + \int d^4k \left[\mathcal{P}^\beta_k \mathcal{Q}^\alpha_k - \mathcal{P}^\alpha_k \mathcal{Q}^\beta_k + \mathcal{P}_{k\gamma}\left(k^\beta \frac{\partial}{\partial k_\alpha} - k^\alpha \frac{\partial}{\partial k_\beta}\right)\mathcal{Q}^\gamma_k\right] \\ - \sum_{j=1}^N \frac{(\mathfrak{q}^\alpha_j \eta^{\beta\gamma} - \mathfrak{q}^\beta_j \eta^{\alpha\gamma})\frac{\partial g}{\partial \mathfrak{q}^\gamma_j}}{2[\mathfrak{p}^\alpha_j - e_j \mathcal{A}^\alpha(\mathfrak{q}_j)]\frac{\partial g}{\partial \mathfrak{q}^\alpha_j}} \{[\mathfrak{p}^\mu_j - e_j \mathcal{A}^\mu(\mathfrak{q}_j)]\eta_{\mu\nu}[\mathfrak{p}^\nu_j - e_j \mathcal{A}^\nu(\mathfrak{q}_j)] - m^2_j\}\,. \end{aligned} \tag{2.31}$$

It is possible to consider the Hamiltonian $\mathcal{H}$ as a Routh function [21], i.e., as a Lagrangian with respect to the variables z^α and $\mathcal{L}^\alpha_\mu$.

The Lagrangian equation associated with the z^α variable gives us

$$\dot P_\mu = \omega^\beta_{\ \mu} P_\beta\,, \tag{2.32}$$

which expresses the conservation of the momentum 4-vector

$$\mathcal{L}^\beta_\alpha P_\beta\,. \tag{2.33}$$

Taking into account the Lorentz constraints

$$\eta_{\alpha\beta}\mathcal{L}^{\alpha}_{\ \mu}\mathcal{L}^{\beta}_{\ \nu} = \eta_{\mu\nu}\,, \tag{2.34}$$

the Lagrange equation associated to the variable $\mathcal{L}^{\alpha}_{\ \beta}$ gives us

$$\begin{aligned}\frac{d}{d\tau}(M^{\mu\beta} + z^{\mu}P^{\beta} - z^{\beta}P^{\mu}) &= \\ &= \omega^{\beta}_{\ \alpha}(M^{\mu\alpha} + z^{\mu}P^{\alpha} - z^{\alpha}P^{\mu}) - \omega^{\mu}_{\ \alpha}(M^{\beta\alpha} + z^{\beta}P^{\alpha} - z^{\alpha}P^{\beta}),\end{aligned} \tag{2.35}$$

which represent the conservation of the antisymmetric angular momentum tensor

$$\mathcal{L}^{\mu}_{\ \alpha}\mathcal{L}^{\nu}_{\ \beta}(M_{\mu\nu} + z_{\mu}P_{\nu} - z_{\nu}P_{\mu})\,. \tag{2.36}$$

On the other hand, the Hamiltonian equation of motion for P_{μ} is

$$\dot{P}_{\mu} = [\mathcal{H}, P_{\mu}] = -\dot{z}^{\nu}_{R}[P_{\nu}, P_{\mu}] - \frac{1}{2}\omega^{\alpha\beta}[M_{\alpha\beta}, P_{\mu}]\,. \tag{2.37}$$

Comparing (2.32) and (2.37) it follows that

$$[P_{\nu}, P_{\mu}] = 0 \tag{2.38}$$

and

$$[M_{\alpha\beta}, P_{\mu}] = \eta_{\beta\mu}P_{\alpha} - \eta_{\alpha\mu}P_{\beta}\,. \tag{2.39}$$

Studying the Hamiltonian equation of motion for $M_{\mu\nu}$ and comparing with (2.35), we also find

$$[M_{\alpha\beta}, M_{\mu\nu}] = \eta_{\alpha\nu}M_{\beta\mu} + \eta_{\beta\nu}M_{\mu\alpha} + \eta_{\alpha\mu}M_{\nu\beta} + \eta_{\beta\mu}M_{\alpha\nu}\,. \tag{2.40}$$

These three equations are the fundamental Lie algebra commutators of the Lorentz–Poincaré group. We find them as the compatibility conditions between the Lagrangian and the Hamiltonian formulations associated to the Routhian $\mathcal{H}$ which proves the theorem. □

Let's proceed to the proof of Proposition (0.26). The Hamiltonian expression for the preservation of the g constraint

$$[\mathcal{H}, g] = 0\,, \tag{2.41}$$

implies the properties

$$[P_{\nu}, g] = 0 \tag{2.42}$$

and

$$[M_{\alpha\beta}, g] = 0\,. \tag{2.43}$$

In Dirac's paper [15], these equations are the starting point for the determination of the ten generators P_{ν}, $M_{\alpha\beta}$, by an inductive method, different for each constraint. In this paper, on the contrary, the general expression for the generators (2.30) and (2.31) are obtained directly from the tensorial Hamiltonian formulation by applying a canonical transformation to a moving reference frame and finding the Lagrange multipliers v_j with the aid of the constraints g that determine the parametrization.

The method here presented has therefore the double advantage of showing explicitly the connection between the two Dirac formulations and of giving the general expression for the generators valid for any "form" of the dynamics.

At first sight, there is a pathological case in Dirac's paper. The two constraints

$$g(\mathfrak{q}) = \mathfrak{q}^\alpha \mathfrak{q}_\alpha \quad \text{and} \quad g(\mathfrak{q}) = \mathfrak{q}^\alpha \mathfrak{q}_\alpha - A^2 \tag{2.44}$$

should have equal generators in our formulation; however Dirac gives different types of generators in the two cases.

The paradox is solved by noting that the generators P_ν for Dirac's hyperboloid form may be transformed to the other P_γ generators by adding the "strong equation" (in Dirac's terminology [9])

$$\left\{ \frac{A^2}{2\mathfrak{p}_\nu \mathfrak{q}^\nu} (\mathfrak{p}_\sigma \mathfrak{p}^\sigma - m^2) \right\}^2 = 0 . \tag{2.45}$$

In order to verify directly the two equations (2.42) and (2.43), we found the interesting results

$$[\mathfrak{q}_j^\mu, P^\gamma] = -\eta^{\mu\gamma} + \frac{\eta^{\gamma\nu} \frac{\partial g}{\partial \mathfrak{q}_j^\nu} [\mathfrak{p}_j^\mu - e_j \mathcal{A}^\mu(\mathfrak{q}_j)]}{[\mathfrak{p}_j^\alpha - e_j \mathcal{A}^\alpha(\mathfrak{q}_j)] \frac{\partial g}{\partial \mathfrak{q}_j^\alpha}} \tag{2.46}$$

and

$$[\mathfrak{q}_j^\mu, M^{\alpha\beta}] = (\mathfrak{q}_j^\alpha \delta_\gamma^\beta - \mathfrak{q}_j^\beta \delta_\gamma^\alpha) [\mathfrak{q}_j^\mu, P^\gamma] . \tag{2.47}$$

These expressions are equivalent to Currie's [18] conditions for the trajectories of particles, which is the proof of Proposition (0.26). □

However it is necessary to remark that Currie's original formulae are related to Dirac's instant form studied in the next section, whereas equations (2.46) and (2.47) are valid for an arbitrary g-constraint. These equations will guarantee the condition of invariant trajectories of particles, independently of the "form" selected for the dynamical description.

3. The instant form

We will consider in this section the more usual, relativistic form of dynamics related to Dirac's instant form, where the g-constraint is

$$g(\mathfrak{q}_j^\alpha) \equiv \mathfrak{q}_j^0 = 0 . \tag{3.1}$$

This "instant form" is specially important because of its physical clearness and its analogy with the non-relativistic case.

The constraint (3.1) implies therefore

$$\frac{\partial g}{\partial \mathfrak{q}_j^\alpha} = \delta_\alpha^0 , \tag{3.2}$$

using Kronecker's delta.

As a consequence, the generators will take the form

$$P^\alpha = \sum_{j=1}^N \mathfrak{p}_j^\alpha - i \int d^4k \, k^\alpha \mathcal{P}_{k\beta} \mathcal{Q}_k^\beta \quad (\alpha = 1, 2, 3) , \tag{3.3}$$

$$P^0 = \sum_{j=1}^{N} \mathfrak{p}_j^0 - i\int d^4k k^0 \mathcal{P}_{k\beta}\mathcal{Q}_k^\beta - \sum_{j=1}^{N} \frac{1}{2[\mathfrak{p}_j^0 - e_j\mathcal{A}^0(\mathfrak{q}_j)]}\{[\mathfrak{p}_j^\beta - e_j\mathcal{A}^\beta(\mathfrak{q}_j)]\eta_{\beta\gamma}[\mathfrak{p}_j^\gamma - e_j\mathcal{A}^\gamma(\mathfrak{q}_j)] - m_j^2\}, \tag{3.4}$$

$$M^{\alpha\beta} = \sum_{j=1}^{N}(\mathfrak{q}_j^\alpha\mathfrak{p}_j^\beta - \mathfrak{q}_j^\beta\mathfrak{p}_j^\alpha) + \int d^4k\left\{\mathcal{Q}_k^\alpha\mathcal{P}_k^\beta - \mathcal{Q}_k^\beta\mathcal{P}_k^\alpha + \mathcal{P}_{k\gamma}\left(k^\beta\frac{\partial}{\partial k_\alpha} - k^\alpha\frac{\partial}{\partial k_\beta}\right)\mathcal{Q}_k^\gamma\right\} \qquad (\alpha \neq \beta = 1,2,3), \tag{3.5}$$

$$M^{\alpha 0} = \sum_{j=1}^{N}\mathfrak{q}_j^\alpha\mathfrak{p}_j^0 + \int d^4k\left\{\mathcal{Q}_k^\alpha\mathcal{P}_k^0 - \mathcal{Q}_k^0\mathcal{P}_k^\alpha + \mathcal{P}_{k\gamma}\left(k^0\frac{\partial}{\partial k_\alpha} - k^\alpha\frac{\partial}{\partial k_0}\right)\mathcal{Q}_k^\gamma\right\} - \sum_{j=1}^{N}\frac{\mathfrak{q}_j^\alpha}{2[\mathfrak{p}_j^0 - e_j\mathcal{A}^0(\mathfrak{q}_j)]}\{[\mathfrak{p}_j^\beta - e_j\mathcal{A}^\beta(\mathfrak{q}_j)]\eta_{\beta\gamma}[\mathfrak{p}_j^\gamma - e_j\mathcal{A}^\gamma(\mathfrak{q}_j)] - m_j^2\} \qquad (\alpha = 1,2,3). \tag{3.6}$$

For positive energy, the constraints will take the form

$$\mathfrak{p}_j^0 = e_j\mathcal{A}^0(\mathfrak{q}_j) + \sqrt{[\mathbf{p}_j - e_j\mathbf{A}(\mathfrak{q}_j)]^2 + m_j^2}, \tag{3.7}$$

where we recovereed the 3-vectorial notation (0.15) and (0.18)

Because $\mathfrak{q}_j^0$ is a constant, according to (3.1), we suppress explicitly the canonical conjugate variable $\mathfrak{p}_j^0$ by using the constraint equation (3.7).

The generators P^0 and $M^{\alpha 0}$ are modified to the new expressions

$$P^0 = \sum_{j=1}^{N}\left\{e_j\mathcal{A}^0(\mathfrak{q}_j) + \sqrt{[\mathbf{p}_j - e_j\mathbf{A}(\mathfrak{q}_j)]^2 + m_j^2}\right\} - i\int d^4k k^0\mathcal{P}_{k\beta}\mathcal{Q}_k^\beta \tag{3.8}$$

and

$$M^{\alpha 0} = \sum_{j=1}^{N}\mathfrak{q}_j^\alpha\left\{e_j\mathcal{A}^0(\mathfrak{q}_j) + \sqrt{[\mathbf{p}_j - e_j\mathbf{A}(\mathfrak{q}_j)]^2 + m_j^2}\right\} + \int d^4k\left\{\mathcal{Q}_k^\alpha\mathcal{P}_k^0 - \mathcal{Q}_k^0\mathcal{P}_k^\alpha + \mathcal{P}_{k\gamma}\left(k^0\frac{\partial}{\partial k_\alpha} - k^\alpha\frac{\partial}{\partial k_0}\right)\mathcal{Q}_k^\gamma\right\} \qquad (\alpha = 1,2,3). \tag{3.9}$$

For this particular choice of g-constraint the equations (2.46) and (2.47) will give us

$$\begin{aligned} [\mathfrak{q}_j^\alpha, P^\beta] &= \delta^{\alpha\beta} \\ [\mathfrak{q}_j^\gamma, M^{\alpha\beta}] &= \mathfrak{q}_j^\beta\delta^{\alpha\gamma} - \mathfrak{q}_j^\alpha\delta^{\beta\gamma}, \\ [\mathfrak{q}_j^\gamma, M^{\alpha 0}] &= \mathfrak{q}_j^\alpha[\mathfrak{q}_j^\gamma, P^0] \end{aligned} \tag{3.10}$$
$$(\alpha = 1,2,3; \beta = 1,2,3; \gamma = 1,2,3).$$

These are the conditions obtained by Currie [18], also Currie, Jordan and Sudarshan [19], for the invariance of the trajectories of particles. These formulae are valid only for the formulation that admits the g-constraint (3.1). We previously found the equations valid for an arbitrary constraint: they are equations (2.46) and (2.47). Equations (3.10) were obtained by Currie as compatibility conditions between the Lorentz transformation of the Hamiltonian formulation and the geometrodynamical transformation of the simultaneous positions of the particles. The method is based on comparing infinitesimal transformations with both techniques. This calculation was made with the explicit hypothesis that the coordinates of the particles are considered at the same time.

These authors did not remark that these conditions are modified in the case when a different parametrization is used.

Taking the variables $\mathcal{L}^{\alpha}_{\mu}$ to be constants, so that

$$\omega_{\alpha\beta} = 0 \tag{3.11}$$

and choosing the variables z^{μ} defined by the equations

$$z^0 = -\tau, \quad z^{\alpha} = 0 \quad (\alpha = 1,2,3), \tag{3.12}$$

this choice corresponds to a parametrization by the time measured in an arbitrary frame specified by the constants $\mathcal{L}^{\alpha}_{\mu}$.

Introducing (3.11) and (3.12) in the Routhian (2.28), we find the remarkable property

$$\mathcal{H} = P^0. \tag{3.13}$$

Lastly, we would like to point out the relation between the canonical variables for the field used in this paper and the formulation employed currently in the literature [6, 7].

Because of the constraint (3.1), the quantities appearing in the Fourier expression for the 4-vector potential

$$\mathcal{A}^{\beta}(\mathrm{q}_j) = \int d^4k \frac{1}{\sqrt{4\pi^3 k_{\gamma} k^{\gamma}}} [\mathcal{Q}^{\beta}_k + \mathcal{P}^{\beta}_{-k}] \exp[ik_{\alpha}\mathrm{q}^{\alpha}_j], \tag{3.14}$$

must have a singular character. This remark enables us to diminish the number of dimensions of the functional dependence of the field variables.

There comes about a first condition: that it is necessary to relate the variables by the equations

$$\begin{aligned} A^{\beta}_k = -\frac{i}{8\pi^2} \int dk_0 k \sqrt{4\pi^3(k_0^2 - k^2)} \times \\ \times \left\{ \mathcal{Q}^{\beta}_k \left[\delta(k_0^2 - k^2) + \frac{i}{\pi} \frac{1}{k_0^2 - k^2} \right] - \mathcal{P}^{\beta}_k \left[\delta(k_0^2 - k^2) - \frac{i}{\pi} \frac{1}{k_0^2 - k^2} \right] \right\} \end{aligned} \tag{3.15}$$

and

$$\begin{aligned} A^{\dagger\beta}_{-k} = \frac{i}{8\pi^2} \int dk_0 k \sqrt{4\pi^3(k_0^2 - k^2)} \times \\ \times \left\{ \mathcal{Q}^{\beta}_k \left[\delta(k_0^2 - k^2) - \frac{i}{\pi} \frac{1}{k_0^2 - k^2} \right] - \mathcal{P}^{\beta}_{-k} \left[\delta(k_0^2 - k^2) + \frac{i}{\pi} \frac{1}{k_0^2 - k^2} \right] \right\}. \end{aligned} \tag{3.16}$$

But these relations are not sufficient to obtain the generators in the new field variables as in Balescu, Kotera and Piña (0.12)-(0.21) [7, 8].

We found that in order to be consistent with the restrictions (3.15) and (3.16), the necessary relations for transforming our generators to the form (0.12), (0.19–0.21) become possible in two different forms, which were computed solving a very general linear relation among these variables assuming 32 different terms.

$$\mathcal{Q}_k^\mu = \frac{\sqrt{4\pi^3(k_0^2-k^2)}}{k} \times \times \left\{ \frac{1}{2} A_{-k}^{\dagger\mu} \delta_+(k_0+k) + \frac{3}{4} A_k^\mu \delta_-(k_0+k) - \frac{1}{4} A_k^\mu \delta_+(k_0-k) \right\}, \tag{3.17}$$

$$\mathcal{P}_k^\mu = \frac{\sqrt{4\pi^3(k_0^2-k^2)}}{k} \times \times \left\{ -\frac{1}{4} A_{-k}^{\mu} \delta_+(k_0+k) + \frac{1}{2} A_k^{\dagger\mu} \delta_+(k_0-k) + \frac{3}{4} A_{-k}^\mu \delta_-(k_0-k) \right\}. \tag{3.18}$$

Or

$$\mathcal{Q}_k^\mu = \frac{\sqrt{4\pi^3(k_0^2-k^2)}}{k} \times \times \left\{ -\frac{1}{4} A_{-k}^{\dagger\mu} \delta_+(k_0+k) + \frac{1}{2} A_k^\mu \delta_+(k_0-k) + \frac{3}{4} A_{-k}^{\dagger\mu} \delta_-(k_0-k) \right\}, \tag{3.19}$$

$$\mathcal{P}_k^\mu = \frac{\sqrt{4\pi^3(k_0^2-k^2)}}{k} \times \times \left\{ \frac{1}{2} A_{-k}^{\mu} \delta_+(k_0+k) + \frac{3}{4} A_k^{\dagger\mu} \delta_-(k_0+k) - \frac{1}{4} A_k^{\dagger\mu} \delta_+(k_0-k) \right\}. \tag{3.20}$$

These two possibilities are simply related. The last four equations used the distributions

$$\delta_\pm(x) \equiv \delta(x) \pm \frac{i}{\pi x}. \tag{3.21}$$

4. Conclusions

In this paper there were considered several fundamental aspects of the special relativistic classical theory of interacting particles.

The first point was to formulate, in Hamiltonian form, equations of motion for particles and a field in terms of 4-vectors. The Hamiltonian contains an undetermined Lagrange multiplier for each particle. This represents an advance with respect to covariant equations without interaction, and with respect to Hamiltonian equations without 4-vectorial covariance and, as well, with respect to integro-differential equations of motion without a Hamiltonian formulation.

When considering applications to quantum foundations or statistical mechanics, this is a basic point to have settled in order to construct any theory with a clear connection to the classical realm.

A second task was to find a smooth transition from the former Hamiltonian formalism of the particle–field equations of motion to Hamiltonian–Routhian equations produced when the coordinates of the particles and the field fulfil a parameter-dependent Lorentz transformation generated as a canonical transformation of the phase space variables. The new Hamiltonian is a first degree polynomial in the ten translation and rotation velocities of the parameter-dependent Lorentz transformation. The coefficients of this polynomial are the ten Dirac infinitesimal generators of the Lorentz–Poincaré group by canonical transformations. Simultaneously, the Dirac-form of the dynamics is imposed as a constraint. The value of this consists of several facts:

1. One has a technique for constructing the ten generators of the Lie algebra of the Lorentz–Poincaré group corresponding to any Dirac-form of the dynamics in terms of the defining constraint. This contrasts with the Dirac method, which seems to require a different treatment for each form.
2. The proof that these generators obey the Lie algebra equations with structure constants corresponding to the Lorentz–Poincaré group as a consequence of the Hamiltonian/Lagrangian equations associated to the Routhian.
3. The generalization to any Dirac-form of the dynamics of the Currie world line conditions for the trajectories of the particles, which were originally obtained by Currie for the instant form of parametrization considered in this paper in the introductory and last sections.
4. The fact that (by construction) these generators are compatible with the Currie non-interaction theorem, which does not apply to the particle–field interaction.

Finally, one has considered the instant-form of the dynamics. From the relativistic perspective, this is an obsolete perspective. Nevertheless, many scholars demand a return to this old perspective in order to fully understand the physics. Our 4-vector field Fourier transformations should be related to the Fourier transformations in the 3-dimensional plane associated to the instant form of the dynamics. The concern was to first see the necessary conditions for allowing the equality between the corresponding generators. Lastly, one has searched for an explicit relation between the different Fourier transformations, using the delta plus and delta minus functions that appear in the quantum field theory.

Compare this formalism for point particles with the similar for finite-size particles subject to electromagnetic interactions [24], published recently.

Acknowledgments

A first verbal version of the Van Dam–Wigner interaction as a Field Theory was presented as a *Thèse* in Brussels in 1967 to a jury formed by Professors O. Goche, R. Debever, I. Prigogine, J. Pelseneer, R. Balescu, and has not yet been published.

We thank Professor I. Prigogine for the interest he took in this problem and for his kind hospitality in Brussels, where the first version of this work was accomplished.

The author is also very much indebted to Prof. R. Balescu for advising this study and for many stimulating discussions.

Received March 11, 2013

Final version received April 15, 2013

E. PIÑA
PHYSICS DEPARTMENT
UNIVERSIDAD AUTÓNOMA METROPOLITANA-IZTAPALAPA
PO BOX 55 534, MÉXICO, D. F. 09340 MEXICO
pge@xanum.uam.mx

REFERENCES

[1] E. PIÑA, *Canonical formalism in special relativity (reproduced in arXiv 0908.3526)*, Acad. Roy. Belg. Bull. Cl. Sci. (5) **58**, (1972), 86–104.
[2] W. PAULI, Theory of Relativity, Pergamon Press, New York, 1958.
[3] H. POINCARÉ, *The present and future of mathematical physics*, Saint Louis Missouri Conference Am. Inst. Phys. (1904), 281–299.
[4] A. EINSTEIN, *Zur Elektrodynamik bewegter Körper*, Ann. Phys. **17**, (1905), 891–921.
[5] W. PANOFSKY AND M. PHILLIPS, Classical Electricity and Magnetism, Addison-Wesley, MA, 1956.
[6] W. HEITLER, The Quantum Theory of Radiation, Clarendon Press, Oxford, 1954.
[7] R. BALESCU AND T. KOTERA, *On the covariant formulation of classical relativistic statistical mechanics*, Physica **33**, (1967), 558–580.
[8] R. BALESCU, T. KOTERA AND E. PIÑA, *Lorentz transformations in phase space and in physical space*, Physica **33**, (1967), 581–594.
[9] P. A. M. DIRAC, *Generalized Hamiltonion dynamics*, Canad. J. Math **2**, (1950), 129–148.
[10] P. A. M. DIRAC, *Generalized Hamiltonian Dynamics*, Proc. Roy. Soc. **A 246**, (1958), 326–332.
[11] H. VAN DAM AND E. P. WIGNER, *Classical relativistic mechanics of interacting point particles*, Phys. Rev. **138 B**, (1965), 1576–1582.
[12] J. A. WHEELER AND R. P. FEYNMANN, *Interaction with the absorber as the mechanism of radiation*, Rev. Modern Phys. **17**, (1945), 157–161.
[13] J. A. WHEELER AND R. P. FEYNMANN, *Classical electrodynamics in terms of direct interparticle action*, Rev. Modern Phys. **21**, (1949), 425–433.
[14] S. CHEŁKOWSI, J. NIETENDEL AND R. SUCHANEK, *No-Interaction theorems in relativistic particle dynamic*, Acta Phys. Polon. **B11**, (1980), 809–825.
[15] P. A. M. DIRAC, *Forms of relativistic dynamics*, Rev. Modern Phys. **21**, (1949), 392–399.
[16] B. BAKAMJIAN AND L. H. THOMAS, *Relativistic particle dynamics II*, Phys. Rev. **92**, (1953), 1300–1310.
[17] L. L. FOLDY, *Relativistic particle systems with interaction*, Phys. Rev. **122**, (1961), 275–288.
[18] D. G. CURRIE, *Interaction contra classical relativistic hamiltonian particle dynamics*, J. Math. Phys. **4**, (1963), 1470–1488.
[19] D. G. CURRIE, T. F. JORDAN AND E. C. G. SUDARSHAN, *Relativistic invariance and Hamiltonian theories of interacting particles*, Rev. Modern Phys. **35**, (1963), 350–375.
[20] P. A. M. DIRAC, *Hamiltonian methods and quantum mechanics*, Proc. Roy. Irish. Acad. **63 A**, (1964), 49–59.
[21] H. GOLDSTEIN, Classical Mechanics, Addison-Wesley, MA, 1957.
[22] H. C. CORBEN AND P. STEHLE, Classical Mechanics, Wiley, New York, 1960.
[23] A. MERCIER, Analytical and Canonical Formalism in Physics, North-Holland, Amsterdam, 1959.
[24] C. CREMASCHINI AND M. TESAROTTO, *Hamiltonian structure of classical N-body system of finite-size particles subject to EM interactions*, dx.doi.org/10.1140/epjp/i2012-12004-4, Eur. Phys. J. Plus **27**:14, (2012).

CONCATENATING PERFECT POWERS

FLORIAN LUCA

ABSTRACT. For a positive integer $n > 1$, we search for two nth powers of positive integers such that if we concatenate their base $b > 1$ representations we obtain again an nth power of an integer.

1. Introduction

For a positive integer $b > 1$ and positive integers $A_1, \dots, A_t$, we write $\overline{A_1 \dots A_t}_{(b)}$ for the base b concatenation of the base b representations of $A_1, \dots, A_t$. For a positive integer m, we write $\ell_b(m)$ for the "length" of m is base b; i.e., the number of its base b digits. Hence,

$$\overline{A_1 \dots A_t}_{(b)} = \sum_{i=1}^{t} A_i b^{\sum_{j>i} \ell(A_j)}.$$

In this paper, we let $n \geq 2$ be an integer, and we search for two nth powers of positive integers, say $A_1 = x^n$ and $A_2 = y^n$, where x and y are positive integers, such that $\overline{A_1 A_2}_{(b)} = z^n$. Writing $m = \ell(A_2)$, we get the diophantine equation

$$x^n b^m + y^n = z^n \qquad \text{and} \qquad m = \ell_b(y^n). \tag{1.1}$$

A similar question when the perfect powers are replaced by members of a non-degenerate binary recurrent sequence of integers $(u_n)_{n \geq 0}$ was studied in [1].

A solution of equation (1.1) is called *reduced* if $b \nmid \gcd(y^n, z^n)$ and *primitive* if $\gcd(y^n, z^n) = 1$. Clearly, any primitive solution is reduced. It is easy to see that if (x_0, y_0, z_0) is a solution of equation (1.1), then $(x_0, y_0 \cdot b^t, z_0 \cdot b^t)$ is also a solution of equation (1.1) for all $t \geq 1$, but none of these solutions is reduced. Hence, any reduced solution gives rise to infinitely many nonreduced ones. It is for this reason that in this paper we look only at the reduced solutions of (1.1).

In this paper, we give finiteness results concerning the reduced/primitive solutions of the diophantine equation (1.1).

2. Concatenating squares

In this section, we assume that $n = 2$.

THEOREM (2.1). *When $n = 2$, equation (1.1) has infinitely many primitive solutions (x, y, z) if and only if b is not of the form $p^{2\alpha}$ with some odd prime number p and some positive integer α. When $b = p^{2\alpha}$ with some odd prime p and positive*

2010 *Mathematics Subject Classification*: 11A63, 11D41, 11J68.

Keywords and phrases: perfect powers, digital problems.

I thank William Banks for suggesting the questions treated in this paper and Pietro Corvaja for useful discussions. This work started during a visit to the University of Missouri, Columbia; the hospitality and support of this institution are gratefully acknowledged.

integer α, then equation (1.1) has infinitely many reduced solutions (x,y,z) when $\alpha>1$ and no reduced solution (x,y,z) when $\alpha=1$.

Proof. When b is not a square, the conclusion of Theorem (2.1) follows from the fact that the Pell equation $x^2b+1=z^2$ has infinitely many positive integer solutions (x,z).

When $b=b_0^2$ is a perfect square, we may relax the requirement that $m=\ell(y^2)$ in (1.1) to $m\geq \ell(y^2)$. Indeed, for if we have positive integers x, y, z, n such that

$$z^2=x^2b^m+y^2 \quad \text{and} \quad m\geq \ell(y^2):=m_1,$$

then

$$z^2=x^2b^m+y^2=(xb_0^{m-m_1})^2b^{m_1}+y^2,$$

which gives the solution (x_1,y,z) to equation (1.1) with $x_1:=xb_0^{m-m_1}$.

So, for the remaining of this proof, we work under the condition that $m\geq \ell(y^2)$. Furthermore, since equation (1.1) is of the form

$$(xb_0^m)^2+y^2=z^2, \tag{2.2}$$

it follows, by the well–known parametrization of Pythagorean triples, that there exist positive integers u, v and d, with u and v coprime and of different parities, such that

$$z=d(u^2+v^2) \quad \text{and} \quad \{xb_0^m,y\}=\{2duv,d|u^2-v^2|\}. \tag{2.3}$$

Clearly, (x,y,z) is reduced if and only if $b\nmid d^2$, which is equivalent to $b_0\nmid d$.

Assume now that b_0 is not a prime power. Then $b_0=c_0d_0$ holds with some coprime positive integers $c_0>1$ and $d_0>1$. For a fixed positive integer $m\geq 1$ let $(u_1(m),v_1(m))$ be the minimal positive integer solution (u_1,v_1) of the diophantine equation

$$u_1c_0^m-v_1d_0^m=1. \tag{2.4}$$

It is well-known and easy to prove that $u_1(m)\in\{1,\ldots,d_0^m-1\}$ and also that $v_1(m)\in\{1,\ldots,c_0^m-1\}$. Let $(u_{-1}(m),v_{-1}(m))$ be the minimal positive integer solution $(u_{-1}$, $v_{-1})$ of the equation

$$u_{-1}c_0^m-v_{-1}d_0^m=-1. \tag{2.5}$$

Again, $u_{-1}(m)\in\{1,\ldots,d_0^m-1\}$ and $v_{-1}(m)\in\{1,\ldots,c_0^m-1\}$. It is easy to see that $u_1(m)+u_{-1}(m)=d_0^m$ and $v_1(m)+v_{-1}(m)=c_0^m$. Thus,

$$\begin{aligned}(u_1(m)c_0^m+v_1(m)d_0^m)&+(u_{-1}(m)c_0^m+v_{-1}(m)d_0^m)\\ &=(u_1(m)+u_{-1}(m))c_0^m+(v_1(m)+v_{-1}(m))d_0^m\\ &=2(c_0d_0)^m=2b_0^m.\end{aligned}$$

It is clear that $u_1(m)c_0^m+v_1(m)d_0^m\neq(c_0d_0)^m$ (otherwise, we would get that d_0^m divides $u_1(m)$, which is not possible), therefore there exists $\zeta\in\{\pm1\}$ such that $u_\zeta(m)c_0^m+v_\zeta(m)d_0^m<b_0^m$. We take

$$d=1, \quad u=u_\zeta(m)c_0^m, \quad \text{and} \quad v=v_\zeta(m)d_0^m \tag{2.6}$$

in (2.3). Note that u and v are coprime because of relations (2.4) and (2.5). Note that with these choices

$$2duv=2u_\zeta(m)v_\zeta(m)(c_0d_0)^m=xb_0^m,$$

where $x = 2u_\zeta(m)v_\zeta(m)$, and

$$y = d|u^2 - v^2| = |u_\zeta(m)c_0^m - v_\zeta(m)d_0^m||u_\zeta(m)c_0^m + v_\zeta(m)d_0^m| < b_0^m = b^{m/2},$$

therefore $y^2 < b^m$, which shows that $\ell(y^2) \le m$. Since m is arbitrary, we have produced infinitely many solutions of equation (1.1) which are all primitive. This takes care of the case when b is not of the form $p^{2\alpha}$ for some prime p.

From now on, we suppose that $b = p^{2\alpha}$ for some prime p. Then $b_0 = p^\alpha$.

When $p = 2$, we take $m \ge 2$ and

$$d = 1, \quad u = \frac{b_0^m}{2}, \quad \text{and} \quad v = \frac{b_0^m}{2} - 1 \tag{2.7}$$

in (2.3). It is clear that u and v are coprime, since they are consecutive. With these choices, we have

$$2duv = xb_0^m, \qquad \text{where} \qquad x = \frac{b_0^m}{2} - 1,$$

and

$$y = d|u^2 - v^2| = b_0^m - 1 < b^{m/2},$$

therefore $y^2 < b^m$, which shows that $\ell(y^2) \le m$. Since m is arbitrary, we have produced infinitely many solutions of equation (1.1) which are all primitive.

Assume next that p is odd. Let (x, y, z) be a solution of equation (1.1). Assume that $2duv = xb_0^m = xp^\alpha$. If $\gcd(d, p) = 1$ (which is the case if the solution is primitive, or if $\alpha = 1$ and the solution is reduced), then $p^\alpha | uv$, and since u and v are coprime and p is odd, it follows that either u or v is a multiple of p^α. However, in this case

$$y = d|u^2 - v^2| \ge u + v > p^\alpha = b^{m/2},$$

therefore $y^2 > b^m$, contradicting again the fact that $\ell(y^2) \le m$.

Assume now that $d|u^2 - v^2| = xb_0^m = xp^\alpha$. If $\gcd(d, p) = 1$ (which is again the case if the solution is primitive, or if $\alpha = 1$ and the solution is reduced), then $p^\alpha | u^2 - v^2$, and since u and v are coprime, it follows that either $u - v$ is a multiple of p^α or $u + v$ is a multiple of p^α. This easily implies that $\max\{u, v\} > p^\alpha/2$. However, in this case

$$y = 2duv \ge 2\max\{u, v\} > p^\alpha = b^{m/2},$$

therefore $y^2 > b^m$, contradicting the fact that $\ell(y^2) \le m$.

The above arguments show, in particular, that there are no reduced solutions to equation (1.1) when $n = 2$ and $b = p$ is an odd prime, and that there are no primitive solutions when $b = p^\alpha$ for any odd prime p and any $\alpha \ge 1$.

It remains to show that if p is odd and $\alpha > 1$, then there are infinitely many reduced solutions to equation (1.1). To see this, we take $m \ge 1$, and

$$d = p, \quad u = p^{\alpha m - 1}, \quad \text{and} \quad v = p^{\alpha m - 1} - 1, \tag{2.8}$$

in (2.3). Note that u and v are coprime because they are consecutive. We get

$$2duv = 2p^{\alpha m}(p^{\alpha m - 1} - 1) = xb_0^m,$$

with $x = 2(p^{\alpha m - 1} - 1)$, and

$$y = d|u^2 - v^2| = p(p^{2\alpha m - 2} - (p^{\alpha m - 1} - 1)^2) = 2p^{\alpha m - 1} - 1 < p^{\alpha m} = b_0^m = b^{m/2},$$

therefore, $y^2 < b^m$, which implies that $\ell(y^2) \le m$. Since m is arbitrary, we have produced infinitely many solutions of equation (1.1). Clearly, all such solutions satisfy $\gcd(y^2,z^2)=d^2=p^2$, and since $1<p^2<b$, we get that these solutions are not primitive but they are reduced, which is what we wanted.

The proof of Theorem (2.1) is therefore complete. $\square$

3. Concatenating nth powers with $n>3$

From now on, we assume that $n\geq 3$. For every fixed n, and for every solution (x,y,z) of equation (1.1), let $r\in\{0,1,\dots,n-1\}$ be such that $m\equiv r \pmod n$. Then the solution (x,y,z) to equation (1.1) gives rise to the *rational point*

$$(X,Y)=(xb^{(m-r)/n}/z,y/z)$$

on the curve

$$\mathcal{C}_n(b,r) \qquad : \ b^rX^n+Y^n=1. \tag{3.1}$$

The curve $\mathcal{C}_n(b,r)$ is defined over $\mathbb{Q}$. The following proposition is useful.

PROPOSITION (3.2). *Let $n\geq 3$ be fixed. If equation (1.1) has infinitely many reduced solutions (x,y,z), there exists $r\in\{1,\dots,n-1\}$ such that the curve $\mathcal{C}_n(b,r)$ contains infinitely many rational points.*

Proof. If (1.1) has infinitely many reduced solutions, it follows that there exists r in $\{0,1,\dots,n-1\}$ such that infinitely many of these solutions will have $m\equiv r \pmod n$. For every prime factor p of b let α_p be such that $p^{\alpha_p}||b$. For every such solution, there exists p such that the order at which p divides $\gcd(y^n,z^n)$ is smaller than α_p. Since we have finitely many such primes p, and infinitely many reduced solutions, we may assume that the prime number p with the above property is the same for infinitely many of them. In particular, the order at which p divides $\gcd(y,z)$ is smaller than α_p/n. This shows that the numerator of $X=xb^{(m-r)/n}/z$ is divisible by p at a power at least $(m-r-\alpha_p)/n$. If m can be arbitrarily large, the statement of the proposition follows. If m is bounded, then, since $y<b^m$, it follows that y is bounded. Hence, we may assume that infinitely many of our solutions have the same fixed value for both y and m. Thus, we get the equation $y^n=z^n-b^r(xb^{(m-r)/n})^n$, where the only unknowns are now z and x (as y, n, m, r, b are all fixed). Since $n\geq 3$, this last equation is a particular case of a Thue equation, and it is well–known (see [6]) that such equations can have only finitely many integer solutions (x,z). This concludes the proof of the proposition. The fact that we may assume that $r\neq 0$ comes from the known fact that the Fermat curve $\mathcal{C}_n(b,0)$ has only finitely many (in fact, only three) rational points. $\square$

Proposition (3.2) together with Falting's Theorem (Mordell's Conjecture) yields immediately the following result.

COROLLARY (3.3). (i) *Let $n>3$ be fixed. Then equation (1.1) has only finitely many reduced solutions (x,y,z).*

(ii) *Let $n=3$. If the Mordell-Weil group of the elliptic curves $\mathcal{C}_3(b,r)$ is finite for both $r\in\{1,2\}$, then equation (1.1) has only finitely many reduced solutions (x,y,z).*

4. Concatenating cubes

From now on, we concentrate on the case $n = 3$. We first prove a finiteness theorem for the number of reduced solutions of equation (1.1) when $n = 3$ and b satisfies a certain technical condition. While the next Proposition (4.1) is contained in Theorem (4.11), we include as a good warm up for the proof of the more general Theorem (4.11).

PROPOSITION (4.1). *Assume that b has no prime factor $p \equiv 1 \pmod 3$. Then (1.1) has only finitely many reduced solutions when $n = 3$.*

Proof. Let (x, y, z) be a reduced solution of equation (1.1). We write $d = \gcd(y, z)$, $y_1 = y/d$, $z_1 = z/d$. We also write $d_1 = \gcd(x, d)$, $x_1 = x/d_1$, and $d_2 = d/d_1$. Finally, let $B = b^m/d_2^3$. Then equation (1.1) is

$$x_1^3 B + y_1^3 = z_1^3, \tag{4.2}$$

and $\gcd(x_1, y_1) = 1$. Furthermore, $B = b^m/d_2^3 \geq b^m/d_1^3 > y^3/d^3 = y_1^3$.

Let $B = B_0^3\alpha$, where α is cubefree. Since the only prime factors of α are among the prime factors of b, it follows that there are only finitely many choices for α. Thus, we may assume that α is fixed. Note further that if $\alpha = 1$, we then get

$$(x_1 B_0)^3 + y_1^3 = z_1^3,$$

which is impossible by the nonexistence of a positive integer solution to Fermat's equation for the exponent 3.

We rewrite equation (4.2) as

$$y_1^3 = z_1^3 - x_1^3 B = z_1^3 - (x_1 B_0)^3\alpha = (z_1 - x_1 B_0 \alpha^{1/3})(z_1^2 + x_1 z_1 B_0 \alpha^{1/3} + x_1^2 B_0^2 \alpha^{2/3}),$$

which together with the fact that $B \geq y_1^3$ gives

$$\left|\frac{z_1}{x_1 B_0} - \alpha^{1/3}\right| \ll \frac{y_1^3}{x_1^3 B_0^3} \ll \frac{y_1^3}{x_1^3 B} \ll \frac{1}{x_1^3}. \tag{4.3}$$

We now rewrite equation (4.2) as

$$x_1^3 B = (z_1 - y_1)(z_1^2 + y_1 z_1 + y_1^2).$$

It is well–known and easy to prove that the only prime factors of $z_1^2 + y_1 z_1 + y_1^2$ are all congruent to 1 modulo 3, except that the prime 3 might divide $y_1^2 + y_1 z_1 + z_1^2$. In this case, we have $3 \| (y_1^2 + y_1 z_1 + z_1^2)$. Since all prime factors of B are also prime factors of b, it follows that $\gcd(B, y_1^2 + y_1 z_1 + z_1^2) \leq 3$. Hence, $z_1 - y_1 \gg B$.

We next observe that B tends to infinity. Indeed, if B is bounded, then so is y_1, and from the properties of solutions of Thue equations (see the proof of Propositon (3.2)), it follows that we may assume that x_1 and z_1 are also bounded. Since $d_1 | x_1$, it follows that d_1 is bounded. Finally, since $Bd_2^3 = b^m$ and d_2^3 is not divisible by b, it follows that m has to be bounded. Thus, only finitely many solutions can be obtained this way.

We now distinguish two cases:

Case 1. *There exist infinitely many solutions satisfying $z_1 - y_1 > Bx_1^{1/2}$.*

In this case, $x_1^3B=(z_1-y_1)(z_1^2+y_1z_1+y_1^2)\gg B^3x_1^{3/2}$, and we therefore conclude that $x_1\gg B^{4/3}\gg B_0^4$. Hence, the inequality

$$\frac{\log(x_1^3)}{\log(x_1B_0)}\gg\frac{12}{5}(1+o(1))$$

holds as $B\to\infty$, therefore the inequality

$$\frac{\log(x_1^3)}{\log(x_1B_0)}\geq\frac{11}{5}$$

holds for all but finitely many of those solutions. Inequality (4.3) now shows that these solutions fulfill

$$\left|\frac{z_1}{x_1B_0}-\alpha^{1/3}\right|\ll\frac{1}{x_1^3}\ll\frac{1}{(x_1B_0)^{11/5}},\tag{4.4}$$

but since $\alpha^{1/3}$ is algebraic, the above inequality can have only finitely many positive solutions (x_1,B_0,z_1) by Roth's theorem on rational approximations of algebraic numbers. This is a contradiction.

Case 2. *There exist infinitely many solutions satisfying $z_1-y_1<Bx_1^{1/2}$.*

In this case, since $x_1^3B=(z_1-y_1)(z_1^2+y_1z_1+y_1^2)$, we get $x_1^3B\ll x_1^{3/2}B^3$, therefore $x_1\ll B_0^4$. Note that we also know that $z_1-y_1\gg B$, therefore $x_1^3B\gg B^3$, which gives $x_1\gg B^{2/3}\gg B_0^2$. Thus, $B_0^2\ll x_1\ll B_0^4$. Inequality (4.3) now shows that

$$\left|\frac{z_1}{x_1B_0}-\alpha^{1/3}\right|\ll\frac{1}{x_1^3}=\frac{1}{x_1^2}\cdot\frac{1}{x_1}\ll\frac{1}{(x_1B_0)^2},\tag{4.5}$$

and since $B_0^2\ll x_1\ll B_0^4$ and B_0 is divisible only by prime factors of b, Ridout's extension of Roth's theorem (see [4]) shows that the above inequality (4.5) can have only finitely many positive integer solutions (x_1,B_0,z_1), which is again a contradiction.

Proposition (4.1) follows now from the argument used in the proof of Proposition (3.2). □

The following proposition shows that it is possible to find infinitely many bases b satisfying the conditions of the above Proposition (4.1) but not the condition (ii) of Corollary (3.3).

PROPOSITION (4.6). *There exist infinitely many primes $b\equiv 2 \pmod 3$ such that the Mordell-Weil group of $C_3(b,1)$ is of positive rank.*

Proof. We start by constructing infinitely many primes $b\equiv 2 \pmod 3$ of the form $b=v-u$, where u and v are coprime positive integers such that $u^2+uv+v^2=w^3$ for some positive integer w. To construct such numbers, let θ be any primitive root of order 3 of 1 and note that

$$u^2+uv+v^2=(u-\theta v)(u-\theta^2v).$$

Thus, if we choose u and v such that $u-\theta v=(r-\theta s)^3$ holds with some positive integers r and s, then the relation $u^2+uv+v^2=w^3$ is satisfied with the value $w=(r-\theta s)(r-\theta^2s)=r^2+rs+s^2$. The equation

$$u-\theta v=(r-\theta s)^3=r^3-3\theta r^2s+3\theta^2rs^2-\theta^3s^3,$$

together with the relations $\theta^2 = 1-\theta$ and $\theta^3 = 1$, gives

$$u - v\theta = (r^3 + 3rs^2 - s^3) - \theta(3r^2s + 3rs^2);$$

hence, $u = r^3 + 3rs^2 - s^3$, and $v = 3r^2s + 3rs^2$; therefore, $v - u = -r^3 + 3r^2s + s^3$. Let $f(X,Y) = -X^3 + 3X^2Y + Y^3$. Clearly, $f(X,Y)$ is a primitive irreducible cubic form. Let $X = 2+3M$ and $Y = 1+3N$. It is easy to check that $f_1(M,N) = f(2+3M, 1+3N)$ is also an irreducible cubic polynomial which is primitive and such that there does not exist any prime number p dividing $f(m,n)$ for all positive integers m and n (in fact, we have that $f_1(-1,0) = f(-1,1) = 5$ and $f_1(1,0) = f(5,1) = -49$ are coprime). Thus, the conditions of Theorem 1 in [3] are fulfilled, and we conclude that there exist infinitely many primes $b = f_1(m,n) = f(2+3m, 1+3n)$. Clearly, all such primes are congruent to $-2^3 + 1^3 \equiv 2 \pmod 3$. Note further that $u + v = r^3 + 6rs^2 + 3r^2s - s^3$, and by the well-known results on the Thue-Mahler equation (see [6]), it follows that for all but finitely many of the primes $b = v - u$ that we have just constructed, the largest prime factor of $u + v$ exceeds 5.

We now note that $bw^3 = b(u^2 + uv + v^2) = v^3 - u^3$, therefore with $x = v/u$, $y = w/u$ we get a rational solution (x, y) to the equation

$$by^3 = (x-1)(x^2 + x + 1) = x^3 - 1. \tag{4.7}$$

Thus, $(x, y) = (v/u, w/u)$ is also a point on $\mathcal{C}_3(b,1)$. We now check that this point is of infinite order on $\mathcal{C}_3(b,1)$. Performing the following substitutions

$$by^3 = (x-1)(x^2 + x + 1) = (x-1)((x-1)^2 + 3(x-1) + 3);$$

with $y_1 = y/(x-1)$ and $x_1 = 1/(x-1)$, we get

$$by_1^3 = 3x_1^2 + 3x_1 + 1; \qquad \text{multiplying by } 3\cdot 4,$$

we get

$$3\cdot 4\cdot by_1^3 = (6x_1 + 3)^2 + 3; \qquad \text{with } x_2 = 6x_1 + 3,$$

we get

$$3\cdot 4\cdot by_1^3 - 3 = x_2^2; \qquad \text{multiplying by } 3^2\cdot 4^2\cdot b^2,$$

we get

$$(3\cdot 4\cdot by_1)^3 - 2^4\cdot 3^3\cdot b^2 = (3\cdot 4\cdot bx_2)^2;$$

so, we see that the given curve (4.7) is birationaly equivalent to

$$Y^2 = X^3 - 2^4\cdot 3^3\cdot b^2, \tag{4.8}$$

via the birational transformations

$$Y = 3\cdot 4\cdot b\left(\frac{6}{x-1} + 3\right) \quad \text{and} \quad X = 3\cdot 4\cdot b\left(\frac{y}{x-1}\right).$$

With $b = v - u$, the point $(x, y) = (v/u, w/u)$ transforms into the point of coordinates $(X, Y) = (3\cdot 4\cdot w, 3^2\cdot 4\cdot(u+v))$. Clearly, the discriminant of the curve (4.8) is divisible only by the primes 2, 3 and b. Since u and v are coprime, it follows that b does not divide $u + v$, and we know that $u + v$ has a prime factor exceeding 5. Hence, $Y = 3^2\cdot 4\cdot(u+v)$ is a multiple of some prime which does not divide the discriminant of the curve (4.8). Via the Lutz-Nagell Theorem (see page 221 in [7]), we immediately get that the above point (X, Y) is not a point of finite order on the curve (4.8). Thus, the Mordell-Weil group of $\mathcal{C}_3(b,1)$ has positive rank. $\square$

EXAMPLE (4.9). *Taking $b = 10$ and $r = 1,2$, we see from the previous calculations that $\mathcal{C}_3(b,r)$ is birationally equivalent to $Y^2 = X^3 - 2^4 \cdot 3^3 \cdot (10^r)^2$ for $r = 1$, 2, which in turn gives us curves birationally equivalent to*

$$Y^2 = X^3 - 3^3 \cdot 5^2 \quad \textit{and} \quad Y^2 = X^3 - 3^3 \cdot 2^2 \cdot 5^4,$$

if $r = 1$, or 2, respectively. These curves are 2700 U 2 and 2700 T 2, respectively, in Cremona's tables (see [2]), and they both have rank 0 and trivial torsion. This shows that there is no solution of equation (1.1) when $n = 3$ in base 10.

We finally prove that equation (1.1) has only finitely many reduced solutions with $n = 3$ for all positive integers $b > 1$. For any pair of rational numbers (X,Y) we write $h(X,Y)$ for the maximum of the absolute values of the numerators and denominators of X and Y. For a finite set of prime numbers $\mathcal{P}$ and a nonzero integer n, we write $n_{\mathcal{P}}$ for the largest divisor of n composed of primes from $\mathcal{P}$. The following result is true for all elliptic curves but we shall need it only for our curves $\mathcal{C}_3(b,r)$.

THEOREM (4.10). *Let $\varepsilon > 0$ be fixed. There are only finitely many rational points $(X,Y) \in \mathcal{C}_3(b,r)$ such that if we write $X = x/z$ with coprime integers x and z, then $|x_{\mathcal{P}}| > h(X,Y)^{\varepsilon}$.*

Proof. This follows immediately from the Theorem on page 101 of [5]. □

THEOREM (4.11). *Let $b > 1$ be an arbitrary integer. Then equation (1.1) has only finitely many reduced solutions (x,y,z).*

Proof. We fix $r \in \{1,2\}$. Let (x,y,z) be a reduced solution of (1.1) such that $(X,Y) = (x/z, y/z) \in \mathcal{C}_3(b,r)$. We write $X = xb^{(m-r)/3}/z = x_1/z_1$ and $Y = y/z = y_1/z_2$, with $\gcd(x_1,z_1) = \gcd(y_1,z_2) = 1$. We note that $z_2 \mid z_1 \mid z_2 b^2$, therefore $z \asymp z_1 \asymp z_2$. Furthermore, $b^m > y^3$, therefore $X > Y$. Since $b^r X^3 + Y^3 = 1$, we get that $X \asymp 1$. It now follows that $h(X,Y) \asymp z_1$. Since (x,y,z) is reduced, there exists a prime $p|b$ such that if $p^{\alpha_p}||b$, then p^{α_p} does not divide y. In particular, the exponent at which p appears in $\gcd(xb^{(m-r)/3}, z)$ is $O(1)$. Hence, if we write $\mathcal{P} = \{p : p|b\}$, then $|x_1|_{\mathcal{P}} \gg p^{m/3}$. By Theorem (4.10), $p^{m/3} \ll (h(X,Y))^{\varepsilon} \ll z_1^{\varepsilon}$ holds with finitely many exceptions, where $\varepsilon > 0$ is a small number which we will fix later. Hence, with finitely many exceptions, we have

$$y_1^3 \le y^3 \le b^m \ll z_1^{3\varepsilon \log b/\log p}.$$

We choose ε such that $3\varepsilon \log b/\log p < 1/2$ holds for all primes $p|b$. Then $y_1^3 \ll \sqrt{z_1}$. We now rewrite $b^r X^3 + Y^3 = 1$ as

$$b^r \left(\frac{x_1}{z_1}\right)^3 + \left(\frac{y_1}{z_2}\right)^3 = 1,$$

therefore

$$b^r x_1^3 - z_1^3 = \left(\frac{z_1}{z_2}\right)^3 y_1^3 \ll y_1^3 \ll z_1^{1/2},$$

therefore

$$\left|\frac{x_1}{z_1} - \frac{1}{b^{r/3}}\right| \ll \frac{\sqrt{z_1}}{z_1^3} \ll \frac{1}{z_1^{2.5}}. \tag{4.12}$$

Since $r = 1,2$ and b is not a perfect cube (otherwise there are no points (X,Y) with $XY \neq 0$ on $\mathcal{C}_3(b,r)$), it follows that $b^{r/3}$ is irrational. Roth's theorem once again tells us that there are only finitely many positive solutions (x_1, z_1) to the inequality (4.12). □

5. Heuristics in the case $n = 3$

The starting point of this section is the following example.

EXAMPLE (5.1). *Note that* $(u,v) = (1,18)$ *satisfies* $u^2+uv+v^2 = w^3$ *with the value* $w = 7$. *Furthermore,* $b = v - u = 17$ *is a prime congruent to* 2 *modulo* 3 *and since* $u + v = 19$ *is prime, it follows that the Mordell-Weil group of* $C_3(17,1)$ *has positive rank. Notice also that the above example actually provides a primitive solution of equation* (1.1) *namely* $\overline{7^3 1^3}_{(17)} = 18^3$.

The above example together with Propositions (4.1) and (4.6) suggest the following question:

QUESTION (5.2). *Do there exist infinitely many* $b > 1$ *such that no prime factor of* b *is congruent to* 1 *modulo* 3 *and such that equation* (1.1) *has at least one reduced solution* (x, y, z)*?*

That is, by Proposition (4.1), we know that for each such b equation (1.1) has only finitely many reduced solutions, and Proposition (4.6) showed us how to construct, for infinitely many such b, a "canonical" point on $C_3(b,1)$ which is not a point of finite order in the Mordell-Weil group. Furthermore, the above example showed us that such a point can actually lead to a primitive solution of equation (1.1), so the above question can loosely be reformulated by asking whether or not one should expect to be able to create infinitely many examples of reduced solutions (x, y, z, b) of (1.1) by this method.

We offer the following conjecture.

CONJECTURE (5.3). *There exist infinitely many quintuples of positive integers* (x, y, z, b, m) *such that no prime factor of* b *is congruent to* 1 *modulo* 3 *and such that* (x, y, z) *is a primitive solution to equation* (1.1) *with* $n = 3$.

In the remaining of this section, we offer some heuristical support in favor of Conjecture (5.3). We use the notation from the beginning of the proof of Proposition (4.6). Namely, we let again r and s be coprime positive integers, and put $u = r^3 + 3rs^2 - s^3$, $v = 3rs(r + s)$, $b = v - u = -r^3 + 3r^2 s + s^3$. We want that $b > 0$ and $u > 0$. With $x = w$, $y = u$, $z = v$, we get $x^3 b + y^3 = z^3$. In order for this to be a solution of equation (1.1), we want that $b > y^3$. Assume that $4r > s > 3r$. Then $b = (s - r)(s^2 + rs + r^2) + 3r^2 s > s^3/2$, so it suffices that

$$\frac{s^3}{2} > y^3 = (r^3 + 3rs^2 - s^3)^3 > 0,$$

which is equivalent to

$$\frac{1}{2^{1/3}} \cdot \frac{1}{s^2} > \left(\frac{r}{s}\right)^3 + 3\left(\frac{r}{s}\right) - 1 > 0. \tag{5.4}$$

Let $\alpha \in (1/4, 1/3)$ be the only real root of $f(x) = x^3 + 3x - 1$ and let β and $\overline{\beta}$ be the other two (complex conjugate) roots of this polynomial. Clearly, $|\beta| = \alpha^{-1/2} < 2$. Since

$$0 = \left|\frac{r}{s} - \beta\right|^2 = \left(\frac{r}{s} - \beta\right) \cdot \left(\frac{r}{s} - \overline{\beta}\right) < \left(|\beta| + \frac{r}{s}\right)^2 < \left(2 + \frac{1}{3}\right)^2 = \frac{49}{9},$$

it follows that inequality (5.4) is implied by

$$\frac{9}{49\cdot 2^{1/3}}\cdot\frac{1}{s^2} > \frac{r}{s} - \alpha > 0. \tag{5.5}$$

Since $9/(49\cdot 2^{1/3}) > 1/7$, inequality (5.5) is implied by

$$\frac{1}{7s^2} > \frac{r}{s} - \alpha > 0. \tag{5.6}$$

Let $[a_0, a_1, \ldots]$ be the continued fraction expansion of α, and for $h \ge 0$ put p_h/q_h for the sequence of convergents of α. Since both inequalities

$$(-1)^{h+1}\left(\frac{p_h}{q_h} - \alpha\right) > 0,$$

and

$$\left|\frac{p_h}{q_h} - \alpha\right| < \frac{1}{q_h q_{h+1}} = \frac{1}{q_h(a_{h+1}q_h + q_{h-1})}$$

hold for all $h \ge 0$, it follows that inequality (5.6) holds whenever $r/s = p_h/q_h$ with some odd h such that $a_{h+1} \ge 7$. Let $\mathcal{H}$ be the subset of those odd integers. We *conjecture* that $\mathcal{H}$ contains a positive proportion of all odd integers. It is clear that if h is large then $r/s \in (1/4, 1/3)$. Thus, the condition $u > 0$ holds from the fact that h is odd, and the condition $b > 0$ is equivalent to $-r^3 + 3r^2 s + s^3 > 0$, which holds because r and s are positive and $r/s < 1/3$. It remains to justify that b can be taken to be free of primes congruent to 1 modulo 3.

Well, clearly $b = f(p_h, q_h) = -p_h^3 + 3p_h^2 q_h + q_h^3$. The discriminant of the polynomial $h(x) = -x^3 + 3x^2 + 1$ is -135, so its Galois group is S_3. For a prime p let $\rho(p)$ be the number of roots of $h(x)$ modulo p. A short calculation in S_3 via the Chebotarev density theorem shows that the proportion of primes p for which $\rho(p)$ equals 0, 1 and 3 is 1/3, 1/2 and 1/6, respectively. Since the cube root of unity does not belong to the splitting field of $h(x)$, by the Chebotarev density theorem again, the same proportions hold when p is restricted to primes which are congruent to 1 (mod 3). In particular, for large X, we have

$$\prod_{\substack{p\le X\\ p\equiv 1 \pmod 3}} \left(1 - \frac{\rho_p}{p}\right) \gg \frac{1}{\sqrt{\log X}}.$$

By the sieve, given a random pair (r, s) with $r/s \in (1/4, 1/3)$, the probability that $f(r, s)$ is not divisible by primes which are congruent to 1 modulo 3 should be about proportional to $1/\sqrt{\log(r+s)}$. Maybe this suggests that the "probability" that the number $b = f(p_h, q_h)$ is free of primes congruent to 1 modulo 3 should be $\gg 1/(\log p_h)^{1/2}$. Finally, we *conjecture* that

$$\sum_{h\in\mathcal{H}} \frac{1}{(\log p_h)^{1/2}}$$

is divergent (in fact, we conjecture this to be the case when $\mathcal{H}$ is *any* subset of the positive integers having positive lower density). Such a statement does not follow from Roth's Theorem, but it would follow assuming that the continued fraction of α behaves like the continued fraction of "most real numbers" in the sense of Khintchine's theory (the statement about $\mathcal{H}$ containing a positive proportion of odd positive integers h would also hold under such assumption).

We believe that the above heuristics do seem to support Conjecture (5.3).

Acknowledgements

We thank the referee for a careful reading of the paper and for comments and remarks which improved its quality. Furthermore, the referee noted that while our heuristics from Section 5 concerning values of $f(r,s)$ which are free of primes congruent to 1 modulo 3 seem fine when applied to random pair (r,s), these heuristics might not hold for $f(p_h,q_h)$ as such pairs of integers are quite special. In fact, the referee did compute the first few hundred convergents of α and always found a prime factor congruent to 1 modulo 3 in the corresponding value of b. While these computations on such a small scale do not refute our conjecture, they do cast some doubts on our heuristics. We only hope that this topic will be pursued in the future by other researchers. This work was supported in part by Projects PAPIIT IN104512, CONACyT Mexico–France 193539, CONACyT Mexico–India 163787, and a Marcos Moshinsky Fellowship.

Received February 12, 2013

Final version received April 16, 2013

FLORIAN LUCA
FUNDACIÓN MARCOS MOSHINSKY, UNAM
CIRCUITO EXTERIOR, C.U.,
APDO. POSTAL 70-543
MÉXICO D.F. 04510, MÉXICO
fluca@matmor.unam.mx

REFERENCES

[1] W. D. BANKS AND F. LUCA, *Concatenations with binary recurrent sequences*, J. Integer Seq. **8**, (2005), no. 1, Article 05.1.3 (electronic).

[2] J. CREMONA, Elliptic curve data, available on line at http://www.maths.nott.ac.uk/personal/jec/ftp/data/INDEX/html.

[3] D. R. HEATH-BROWN AND B. Z. MOROZ, *Representation of primes by cubic polynomials in two variables*, Proc. London Math. Soc. **88**, (2004), 289–312.

[4] D. RIDOUT, *Rational approximations to algebraic numbers*, Mathematika **4**, (1957), 125–131.

[5] J.-P. SERRE, Lectures on the Mordell-Weil Theorem, Aspects of Mathematics. Friedr. Vieweg & Sohn, Braunschweig, 1989.

[6] T. N. SHOREY AND R. TIJDEMAN, Exponential diophantine equations, Cambridge University Press, 1986.

[7] J. SILVERMAN, The arithmetic of elliptic curves, Springer-Verlag, New York, 1986.

Bol. Soc. Mat. Mexicana (3) Vol. 19, 2013

ON TOPOLOGICAL RIGIDITY OF PHASE-PORTRAITS IN THE COMPLEX PLANE

M. TEYMURI GARAKANI, A. MAFRA, AND B. SCÁRDUA

ABSTRACT. We study phase-portraits of autonomous ordinary differential equations with polynomial coefficients in the complex plane and address the following question: *Is it true that for generic choices of coefficients the equation is completely determined, up to analytic equivalence, by the topological equivalence class of its phase-portrait?* This question was first considered by Y. Ilyashenko in his landmark work [13]. Our results give a contribution to this problem in the framework of holomorphic foliations in the complex projective plane $\mathbf{CP}^2$ under the hypothesis of topological triviality in $\mathbb{C}^2$. We also give some description of the non-rigid foliations with generic singularities as Darboux foliations.

1. Introduction

This paper is mainly concerned with the study of phase-portraits of polynomial autonomous ordinary differential in the complex plane $\mathbb{C}^2$. We follow the viewpoint introduced by Ilyashenko in [13]. Denote by $\dot{x} = P(x,y)$, $\dot{y} = Q(x,y)$ the polynomial ODE that we consider. The basic philosophy/idea is that *for generic choice of P and Q the corresponding ODE is completely determined, up to analytic equivalence, by its phase-portrait, more precisely, by the topological equivalence class of it*. This can be put in simple words as: *Given generic P and Q and other polynomials P_1, Q_1 then the ODEs* $(*)\, dy/dx = P/Q$ *and* (1), $dy/dx = P_1/Q_1$ *are analytically equivalent if and only if they are topologically equivalent*. In turn, the topology of plane complex flows is a quite well-developed subject and has been considered, either in the local and global aspects, by several authors ([2, 4, 10]). Thus a positive confirmation of the above idea should have important consequences in the study of the above mentioned ODEs.

Note that so far we place ourselves on the complex plane $\mathbb{C}^2$ and we speak of *equivalence* instead of *conjugacy*. The first striking work addressing this problem is due to Ilyashenko ([13]), where he proves the above philosophy is correct, but once we consider the equivalence (topological or analytical) being defined on the (compact) projective plane $\mathbf{CP^2}$, and for the case the differential equation (1) embeds into an analytical deformation of $(*)$. In order to give a more detailed description on Ilyashenko's landmark result we shall introduce some ingredients (foliations with singularities, holonomy,...) somehow beyond the original context. As we shall mention below (see the last paragraph of this introduction), our results are somehow paving the way to the original question, as it will be clear from our statements.

In this work we will be mainly concerned with holomorphic foliations with singularities in $\mathbf{CP}^2$. Any such foliation is induced by a polynomial vector field (or

2010 *Mathematics Subject Classification*: Primary: 32M25, 32S65; Secondary: 26A18, 57R30.
Keywords and phrases: holomorphic foliation, topological rigidity, solvable group, holonomy group.

polynomial one-form) on any affine subspace $\mathbb{C}^2 \subset \mathbf{CP}^2$. Conversely, any nontrivial polynomial vector field or 1-form on $\mathbb{C}^2$ induces a foliation which extends to $\mathbf{CP^2}$ as a holomorphic foliation with singularities ([17]). The study of these foliations is motivated by Hilbert Sixteenth Problem on the number and position of limit cycles of polynomial differential equations in the real plane $\mathbb{R}^2$. A major attempt in this line was started in 1956 by a seminal work of I. Petrovski and E. Landis [21]. They consider the real equation as a differential equation in the complex plane $\mathbb{C}^2$, and the time t is now a complex time parameter. The integral curves of the vector field are now either singular points which correspond to the common zeros of P and Q, or complex curves tangent to the vector field which are holomorphically immersed in $\mathbb{C}^2$. This gives rise to a holomorphic foliation by complex curves with a finite number of singular points. One can easily see that this foliation extends to the complex projective plane $\mathbf{CP}^2$, which is obtained by adding an infinite line to the plane $\mathbb{C}^2$. Conversely any holomorphic foliation by curves on $\mathbf{CP}^2$ is given in an affine space $\mathbb{C}^2 \hookrightarrow \mathbf{CP}^2$ by a polynomial vector field $X = (P,Q) \in \mathfrak{X}(\mathbb{C}^2)$ with $\gcd(P,Q) = 1$.

Although they did not completely solve this problem, they introduced a truly novel method in geometric theory of ordinary differential equations. In 1978, Il'yashenko made a first fundamental contribution to the problem. Following the general idea of Petrovski and Landis, he studied complex polynomial equations in the plane from a topological standpoint without particular attention to Hilbert's question. On what follows we shall describe some of his results and related works. Let us first introduce some notation as it appears in [14]. Denote by $\mathfrak{U}(n)$ the set of all foliations of the complex projective plane given by a polynomial vector field of degree at most n in a fixed affine neighborhood $\mathbb{C}^2 \subset \mathbf{CP}^2$. The equations of class $\mathfrak{U}(n)$ with $n+1$ infinite singular points form, by definition, a subclass $\mathfrak{U}(n)'$. Notice that for algebraic reasons, a foliation in the class $\mathfrak{U}(n)$ leaves invariant the infinite line $L_\infty = \mathbf{CP}^2 \setminus \mathbb{C}^2$.

(1.1) Ilyashenko's absolutely rigidity. A striking result of Y. Ilyashenko states topological rigidity for a residual set of foliations on $\mathfrak{U}(n)'$ if $n \geq 2$.

Definition (1.1) (cf. [14] Definition 3). A foliation $\mathcal{F}$ of class $\mathfrak{U}(n)$ is said to be *absolutely rigid* if there exist a neighborhood of the foliation $\mathcal{F}$ in the class $\mathfrak{U}(n)$ and a neighborhood of the identity homeomorphism in the space of all self-homeomorphisms of the complex projective plane such that any foliation in the former neighborhood that is conjugate to $\mathcal{F}$ by a homeomorphism from the latter neighborhood is affine equivalent to $\mathcal{F}$.

Using this terminology Ilyashenko's rigidity result in [13] can be stated as:

THEOREM (1.2) ([14]). *Any generic foliation of class $\mathfrak{U}(n)$ is absolutely rigid.*

The genericity conditions are stated in Ilyashenko's original work, but we can mention that they eliminate a dense subset of equations in a sense that will be made clear later. We stress that it is required that the foliation leaves invariant the infinite line. As remarked in [14], Shcherbakov [24, 25] and other authors have reduced this exceptional set. Scherbakov's result can be stated as in [14] as follows:

THEOREM (1.3). *For $n \geq 2$ the space $\mathfrak{U}(n)$ contains a real algebraic subset Σ_n and a nowhere dense real analytic subset Σ'_n of real codimension at least 2 such that each foliation $\mathcal{F}$ in the set $\mathfrak{U}(n) \setminus (\Sigma_n \cup \Sigma'_n)$ is absolutely rigid and has dense leaves in $\mathbb{C}^2$. For $n \geq 3$ the foliation exhibits only a countable set of homologically independent complex limit cycles.*

Very sharp absolute rigidity results for foliations defined by quadratic vector fields can be found in [15]. Also, another important reference for rigidity results is the book [16], where some versions of [15] are proved as well.

(1.2) Deformations and rigidity. Let us change now our point of view. Let M be a complex surface. By a *holomorphic foliation with singularities* on M we mean a pair $\mathcal{F} = (\mathcal{F}_o, \text{sing}(\mathcal{F}))$ where $\text{sing}(\mathcal{F}) \subset M$ is a discrete subset of M and $\mathcal{F}_o$ is a one-dimensional holomorphic foliation on the open manifold $M \setminus \text{sing}(\mathcal{F})$ in the ordinary sense. The set $\text{sing}(\mathcal{F})$ is the *singular set* of $\mathcal{F}$ and by a *leaf* of $\mathcal{F}$ we shall mean a leaf of the underlying regular foliation $\mathcal{F}_o$. Let $\text{Fol}(M)$ denote the space of holomorphic foliations on M. An *analytic deformation* of $\mathcal{F} \in \text{Fol}(M)$ is an analytic family $\{\mathcal{F}_t\}_{t\in Y}$ of foliations on M, with parameters on an analytic space Y, such that there exists a point "0" $\in Y$ with $\mathcal{F}_0 = \mathcal{F}$. Here we will only consider deformations where $Y = \mathbb{D} \subset \mathbb{C}$ is the unit disk. A *topological equivalence* (resp. *analytical equivalence*) between two foliations $\mathcal{F}_1$ and $\mathcal{F}_2$ is a homeomorphism (resp. biholomorphism) $\phi : M \to M$, which takes leaves of $\mathcal{F}_1$ onto leaves of $\mathcal{F}_2$, and such that $\phi(\text{sing}(\mathcal{F}_1)) = \text{sing}(\mathcal{F}_2)$. The deformation $\{\mathcal{F}_t\}_{t\in\mathbb{D}}$ is *topologically trivial* (resp. *analytically trivial*) if there exists a continuous map (resp. holomorphic map) $\phi : M \times \mathbb{D} \to M$, such that each map $\phi_t = \phi(.,t) : M \to M$ is a topological equivalence (resp. an analytical equivalence) between $\mathcal{F}_t$ and $\mathcal{F}_0$ and $\phi_0 = \text{Id}$.

Let $\mathcal{C} \subset \text{Fol}(M)$ be a *class of foliations* on M, i.e., a subset of $\text{Fol}(M)$. A foliation $\mathcal{F}_0 \in \mathcal{C}$ is *topologically rigid under deformations in the class* $\mathcal{C}$ if any topologically trivial deformation $\{\mathcal{F}_t\}_{t\in\mathbb{D}}$ of $\mathcal{F}_0$ with $\mathcal{F}_t \in \mathcal{C}$ is analytically trivial.

Let now $U \subset M$ be an open subset. We also say that $\mathcal{F}_0 \in \mathcal{C}$ is *U-topologically rigid under small deformations in the class* $\mathcal{C}$, if any analytic deformation $\{\mathcal{F}_t\}_{t\in\mathbb{D}}$ of $\mathcal{F}_0$ with $\mathcal{F}_t \in \mathcal{C}$, $\forall t$; which is topologically trivial in the open subset U, is in fact analytically trivial in M for $|t|$ small enough. If we may take $U = M$ we say that the foliation is *topologically rigid under small deformations in the class* $\mathcal{C}$. Notice that the word *small* appears to indicate that we may have to restrict the deformation parameter t to a small disk centered at the origin.

The space $\text{Fol}(\mathbf{CP}^2)$ of foliations with singularities on $\mathbf{CP}^2$ can be stratified as $\text{Fol}(\mathbf{CP}^2) = \bigcup_{n=1}^{\infty} \text{Fol}(n)$, where $\text{Fol}(n)$ denotes the class of foliations of *degree* $n \geq 1$ (see Section 2). We fix the infinite line $L_\infty = \mathbf{CP}^2 \setminus \mathbb{C}^2$ and denote by $\mathfrak{X}(n) \subset \text{Fol}(n)$ the space of foliations of degree $n \in \mathbb{N}$ which leave invariant L_∞.

In this terminology we can re-state Ilyashenko's result as:

THEOREM (1.4) ([13]). *For any $n \geq 2$ there exists a residual subset $\mathfrak{I}(n) \subset \mathfrak{X}(n)$ whose foliations are topologically rigid under deformations in the class $\mathfrak{X}(n)$. In particular the foliations in $\mathfrak{I}(n)$ are absolutely rigid.*

A variant of this result is found in [18] and may be written as follows:

THEOREM (1.5) ([18]). *For each $n \geq 2$, $\mathfrak{X}(n)$ contains an open dense subset $\mathfrak{R}(n) \subset \mathfrak{X}(n)$ whose foliations are topologically rigid under deformations in the class $\mathfrak{X}(n)$.*

Again, foliations in the class $\mathfrak{R}(n)$ are absolutely rigid in the sense of Ilyashenko. We stress the fact that in both theorems above we consider deformations $\{\mathcal{F}_t\}_{t\in\mathbb{D}}$ in the class $\mathfrak{X}(n)$, that is, $\mathcal{F}_t$ leaves invariant $L_\infty, \forall t \in \mathbb{D}$; and we assume topological triviality in $\mathbf{CP}^2$. This last hypothesis is slightly relaxed by requiring topological triviality for the set of separatrices through the singularities at L_∞:

Given $\mathcal{F} \in \mathfrak{X}(n)$ denote by $\mathrm{Sep}(\mathcal{F})$ the (germ of the) set of local separatrices of $\mathcal{F}$ transverse to the singularities of $\mathcal{F}$ in L_∞. A deformation $\{\mathcal{F}_t\}_{t\in\mathbb{D}}$ of $\mathcal{F} = \mathcal{F}_0$ is *s-trivial* if there exists a continuous family of maps $\phi_t : \mathrm{Sep}(\mathcal{F}_0) \to \mathbb{C}^2$ such that ϕ_0 is the inclusion map and ϕ_t is a continuous injection map from $\mathrm{Sep}(\mathcal{F}_0)$ to $\mathbb{C}^2$ with $\phi_t(\mathrm{Sep}(\mathcal{F}_0)) = \mathrm{Sep}(\mathcal{F}_t)$ ([18]). A foliation $\mathcal{F}_0 \in \mathrm{Fol}(n)$ *is s-rigid under deformations in the class* $\mathfrak{X}(n)$ if any s-trivial deformation $\{\mathcal{F}_t\}_{t\in\mathbb{D}} \subset \mathfrak{X}(n)$ of $\mathcal{F}_0$ is analytically trivial.

THEOREM (1.6) ([18]). *For any $n \geq 2$, $\mathfrak{X}(n)$ contains an open dense subset $\mathfrak{s}\mathfrak{R}(n)$ whose foliations are s-rigid under deformations in the class $\mathfrak{X}(n)$.*

REMARK (1.7). As it is proved in [18], topological triviality in $\mathbf{CP}^2$ implies s-triviality. We shall prove (cf. Proposition (3.13) and Corollary (3.14)) that, under generic conditions on (the singular set of) the foliation, s-triviality is also a consequence of topological triviality in $\mathbb{C}^2$.

Theorem A. *Given $n \geq 2$ there exists an open dense subset $\mathrm{Rig}(n) \subset \mathfrak{X}(n)$ such that any foliation $\mathcal{F}$ in $\mathrm{Rig}(n)$ is $\mathbb{C}^2$-topologically rigid under <u>small</u> deformations in the class $\mathrm{Fol}(n)$: any analytic deformation $\{\mathcal{F}_t\}_{t\in\mathbb{D}}$ of $\mathcal{F} = \mathcal{F}_0$, with $\mathcal{F}_t \in \mathrm{Fol}(n), \forall t$, which is topologically trivial in $\mathbb{C}^2$, must be analytically trivial in $\mathbf{CP}^2$ for $|t| < \epsilon$ small enough.*

Theorem A is strongly connected with the absolute rigidity result in Theorem 1 from Ilyashenko's 2007 paper [13] (see page 63). Nevertheless, Ilyashenko's result holds under the hypothesis that the topological conjugacy is of a bounded *complexity*, which is a notion introduced in this same paper. Theorem 1 in [15] and our Theorem A above are therefore of different nature.

We shall also mention that Theorem A is similar to local results in [7], in the sense that in principle, there is no a topological conjugation along an invariant projective line. In [7] the authors prove a similar topological rigidity, for suitable deformations of certain germs of holomorphic foliations, that can be reduced with a single non-dicritical blow-up, and exhibiting non-solvable holonomy group for the exceptional projective line divisor. Their technique is based on topological rigidity for groups and germs of holomorphic diffeomorphisms. Thus, our results, suggest that a global version of [7] (see Corollary 3 page 248) may be at hand.

Now let us recall an important class of foliations, namely *Darboux foliations*, which appears in the statement of our next theorem. Let M be a complex manifold and let $f_j : M \to \overline{\mathbb{C}}$ be meromorphic functions and $\lambda_j \in \mathbb{C}^*$ complex numbers, $j =$

$1, \cdots, r$. The meromorphic integrable 1-form

$$\Omega = (\prod_{j=1}^{r} f_j) \sum_{j=1}^{r} \lambda_i \frac{df_i}{f_i}$$

defines a *Darboux* foliation $\mathcal{F}$ on M. The foliation $\mathcal{F}$ has $f = \prod_{j=1}^{r} f_j$ as a logarithmic first integral.

Our next result hints a description of the complementary class $\mathrm{Fol}(n) \setminus \mathrm{Rig}(n)$, i.e., of the class of non-rigid foliations with invariant infinite line and hyperbolic singularities on this line:

Theorem B. *Let $\mathcal{F} \in \mathfrak{X}(n)$ be a foliation with hyperbolic singularities in the infinite line and irreducible singularities in* $\mathbf{CP}^2$*. Then either $\mathcal{F}$ is $\mathbb{C}^2$-topologically rigid for small deformations in the class* $\mathrm{Fol}(n)$ *or $\mathcal{F}$ is a Darboux foliation in* $\mathbf{CP}^2$*.*

This theorem states some kind of dicotomy: non solvable holonomy of the invariant line implies rigidity and solvable holonomy implies a holmorphic integrating factor or a Darbouxian first integral.

REMARK (1.8). Generic Darboux foliations are non-rigid. Indeed, this is related to topological non-rigidity of abelian linearizable finitely generated subgroups of germs of complex diffeomorphisms in one variable.

In few words, a foliation $\mathcal{F} \in \mathfrak{X}(n)$ with hyperbolic singularities in the infinite line and irreducible singularities in $\mathbb{C}^2$, belongs to the class $\mathrm{Rig}(n)$ if and only if the holonomy group of the leaf $L_\infty \setminus (\mathrm{sing}(\mathcal{F}) \cap L_\infty)$ is non-solvable. An isolated singularity is *irreducible* if it is defined by a vector field with non-zero eigenvalues having as quotient a complex number $\lambda \notin \mathbb{Q}_+$ (see Section 2).

REMARK (1.9). No confusion should be made with the notion of irreducible we use and the notion of irreducible singularity appearing in the resolution of singularities (cf. [26]). The last consists of our non irreducible singularities and saddle-nodes as well. In this paper no saddle-nodes are considered, so by irreducible we shall mean non-degenerate and irreducible.

One of the main gains of our results is that we are able to give a description of the class of non-rigid foliations having the infinite line invariant and generic singularities in terms of Darboux foliations. Nevertheless, the main point is related to the motivation mentioned in the first paragraph. Since our deformations are, *a priori*, allowed to move the line L_∞ which is $\mathcal{F}_0$-invariant by hypothesis, and since we assume topological triviality on $\mathbb{C}^2$ (not on $\mathbf{CP}^2$) we can this way state our rigidity result completely in terms of the original ODE $dy/dx = P/Q$ on $\mathbb{C}^2$, proving this way that the mentioned phase-portrait rigidity actually holds for an open dense class of ODEs and we can restate Theorem A in these terms.

The article is organized according to the following plan. Sections 2 contains basic material, definitions and standard properties of holomorphic foliations in the complex projective plane. Section 3 is devoted to two basic results in the proof of Theorem A. In Section 4 we recall results of several authors ([1], [20], [27]) about dynamics, fixed points and topological rigidity of non-solvable finitely generated subgroups of germs of complex diffeomorphisms at the origin $0 \in \mathbb{C}$. The next section is devoted to prove Theorem A. Section 6 is dedicated to the proof of Theorem B. Finally, in Section 7 we prove a kind of Noether's lemma for foliations which assures triviality for analytic integrable deformations under some local triviality hypothesis and conclude with some conjecture. More precisely, we have (cf. Theorem (7.1)):

Theorem C. *Let $\{\mathcal{F}_t\}_{t\in\mathbb{D}}$ be a holomorphic integrable deformation of a foliation $\mathcal{F}_0$ of degree n on $\mathbf{CP}^2$. Assume that for each singularity $p \in \operatorname{sing}(\mathcal{F}_0)$ the germ of integrable deformation at p is analytically trivial. Then there exists $\epsilon > 0$ such that $\{\mathcal{F}_t\}_{|t|<\epsilon}$ is analytically trivial.*

This theorem is comparable to Theorem 5.3 and Corollary 5.7 in [8], where the problem of giving a criterium for an unfolding of a holomorphic foliation with singularities to be holomorphically trivial, is originally studied. In the above mentioned results in [8], the author reaches similar conclusions, based on the triviality of a certain first Cohomology group associated to the line bundle that defines the foliation, with multiplicity one non-degenerate singularities, on a compact complex surface. Thanks to the special geometry of the complex projective spaces, these conditions are essentially verified in our case, so that our results may be seen as a particular case of [8].

2. Preliminaries

Let $\mathcal{F}$ be a (singular) foliation on $\mathbf{CP}^2$ and $L \subset \mathbf{CP}^2$ be a projective line, which is not an algebraic solution of $\mathcal{F}$ ($L \setminus \operatorname{sing}(\mathcal{F})$ is not a leaf of $\mathcal{F}$). We say that $p \in L$ is a tangency point of $\mathcal{F}$ with L, if either $p \in \operatorname{sing}(\mathcal{F})$ or $p \notin \operatorname{sing}(\mathcal{F})$ and the tangent spaces of L and of the leaf of $\mathcal{F}$ through p, at p, coincide. We say that L is invariant by $\mathcal{F}$ if $\forall p \in L \setminus \operatorname{sing}(\mathcal{F})$, p is a tangency point of $\mathcal{F}$ with L. Denote by $T(\mathcal{F}, L)$, the set of tangency points of $\mathcal{F}$ with L. According to [17], if $\operatorname{sing}(\mathcal{F})$ has codimension ≥ 2 or equivalently the singularities of $\mathcal{F}$ are finitely many points in $\mathbf{CP}^2$, then there exists an open, dense and connected subset $NI(\mathcal{F})$ of the set of lines in $\mathbf{CP}^2$, such that every $L \in NI(\mathcal{F})$ satisfies the following properties:

- L is not invariant by $\mathcal{F}$,
- $T(\mathcal{F}, L)$ is an algebraic subset of L defined by a polynomial of degree $n = n(\mathcal{F})$ in L and this number is independent of L.

The integer $n(\mathcal{F})$ is called *the degree of the foliation* $\mathcal{F}$. According to [17], a foliation of degree n in $\mathbf{CP}^2$ can be expressed in an affine coordinate system by a differential equation of the form

$$(P(x,y) + xg(x,y))dy - (Q(x,y) + yg(x,y))dx = 0,$$

where P, Q and g are polynomials such that:

1. $P + xg$ and $Q + yg$ are relatively prime,
2. g is homogeneous of degree n,
3. $\max\{\deg(P), \deg(Q)\} \leq n$,
4. $\max\{\deg(P), \deg(Q)\} = n$ if $g \equiv 0$.

Let B_{n+1} be the space of polynomials of degree $\leq n+1$ in two variables. Let $V \subset B_{n+1} \times B_{n+1}$ be the subspace of pairs of polynomials of the form $(p + xg, q + yg)$, where P, Q and g are as in (2) and (3) above. Clearly V is a vector subspace of $B_{n+1} \times B_{n+1}$. Let $\mathbb{P}(V)$ be the projective space of lines through $0 \in V$. Since the differential equations $(P+xg)dy - (Q+yg)dx = 0$ and $\lambda(P+xg)dy - \lambda(Q+yg)dx = 0$

define the same foliation in $\mathbb{C}^2$, we can identify the set of all foliations of degree n in $\mathbf{CP}^2$ with a subset $\mathrm{Fol}(n) \subset \mathbb{P}(V)$. We consider $\mathrm{Fol}(n)$ with the topology induced by the topology of $\mathbb{P}(V)$. $\mathrm{Fol}(n)$ is called the *space of foliations of degree n in* $\mathbf{CP}^2$.

Given a foliation $\mathcal{F} \in \mathrm{Fol}(M)$ and an isolated singularity $p \in \mathrm{sing}(\mathcal{F})$ we say that p is *non-degenerate* if $\mathcal{F}$ is represented in local coordinates centered at p by a holomorphic vector field X such that $DX(p)$ is nonsingular, i.e., X has a simple zero at p. Let p be a non-degenerate singularity. Let λ_1, λ_2 denote the eigenvalues of $DX(p)$. The *characteristic numbers* of $\mathcal{F}$ at p are the quotients of $\lambda_1/\lambda_2, \lambda_2/\lambda_1$ of the eigenvalues of $DX(p)$. A non-degenerate singularity p is *irreducible* if $\lambda_1/\lambda_2 \notin \mathbb{Q}_+$. The singularity is *hyperbolic* if $\lambda_1/\lambda_2 \in \mathbb{C} \setminus \mathbb{R}$. Hyperbolic singularities are analytically linearizable.

Using the above terminology we introduce some distinguished subsets of $\mathrm{Fol}(n) \subset \mathrm{Fol}(\mathbf{CP}^2)$. Denote by $\mathcal{S}(n)$ the set of foliations in $\mathrm{Fol}(n)$ with non-degenerate singularities and by $\mathcal{T}(n)$ the set of degree-n foliations with irreducible singularities. Denote by $\mathcal{A}(n)$ the set of foliations in $\mathcal{T}(n)$ that leave invariant the infinite line and by $\mathcal{H}(n)$ the set of foliations in $\mathcal{A}(n)$ that have hyperbolic singularities in L_∞.

The structure of above sets is described below:

PROPOSITION (2.1) ([17],[18]). *$\mathfrak{X}(n)$ is an analytic subvariety of* $\mathrm{Fol}(n)$ *and also if $n \geq 2$ then:*

1. *$\mathcal{T}(n)$ contains an open dense subset of* $\mathrm{Fol}(n)$.
2. *$\mathcal{H}(n)$ contains an open dense subset $\mathcal{M}_1(n) \subset \mathcal{H}(n)$ such that if $\mathcal{F} \in \mathcal{M}_1(n)$, $n \geq 2$ then:*
   (a) *L_∞ is the only algebraic solution of $\mathcal{F}$.*
   (b) *The holonomy group of the leaf $L_\infty \setminus \mathrm{sing}(\mathcal{F})$ is non-solvable.*
3. *$\mathcal{M}_1(n) \subset \mathcal{H}(n) \subset \mathcal{A}(n)$ are open dense subsets of $\mathfrak{X}(n)$.*

The next lemma says, roughly speaking, that both the singularities and the separatrices of a foliation with non-degenerate singularities, move analytically under analytic deformations of the foliation.

LEMMA (2.2) ([18], Proposition 1). *Let $\mathcal{F}_o \in \mathcal{S}(n)$. Then $\#\mathrm{sing}(\mathcal{F}_o) = n^2+n+1 = N$. Moreover if $\mathrm{sing}(\mathcal{F}_o) = \{p_1^o, \dots, p_N^o\}$ where $p_i^o \neq p_j^o$ if $i \neq j$, then there are connected neighborhoods $U_j \ni p_j$, pairwise disjoint, and holomorphic maps $\varphi_j : \mathcal{U} \subset \mathcal{S}(n) \to U_j$, where $\mathcal{U} \ni \mathcal{F}_o$ is an open neighborhood in $\mathcal{F}(n)$, such that for $\mathcal{F} \in \mathcal{U}$, $\mathrm{sing}(\mathcal{F}) \cap U_j = \varphi_j(\mathcal{F})$ is a non-degenerate singularity. Moreover, if $\mathcal{F}_o \in T(n)$ then the two local separatrices as well as their associated eigenvalues depend analytically on $\mathcal{F}$. In particular the set $\mathcal{S}(n)$ is open in* $\mathrm{Fol}(n)$.

3. Integrable deformations and rigidity

In this section we state two key results for the proof of Theorem A. The main notion is the notion of integrable deformation that we introduce now. Let $\mathcal{F}$ be a holomorphic foliation with isolated singularities on a complex surface M. An *integrable deformation* of $\mathcal{F}$ is a holomorphic foliation $\tilde{\mathcal{F}}$ of codimension one on the total space $M \times \mathbb{D}$ such that:

1. The singular set of $\tilde{\mathcal{F}}$ has codimension greater than 1.
2. $\mathcal{F}$ is generically transverse to each slice $M \times \{t\}$ and induces by restriction a foliation $\mathcal{F}_t := \tilde{\mathcal{F}}|_{M \times \{t\}}$ with isolated singularities.

3. $\mathcal{F}_0 = \mathcal{F}$ in $M \times \{0\} \cong M$.

In few words, *an integrable deformation is a deformation* $\{\mathcal{F}_t\}_{t\in\mathbb{D}}$ *which embeds into an analytic foliation* $\tilde{\mathcal{F}}$.

REMARK (3.1). Integrable deformations are also often called *unfoldings*, which by their turn have a well-developed theory, mainly related to the study of singularities (called *local case*). We shall not use this terminology because our study does not go into this direction, indeed we focus on global properties of the deformations, while assuming the singularities to be already stable (hyperbolic).

The first key result below, derived from arguments in [13], assures analytical triviality for integrable deformations of foliations in $\mathcal{M}_1(n)$ provided the infinite line is invariant for every foliation on the deformation.

THEOREM (3.2). *Let* $\{\mathcal{F}_t\}_{t\in\mathbb{D}}$*, be an integrable deformation in the class* Fol(n) *of a foliation* $\mathcal{F}_0 = \mathcal{F} \in \mathcal{S}(n)$*. Then for* $|t| < \epsilon$ *small enough, the integrable deformation is analytically equivalent to the trivial integrable deformation of* $\mathcal{F}$.

Proof. Denote by $\tilde{\mathcal{F}}$ the foliation on $\mathbf{CP}^2 \times \mathbb{D}$ such that $\forall t \in \mathbb{D}$, $\tilde{\mathcal{F}}|_{\mathbf{CP}^2\times\{t\}} = \mathcal{F}_t$, by $\pi : \mathbb{C}^3\backslash\{0\} \to \mathbf{CP}^2$ the canonical projection and by $\Pi : (\mathbb{C}^3\backslash\{0\})\times\mathbb{D} \to \mathbf{CP}^2\times\mathbb{D}$ the map $\Pi(p,t) := (\pi(p), t)$. Let also $\mathcal{F}^* := \Pi^*(\tilde{\mathcal{F}})$, be the pull-back foliation on $(\mathbb{C}^3\backslash\{0\}) \times \mathbb{D}$. Then $\mathcal{F}^*$ extends to a foliation on $\mathbb{C}^3 \times \mathbb{D}$ by a Hartogs type argument.

CLAIM (3.3). *We can choose an integrable holomorphic 1-form* Ω *which defines* $\mathcal{F}^*$ *on* $\mathbb{C}^3 \times \mathbb{D}$ *such that*

$$\Omega = A(x,t)dt + \sum_{i=1}^{3} B_j(x,t)dx_j,$$

where B_j *is a homogeneous polynomial of degree* $n+1$ *in* x*,* A *is a homogeneous polynomial of degree* $n+2$ *in* x*,* $\sum_{i=1}^3 x_j B_j(x,t) \equiv 0$ *and* $\Omega_t := \sum_{i=1}^3 B_j(x,t)dx_j$ *defines* $\pi^*(\mathcal{F}_t)$ *on* $\mathbb{C}^3$.

proof of Claim (3.3). First we remark that by triviality of Dolbeault and Cech cohomology groups of $\mathbb{C}^3 \times \mathbb{D}$, $\mathcal{F}^*$ is given by an integrable holomorphic 1-form, say, ω in $\mathbb{C}^3 \times \mathbb{D}$. The restriction $\omega_t := \omega|_{\mathbb{C}^3\times\{t\}}$ defines $\mathcal{F}^*_t := \pi^*(\mathcal{F}_t)$ in $\mathbb{C}^3$. Thus we may write $\omega = \alpha(x,t)dt + \sum_{k=1}^3 \beta^k(x,t)dx_k = \alpha(x,t)dt + \omega_t(x)$. Since the radial vector field R is tangent to the leaves of $\mathcal{F}^*$ we have $\omega \circ R = 0$ so that $\omega_t \circ R = 0$, i.e. $\sum_{k=1}^3 x_k \beta^k(x,t) = 0$. Now we use the Taylor expansion in the variable $x = (x_1, x_2, x_3)$ of ω around a point $(0,t)$ so that $\omega = \sum_{j=\nu}^{+\infty} \omega_j$ where $\omega_j(x,t) := \alpha_j(x,t)dt + \sum_{k=1}^3 \beta_j^k(x,t)dx_k = \alpha_j(x,t)dt + \omega_j^t$ and α_j, β_j^k are holomorphic in (x,t), homogeneous polynomial of degree j in x, $\omega_\nu \not\equiv 0$.

Now the main argument in the proof of Claim (3.3) is the following:

LEMMA (3.4). $\Omega = \alpha_{\nu+1}dt + \omega_\nu^t$ *defines* $\mathcal{F}^*$ *in* $\mathbb{C}^3 \times \mathbb{D}$.

Proof of Lemma (3.4). Indeed, $\omega \wedge d\omega = 0 \Rightarrow i_R(\omega\wedge d\omega) = i_R(\omega).d\omega - \omega \wedge i_R(d\omega) = 0$. On the other hand, $\omega \wedge i_R(d\omega) = 0$ (since $i_R(\omega) = 0$) $\Rightarrow i_R(d\omega) = f\omega$ for some holomorphic function f (Division lemma of Saito [22]). Therefore the Lie derivative of ω with respect to R is

$$L_R(\omega) = i_R(d\omega) + d(i_R(\omega)) = f\omega. \tag{3.5}$$

On the other hand since $\omega = \sum_{j=\nu}^{+\infty} \omega_j = \sum_{j=\nu}^{+\infty} (\alpha_j(x,t)dt + \omega_j^t)$ we obtain

$$L_R(\omega) = \sum_{j=\nu}^{+\infty} L_R(\alpha_j(x,t)dt + \omega_j^t)$$

$$= \sum_{j=\nu}^{+\infty} \frac{d}{dz}[\alpha_j(e^z x,t)dt + \sum_{k=1}^{3} \beta_j^k(e^z x,t)e^z dx_k]|_{z=0}$$

(The flow of R is $R_z(x,t) = (e^z x, t)$)

$$= \sum_{j=\nu}^{+\infty} [j\alpha_j(x,t)dt + (j+1)\omega_j^t]. \tag{3.6}$$

Now we write the Taylor expansion also for f in the variable x. $f(x,t) = \sum_{j=0}^{+\infty} f_j(x,t)$ where $f_j(x,t)$ is holomorphic in (x,t) homogeneous polynomial of degree j in x. We obtain from (1) and (2)

$$\sum_{j=\nu}^{+\infty} j\alpha_j dt + (j+1)\omega_j^t = (\sum_{k=0}^{+\infty} f_k)(\sum_{l=\nu}^{+\infty} \omega_l)$$

$$= \sum_{j\geq\nu} \sum_{l+k=j, l\geq\nu} f_k\omega_l$$

$$= \sum_{j\geq\nu} \sum_{l+k=j, l\geq\nu} (f_k\alpha_l dt + f_k\omega_l^t) \qquad \text{and}$$

Then

$$j\alpha_j = \sum_{l+k=j, l\geq\nu} (f_k\alpha_l) \tag{3.7}$$

$$(j+1)\omega_j^t = \sum_{l+k=j, \geq\nu} (f_k\omega_l^t) \quad \forall j \geq \nu \quad \text{and} \tag{3.8}$$

Notice that $\omega_\nu \not\equiv 0 \implies \omega_\nu^t \not\equiv 0, \forall t$ with $|t|$ small enough. In particular (3) and (4) imply $f_0\alpha_\nu = \nu\alpha_\nu$ and $f_0\omega_\nu^t = (\nu+1)\omega_\nu^t$ then $f_0 = \nu+1, \alpha_\nu = 0$.

An induction argument shows that:

$j \geq \nu \Rightarrow (\alpha_{j+1}dt + \omega_j^t) \wedge \Omega = 0, (\Omega := \alpha_{\nu+1}dt + \omega_\nu^t)$

Finally since the degree of the foliation $\mathcal{F} = \mathcal{F}_0$ is n we have $\nu = n+1$. This proves Lemma (3.4). □

LEMMA (3.9). *There exists a complete holomorphic vector field X on $\mathbb{C}^3 \times \mathbb{D}_\epsilon$, $\mathbb{D}_\epsilon \subset \mathbb{D}$ small subdisc of radius $\epsilon > 0$, such that $X(x,t) = \frac{\partial}{\partial t} + \sum_{j=1}^{3} F_j(x,t)\frac{\partial}{\partial x_j}$, $\Omega \circ X = 0$ and $F_j(x,t)$ is linear on x.*

Proof of Lemma (3.9). We may present $\Omega = A(x,t)dt + \sum_{j=1}^{3} B_j(x,t)dx_j = A(x,t)dt + \omega_t$ where $i_R(\omega_t) = 0$, B_j is a homogeneous polynomial of degree $n+1$ in x, A is a homogeneous polynomial of degree $n+2$ in x.

At this point we need the following claim:

CLAIM (3.10). $\forall t \in \mathbb{D}_\epsilon$ *($\epsilon \geq 0$ small enough) we have* $\operatorname{sing}(\mathcal{F}_t) \subset \{A(.,t) = 0\}$.

Proof of Claim (3.10). Since $\Omega \wedge d\Omega = 0$ we have the coefficients of $dt \wedge dx_i \wedge dx_j$ equal to zero, that is:

$$-A(\frac{\partial B_j}{\partial x_i} - \frac{\partial B_i}{\partial x_j}) + B_j\frac{\partial B_j}{\partial t} - B_i\frac{\partial B_j}{\partial t} + B_i\frac{\partial A}{\partial x_j} - B_j\frac{\partial A}{\partial x_i} = 0 \tag{3.11}$$

Now given $p_0 \in \text{sing}(\mathcal{F}_{t_0})$, ($t_0 \approx 0$, so that $\mathcal{F}_{t_0} \in \mathcal{S}(n)$) we have from (5) that $(B_i(p_0,t_0) = B_j(p_0,t_0) = 0)$ and $A(p_0,t_0)(\frac{\partial B_j}{\partial x_i}(p_0,t_0) - \frac{\partial B_i}{\partial x_j}(p_0,t_0))$. Since $\mathcal{F}_{t_0} \in T(n)$ we have $\frac{\partial B_j}{\partial x_i}(p_0,t_0) \neq \frac{\partial B_i}{\partial x_j}(p_0,t_0)(i \neq j)$ and $A(p_0,t_0) = 0$. □

Using now a natural parametric version of Noether's lemma we conclude that there exist $F_j(x,t)$ holomorphic in (x,t), homogeneous polynomial of degree $1 = (n+2)-(n+1)$ in x, such that $A(x,t) = \sum_{j=1}^{3} F_j(x,t)B_j(x,t)$. Now we define $X(x,t) := 1\frac{\partial}{\partial t} - \sum_{j=1}^{3} F_j(x,t)\frac{\partial}{\partial x_j}$ so that $\Omega \circ X = A - \sum_{j=1}^{3} F_j B_j = 0$.

In addition X is complete because each F_j is of degree one in x. The flow of X writes $X_z(x,t) = (\Psi_z(x,t), t+z)$. Clearly $\Psi_z : \mathbb{C}^3\backslash\{0\} \longrightarrow \mathbb{C}^3\backslash\{0\}$ defines an analytic equivalence between $\mathcal{F}$ and $\mathcal{F}_z$. □

Theorem (3.2) is now proved. □

Another important remark is the following:

LEMMA (3.12). *Let $\mathcal{F}$, $\mathcal{G}$ be foliations with non-degenerate singularities on* $\mathbf{CP}^2$. *Assume that $\mathcal{F}$ and $\mathcal{G}$ have same degree $n \geq 2$ and are topologically conjugate in $\mathbb{C}^2$. Then L_∞ is also $\mathcal{G}$-invariant.*

Proof. Let us first recall that for a foliation $\mathcal{H} \in \text{Fol}(n)$ with non-degenerate singularities in $\mathbf{CP}^2$ we have $\sharp(\text{sing}(\mathcal{H})) = n^2 + n + 1$. Moreover, we have $\mathcal{H} \in \mathfrak{X}(n)$ if and only if $\sharp(\text{sing}(\mathcal{H})) \cap L_\infty = n+1$ ([17]). Thus, $\sharp(\text{sing}(\mathcal{F}) \cap \mathbb{C}^2 = n^2$ and $\sharp\,\text{sing}(\mathcal{G}) = \sharp\,\text{sing}(\mathcal{F}) = n^2+n+1$. Also we have $\sharp\,\text{sing}(\mathcal{G}) \cap \mathbb{C}^2 = \sharp\,\text{sing}(\mathcal{F}) \cap \mathbb{C}^2 = n^2$, because of the topological equivalence. Therefore we also have $\sharp\,\text{sing}(\mathcal{G}) \cap L_\infty = n+1$. Since $\mathcal{G}$ has degree n this implies that the line L_∞ is also $\mathcal{G}$-invariant. □

Denote by $\text{Sep}(\mathcal{F})$ and $\text{Sep}(\mathcal{G})$ respectively the set of separatrices of $\mathcal{F}$ and $\mathcal{G}$ in $\mathbb{C}^2$ that are transverse to L_∞ at some singular point $p \in \text{sing}(\mathcal{F})$. The second key result is the following:

PROPOSITION (3.13). *Let $\mathcal{F}$, $\mathcal{G}$ be foliations on* $\mathbf{CP}^2$ *both leaving invariant the line L_∞ and having only hyperbolic singularities on this line. Let $\phi\colon \mathbb{C}^2 \to \mathbb{C}^2$ be a topological equivalence for $\mathcal{F}|_{\mathbb{C}^2}$ and $\mathcal{G}|_{\mathbb{C}^2}$. Then ϕ takes the separatrix set* $\text{Sep}(\mathcal{F})$ *onto the separatrix set* $\text{Sep}(\mathcal{G})$*.*

Proof. Fix a neighborhood W of L_∞ in $\mathbf{CP}^2$. Put $W^* = W \setminus L_\infty$. Because ϕ is proper, the image $\phi(W^*)$ is of the form $V^* = V \setminus L_\infty$ for some neighborhood V of L_∞ in $\mathbf{CP}^2$. Take a separatrix $\Gamma \in \text{Sep}(\mathcal{F})$ and set $\Gamma^* = \Gamma \setminus (\Gamma \cap L_\infty)$. Then, for W sufficiently small, $W \cap \Gamma^*$ is a connected closed analytic subset of a neighborhood W^*, which is invariant by the foliation $\mathcal{F}^* := \mathcal{F}\big|_{W^*}$.

Because $\phi\big|_{W^*} : W^* \to V^*$ is proper, the image $\phi(W \cap \Gamma^*) \subset V^*$ is a connected closed subset of V^*. Since $\phi\big|_{W^*} : W^* \to V^*$ conjugates $\mathcal{F}^*$ to $\mathcal{G}^* = \mathcal{G}\big|_{V^*}$, the image $\phi(W \cap \Gamma^*)$ is also invariant by $\mathcal{G}^*$. If $\phi(W \cap \Gamma^*)$ accumulates on some regular point $p \in L_\infty \backslash \text{sing}(\mathcal{G})$ then, because L_∞ is also $\mathcal{G}^*$-invariant, $\phi(W \cap \Gamma^*)$ also accumulates on all points in L_∞. In particular, $\phi(W \cap \Gamma^*)$ accumulates on some singularity $q \in \text{sing}(\mathcal{G})$ which is hyperbolic by hypothesis. Since $\phi(W \cap \Gamma^*)$ accumulates on L_∞ it is not contained in a separatrix of $\mathcal{G}$ through q and therefore it accumulates on both separatrices. This contradicts the fact that $\phi(W \cap \Gamma^*)$ is closed in V^*. Thus we conclude that the only accumulation points of $\phi(W \cap \Gamma^*)$ are singularities of $\mathcal{G}$

in L_∞. This implies that $\phi(W \cap \Gamma^*)$ is contained in Sep($\mathcal{G}$). Because it is connected, it must correspond to a single separatrix of $\mathcal{G}$. □

COROLLARY (3.14). *Let $\mathcal{F}_0 \in \mathcal{H}(n)$, $n \geq 2$. Then any $\mathbb{C}^2$-topologically trivial deformation $\{\mathcal{F}_t\}_{t\in\mathbb{D}}$ of $\mathcal{F}$ in the class $\mathcal{S}(n)$, is a deformation in the class $\mathfrak{X}(n)$ and it is also s-trivial if we consider* t ≈ 0.

Proof. By Lemma (3.12) the deformation is in the class $\mathfrak{X}(n)$. By Proposition (3.13) the deformation is is s-trivial for $|t|$ small enough. □

REMARK (3.15). Since $\mathcal{H}(n)$ is open in Fol(n), any analytic deformation $\{\mathcal{F}_t\}_{t\in\mathbb{D}}$ of a foliation $\mathcal{F} \in \mathcal{S}(n)$ remains in $\mathcal{S}(n)$ for $|t|$ small enough. Thus the hypothesis of the above corollary is automatically satisfied for topologically trivial small deformations.

4. Fixed points and one-parameter pseudogroups

Denote by Diff($\mathbb{C}$, 0) the group of germs of complex diffeomorphisms fixing $0 \in \mathbb{C}$, $f(z) = \lambda z + \sum_{n\geq 2} a_n z^n$; $\lambda \neq 0$. Let $G \subset$ Diff($\mathbb{C}$, 0) be a finitely generated subgroup with a set of generators $g_1, \cdots, g_r \in G$ defined in a compact disk $\bar{D} = \bar{\mathbb{D}}_\epsilon$. We say that G has the "Dense Orbits Property" (D.O.P. for short) if there exists a neighborhood $0 \in V \subset \bar{D}$ where the pseudo-orbits of G are dense.

In the following we collect the main properties of non-solvable groups of germs.

THEOREM (4.1) ([1],[20], [24, 25]). *Suppose G is non-solvable. Then:*

1. *The basin of attraction of* (*the pseudo-orbits of*) *G is an open neighborhood B_G of the origin.*
2. *Either G has the D.O.P. or there exists an invariant germ of analytic curve* Γ (*equivalent to* $\operatorname{Im}(z^k) = 0$ *for some* $k \in \mathbb{N}$) *where G has dense pseudo-orbits and such that G has also dense pseudo-orbits in each component of the complement of* Γ.
3. *If G contains some $f \in G$ with $f'(0) = e^{2\pi i\lambda}$, $\lambda \in \mathbb{C}\setminus\mathbb{R}$ and some flat element* $h = z + az^{k+1} + \ldots \neq \mathrm{Id}$ *then G has the D.O.P.*
4. *G is topologically rigid: given another non-solvable subgroup $G' \subset$* Diff($\mathbb{C}$, 0) *and an orientation preserving topological conjugation $\phi : B_G \to B_{G'}$ between G and G', then ϕ is holomorphic in a neighborhood of* 0.
5. *There exists a neighborhood* $0 \in W \subset B_G$ *where G has a dense set of hyperbolic fixed points.*

Holomorphic deformations in Diff($\mathbb{C}$, 0). Let $g \in$ Diff($\mathbb{C}$, 0) be defined in some open neighborhood $0 \in V \subset \mathbb{C}$. A *holomorphic* (*one-parameter*) *deformation* of g is a map $G : \mathbb{D}_\epsilon \to$ Diff($\mathbb{C}$, 0), ($\epsilon > 0$) which verifies the four properties:

1. $G(0) = g$ as germs.
2. The Taylor expansion coefficients of $G(t)$ depend holomorphically on t.
3. There is a uniform lower bound $R > 0$, independent of $t \in \mathbb{D}_\epsilon$, for the radii of convergence of $G(t)$ and $G(t)^{-1}$.
4. There is a uniform lower bound $C > 0$, independent of $t \in \mathbb{D}_\epsilon$, for the module of the linear coefficient of $G(t)$. In particular, there is a uniform upper bound for $|(G(t)^{-1})'(0)|$, independent of $t \in \mathbb{D}_\epsilon$.

Given a finitely generated pseudo-group $G \subset \mathrm{Diff}(\mathbb{C},0)$ with a set of generators $g_1, \cdots g_r \in G$; a holomorphic (one parameter) deformation of G is given by holomorphic deformation of g_j, $j = 1, \cdots, r$. We may restrict ourselves to the following situation:

G_t is a one-parameter analytic deformation of G with $t \in \mathbb{D}$, $G_0 = G$. We have $g_{1,t} \cdots g_{r,t}$ as a set of generators for G_t, all of them defined in a disk $\bar{\mathbb{D}}_\delta$ (uniformly on t). We will consider dynamical and analytical properties of such deformations. The results we state below have their proofs reduced to the following case which is studied in [27].

$g_{1,t}(z) = g_1(z) + tz^{D+1}$ where $D \in \mathbb{N}$ is fixed,
$g_{2,t}(z) = g_2(z), \cdots, g_{r,t}(z) = g_r(z)$.

For such deformations we have:

THEOREM (4.2) ([24, 25] and [27]). *Given a hyperbolic fixed point $p \approx 0$ for a word $f = f_n \circ f_{n-1} \circ \cdots \circ f_1$ in G, we consider the corresponding word $f_t = f = f_{n,t} \circ \cdots \circ f_{1,t}$ in G_t. Then f_t has a hyperbolic fixed point $p(t)$ given by the implicit differential equation with initial conditions:*

$$\frac{dp(t)}{p(t)^{D+1}dt} = \frac{f_t'(p(t))}{f_t'(p(t)) - 1} f_{1,t}'(p(t)), \quad p(0) = p.$$

In particular $p(t)$ depends analytically in t as well as its multiplier ${f_t}'(p(t))$. This holds for $|t| < \epsilon$ if $\epsilon > 0$ is small enough.

Let $\mathcal{M}_2(n) \subset \mathcal{M}_1(n)$ be the set of foliations such that the holonomy group at L_∞ has the D.O.P.

LEMMA (4.3) (Lemma 3, [18]). *For all $n \geq 2$, $\mathcal{M}_2(n)$ contains an open and dense subset* $\mathrm{Rig}(n)$ *of $\mathcal{M}_1(n)$.*

The proof of the existence of the set $\mathrm{Rig}(n)$ is essentially as follows: Given a foliation $\mathcal{F} \in \mathcal{M}_1(n)$ we define $\mathrm{ord}(\mathcal{F})$ as the minimum order of tangency with the identity of the flat elements in the holonomy group $\mathrm{Hol}(\mathcal{F}, L_\infty)$. Take $\mathcal{M}_1'(n) = \{\mathcal{F}_o \in \mathcal{M}_1(n); \mathrm{ord}(\mathcal{F}) = \mathrm{ord}(\mathcal{F}_o)$ for $\mathcal{F}$ in a neighborhood of $\mathcal{F}_o$ in $\mathcal{M}_1(n)\}$. Then it is proved in [18] that $\mathcal{M}_1'(n)$ is dense and open in $\mathcal{M}_1(n)$. Moreover, fixed an open connected subset $V \subset \mathcal{M}_1'(n)$ and $\mathcal{F}_o \in V \setminus \mathcal{M}_2(n)$ (if it exists), then for a singularity of $\mathcal{F}_o$ the corresponding quotient of eigenvalues λ satisfies $\lambda^{\ell-1} \in \mathbb{R}$ where $\ell = \mathrm{ord}\, \mathcal{F}_0$.

LEMMA (4.4). *Let $\mathcal{F} \in \mathrm{Rig}(n)$, $n \geq 2$; then each leaf $F \not\subset L_\infty$ is dense in $\mathbf{CP}^2$.*

Proof. First we notice that F must accumulate to L_∞. Since F is a non-algebraic leaf it must accumulate to some regular point. Because L_∞ is $\mathcal{F}$-invariant the leaf L must accumulate to every point in L_∞. Fix a regular point $p \in L_\infty \setminus \mathrm{sing}(\mathcal{F})$. Choose a small transverse disk $\Sigma \pitchfork L_\infty$ with $\Sigma \subset V$, V is a flow-box neighborhood of p. We consider the holonomy group $G = \mathrm{Hol}(\mathcal{F}, L_\infty, \Sigma)$. Then F accumulates to the origin $p \in \Sigma$ and since G has the D.O.P. it follows that F is dense in a neighborhood of p in Σ. Any other leaf L' of $\mathcal{F}$, $L' \neq L_\infty$ must have the same property. Using the continuous dependence of the solutions with respect to the initial conditions we may conclude that F accumulates any point $q \in F'$, $\forall F' \neq L_\infty$. Thus F is dense in $\mathbb{C}^2$ and since L_∞ is $\mathcal{F}$-invariant, F is dense in $\mathbf{CP}^2$. $\square$

PROPOSITION (4.5). *Given $\mathcal{F} \in \mathrm{Rig}(n)$, $n \geq 2$, each leaf containing a separatrix $\Gamma \subset \mathrm{Sep}(\mathcal{F})$ of $\mathcal{F}$ is dense in $\mathbf{CP}^2$ and it accumulates densely on a neighborhood of the origin for any transverse disk $\Sigma \pitchfork L_\infty$, $q \notin \mathrm{sing}\,\mathcal{F}$.*

Proof. Indeed, given a separatrix $\Gamma \subset \mathrm{Sep}(\mathcal{F})$ the leaf $L \supset \Gamma$ is nonalgebraic for $\mathcal{F} \in \mathcal{M}_1(n)$. This implies that $L \setminus \Gamma$ accumulates L_∞ and therefore any transverse disk Σ as above is cut by L. Now it remains to use the density of the pseudo-orbits of the holonomy group $\mathrm{Hol}(\mathcal{F}, L_\infty)$. □

Since the class $\mathrm{Rig}(n)$ is open in $\mathfrak{X}(n)$ we obtain:

LEMMA (4.6) ([24, 25] and [27]). *Let $\mathcal{F} \in \mathrm{Rig}(n)$ be given with $n \geq 2$ and $\{\mathcal{F}_t\}_{t \in \mathbb{D}}$ an analytic deformation in the class $\mathfrak{X}(n)$ of $\mathcal{F} = \mathcal{F}_0$. Let $p_1, \cdots, p_{n+1} \in L_\infty$ be the singularities of $\mathcal{F}_0$ in L_∞.*

1. *There exist analytic functions $p_j(t)$, $t \in \mathbb{D}_\epsilon$ such that* $\{p_1(t), \cdots, p_{n+1}(t)\} = \mathrm{sing}(\mathcal{F}_t) \cap L_\infty$, $p_j(0) = p_j$, $j = 1, \cdots, n+1$.
   Fix $q \in L_\infty \setminus \mathrm{sing}(\mathcal{F}_0)$ and take small simple loops $\alpha_j \in \pi_1(L_\infty \setminus \mathrm{sing}(\mathcal{F}_0), q)$ and a small transverse disk $\Sigma \pitchfork L_\infty$. Then for $\epsilon > 0$ small we have:
2. *The holonomy group $G_t := \mathrm{Hol}(\mathcal{F}_t, L_\infty, \Sigma) \subset \mathrm{Diff}(\Sigma, q)$ is generated by the holonomy maps $f_{j,t}$ associated to the loops α_j (α_j is also a simple loop around $p_j(t)$).*
   In particular we obtain
3. *$\{G_t\}_{t \in \mathbb{D}_\epsilon}$ is a one-parameter holomorphic deformation of* $G_0 = \mathrm{Hol}(\mathcal{F}_0, L_\infty, \Sigma)$.
4. *The group G_t is non-solvable with the density orbits property, and has a dense set $\eta_t \subset \Sigma \times \{t\}$ of hyperbolic fixed points around the origin (q, t). Moreover, given any $p(0) \in \eta_0$, $p(0) = f_0(p(0))$, there exists an analytic curve $p(t) \in \eta_t$ $f_t(p(t)) = p(t)$ where $f_t \in G_t$ is the corresponding deformation of f_0.*

5. Proof of Theorem A

We use the terminology introduced in the previous sections (compatible with the one in [18]) and original ideas of [13]. Let therefore $\{\mathcal{F}_t\}_{t \in \mathbb{D}}$ be a $\mathbb{C}^2$-topologically trivial analytic deformation in the class $\mathrm{Fol}(n)$ of $\mathcal{F} \in \mathrm{Rig}(n)$, $n \geq 2$. By Corollary (3.14) there exists $\epsilon > 0$ such that $\{\mathcal{F}_t\}_{t \in \mathbb{D}_\epsilon}$ is a a s-trivial deformation of $\mathcal{F}$ in the class $\mathfrak{X}(n)$. Theorem A is then a consequence of the main steps in the proof of Theorem (1.6) in [18] and of Theorem (3.2) and Corollary (3.14). Indeed, as a main step in [18] it is proved that *an s-trivial deformation of a foliation $\mathcal{F} \in \mathrm{Rig}(n)$ in the class $\mathfrak{X}(n)$ is analytically trivial for $|t|$ small enough.* This, together with Corollary (3.14), then implies our Theorem A, without the need to use the results in Section 4. Nevertheless, since the above mentioned fact is not clearly stated in the mentioned reference, we shall state it here. More precisely, we shall prove the following proposition:

PROPOSITION (5.1). *Let $\{\mathcal{F}_t\}_{t \in \mathbb{D}}$ be an analytic deformation in $\mathrm{Fol}(n)$ of a foliation $\mathcal{F}_0 \in \mathrm{Rig}(n)$, $n \geq 2$. If the deformation is topologically trivial in $\mathbb{C}^2$, then there exists $\epsilon > 0$ such that $\{\mathcal{F}_t\}_{t \in \mathbb{D}_\epsilon}$ is an integrable deformation of $\mathcal{F}$ in $\mathbf{CP}^2$.*

Let us prove this proposition. Consider the continuous codimension one complex distribution (with singularities) $\tilde{\mathcal{F}}$ on $\mathbf{CP}^2 \times \mathbb{D}_\epsilon$ defined as follows:

(i) $\operatorname{sing}(\tilde{\mathcal{F}}) = \bigcup_{|t|<\epsilon} \operatorname{sing}(\mathcal{F}_t) \times \{t\}$

(ii) The leaves of $\mathcal{F}_t$ are the intersections of the leaves of $\tilde{\mathcal{F}}$ with $\mathbf{CP}^2 \times \{t\}$, $\forall |t| < \epsilon$.

LEMMA (5.2). *$\tilde{\mathcal{F}}$ defines a codimension one continuous foliation on $\mathbf{CP^2} \times \mathbb{D}_\epsilon$, having singular set* $\operatorname{sing}(\tilde{\mathcal{F}})$.

Proof. Thanks to the topological triviality in $\mathbb{C}^2$, $\tilde{\mathcal{F}}$ is a continuous foliation on $\mathbb{C}^2 \times \mathbb{D}_\epsilon$. This foliation extends to a continuous foliation on $\mathbf{CP}^2 \times \mathbb{D}_\epsilon$ by adding the leaf with singularities $L_\infty \times \mathbb{D}_\epsilon$ (recall that L_∞ is invariant for every foliation $\mathcal{F}_t$, cf. Lemma (3.12)). □

REMARK (5.3). We proceed to prove that $\tilde{\mathcal{F}}$ is transversely holomorphic. This together with the fact that $\tilde{\mathcal{F}}$ has holomorphic leaves implies that $\tilde{\mathcal{F}}$ is a holomorphic foliation. Indeed, by an adaptation of the classical Theorem of Osgood ([11]) a continuous foliation which is holomorphic in both directions (tangent and transverse) is holomorphic.

LEMMA (5.4). *$\tilde{\mathcal{F}}$ has holomorphic leaves.*

Proof. It is enough to prove that the leaves of $\tilde{\mathcal{F}}$ are holomorphic close to $L_\infty \times \mathbb{D}_\epsilon$. Given a point $p(0) \in \eta_0$ and $f_0 \in G_0$ as in Lemma (4.6) above, the curve $p(t)$ and $f_t \in G_t$ given by (4) of this lemma, we have $\{p(t), |t| < \epsilon\} \subset \tilde{L}_{p(0)} \cap (\Sigma \times \mathbb{D}_\epsilon)$ where $\tilde{L}_{p(0)}$ is the $\tilde{\mathcal{F}}$-leaf through $p(0)$. On the other hand $\tilde{L}_{p(0)}$ is already holomorphic along the cuts $\tilde{L}_{p(0)} \cap (\mathbf{CP}^2 \times \{t\})$ for $L_{p'(0),t}$ for $p(0) = (p'(0), 0)$. This implies that $\tilde{L}_{p(0)}$ is analytic. Since the curves $\{p(t), |t| < \epsilon\}$ with $p(0) \in \eta_0$ are analytic and locally dense around $\{q\} \times \mathbb{D}_\epsilon \subset \Sigma \times \mathbb{D}_\epsilon$ (Lemma (4.6)) it follows that any leaf $\tilde{L}$ of $\tilde{\mathcal{F}}$ is a uniform limit of holomorphic leaves $\tilde{L}_{p(0)}$ and it is therefore holomorphic. Thus $\tilde{\mathcal{F}}$ has holomorphic leaves. □

Now we study the transverse behavior of $\tilde{\mathcal{F}}$. Fix a point $q_\infty \in L_\infty$ which is not a singular point for $\mathcal{F}_0$. Choose a transverse disk Σ_{q_∞} to L_∞ with $\Sigma_{q_\infty} \cap L_\infty = \{q_\infty\}$. For $|t|$ small enough we have $q_\infty \notin \operatorname{sing}(\mathcal{F}_t)$ and Σ_{q_∞} is transverse to $\mathcal{F}_t$. Denote by G_t the holonomy group $\operatorname{Hol}(\mathcal{F}_t, L_\infty, \Sigma_{q_\infty})$ of the leaf $L_\infty \setminus (L_\infty \cap \operatorname{sing}(\mathcal{F}_t))$ calculated at the section Σ_{q_∞}. Then, from Lemma (4.6) we promptly obtain:

LEMMA (5.5). *The holonomy group G_t is an analytic deformation of the holonomy group G_0.*

Using this we can prove:

LEMMA (5.6). *$\tilde{\mathcal{F}}$ is transversely holomorphic close to* $\mathbf{CP}^2 \times \{0\}$.

Proof. This is in fact a consequence of the topological rigidity in [20] for non-solvable groups of $\operatorname{Diff}(\mathbb{C}, 0)$ (Theorem (4.1)). Fix transverse section $\Sigma = \Sigma_{q_\infty}$ transverse to $\mathcal{F}_0$ at $q = q_\infty \in L_\infty$ as above. We may assume that $\Sigma \subset V$ where V is a flow-box neighborhood for $\mathcal{F}_0$ with $q \in V$. By Proposition (3.13) the homeomorphisms $\phi_t : \mathbb{C}^2 \to \mathbb{C}^2$ take the separatrices $\operatorname{Sep}(\mathcal{F}_0)$ of $\mathcal{F}_0$ onto the set of separatrices $\operatorname{Sep}(\mathcal{F}_t)$ of $\mathcal{F}_t$. By Proposition (4.5) the set of separatrices is dense in a neighborhood of the infinite line. Fix any $p \in \Sigma$ contained in a separatrix $p \ni \Gamma_0 \subset \mathrm{Sep}(\mathcal{F}_0)$ of $\mathcal{F}_0$ and denote by $P(\Gamma_0, p)$ the local plaque of $\mathcal{F}_0|_V$ that is contained in $\Gamma_0 \cap V$ and contains the fixed point p. Put $\Gamma_t = \phi_t(\Gamma_0)$ and consider the map $t \mapsto p(t) := P(\Gamma_t, \phi_t(p)) \cap (\Sigma \times \{t\})$. Clearly we may write $p(t) = \phi_t(P(\Gamma_0, p)) \cap (\Sigma \times \{t\})$ by choosing Σ and $|t|$ small enough. This map $t \mapsto p(t)$ is holomorphic as a consequence of Lemma (2.2).

Finally we define $h_t(p) := p(t)$ obtaining this way an injective map defined in a dense subset of Σ ($\mathcal{F}_0$ has dense separatrices in (Σ, q)), so that by the λ-lemma for complex mappings, (see [19]), we may extend h_t to a map that $h_t : \Sigma \to \Sigma$. Moreover, it is clear that if $f_{j,t}$ is a holonomy map as above then we have

$$h_t(f_{j,0}(p)) \quad = \quad f_{j,t}(h_t(p)),$$

because f_0 and f_t fix the separatrices. Therefore, by density we have $h_t \circ f_{j,0} = f_{j,t} \circ h_t$, $\forall j \in \{1, \cdots, n+1\}$ and the mapping h_t conjugates the holonomy groups $G_t = \mathrm{Hol}(\mathcal{F}_t, L_\infty, \Sigma)$ and G_0. By the topological rigidity theorem h_t is holomorphic which implies that $\tilde{\mathcal{F}}$ is transversely holomorphic close to $L_\infty \times \mathbb{D}_\epsilon$ ([20]) (Notice that, since h_t is close to the identity, it preserves the orientation so that we can exclude the anti-holomorphic case). □

Proof of Proposition (5.1). Lemma (5.6) and the density of $\mathrm{Sep}(\mathcal{F}_t)$ (Lemmas (4.4) and (4.6)), for $|t|$ small enough, assures that $\tilde{\mathcal{F}}$ is in fact transversely holomorphic in $\mathbf{CP}^2 \times \mathbb{D}_\epsilon$. Thus, the continuous foliation $\tilde{\mathcal{F}}$ is transversely holomorphic and has holomorphic leaves (Lemma 5.4) it is, by a Osgood-Hartogs' type argument (Remark 5.3), a holomorphic foliation in $\mathbf{CP}^2 \times \mathbb{D}_\epsilon$. □

End of the proof of Theorem A. Let $\{\mathcal{F}\}_{t \in \mathbb{D}}$ be a $\mathbb{C}^2$-topologically trivial deformation of a foliation $\mathcal{F} \in \mathrm{Rig}(n)$, $n \geq 2$, in the class $\mathrm{Fol}(n)$. By Proposition (5.1) we know that there is $\epsilon > 0$ such that $\{\mathcal{F}_t\}_{t \in \mathbb{D}_\epsilon}$ is an integrable deformation. Now, by Theorem (3.2) if ϵ is small enough then the deformation is analytically trivial. □

6. Proof of Theorem B

Now we are ready to prove Theorem B.

Proof of Theorem B. Suppose that $\mathcal{F} \in \mathfrak{X}(n)$, $n \geq 2$ is not topologically rigid in the class $\mathrm{Fol}(n)$. Let therefore $\{\mathcal{F}_t\}_{t \in \mathbb{D}}$ be a $\mathbb{C}^2$-topologically trivial analytic deformation of $\mathcal{F}$ in the class $\mathrm{Fol}(n)$, which is not analytically trivial for any $|t| < \epsilon$.

CLAIM (6.1). *The holonomy group* $\mathrm{Hol}(\mathcal{F}, L_\infty)$ *is solvable.*

Assume by contradiction that the holonomy $G = \mathrm{Hol}(\mathcal{F}, L_\infty)$ is a non-solvable group. We claim that this group has the D.O.P. Indeed, thanks to Theorem (4.1) (3) it is enough to observe that G contains some element f with linear part of the form $f^{(}0) = \exp(2\pi i \lambda)$ where $\lambda \in \mathbb{C} \backslash \mathbb{R}$ and some non-trivial flat element $g = z + az^{k+1} + \ldots$. The existence of such element f follows from the fact that $\mathrm{sing}(\mathcal{F}) \cap L_\infty$ contains (hyperbolic) singularities of the form $xdy - \lambda ydx + \ldots = 0$ where $\lambda \in \mathbb{C} \setminus \mathbb{R}$. The existence of such element g follows from the fact that G is not abelian, therefore some commutator in G has the desired form. This proves that G in fact has the D.O.P. Therefore, by the argumentation in Sections 3, 4 and mainly in Section 5 we conclude that the deformation $\mathcal{F}_t$ is s-trivial and integrable for $|t|$ small enough.

By Theorem (3.2) we conclude that the deformation is analytically trivial for $|t|$ small enough, a contradiction. This proves Claim (6.1).

Therefore the holonomy group $\mathrm{Hol}(\mathcal{F}, L_\infty)$ is necessarily solvable. This is the main point. From now on we proceed basically as follows: $\mathrm{Hol}(\mathcal{F}, L_\infty)$ solvable implies that it is linearizable and then there exists an integrating factor (cf. [3] or [6]). Indeed, since the singularities of $\mathcal{F}$ in the infinite line L_∞ are hyperbolic we conclude from [23] Proposition 5.1 page 185 that $\mathcal{F}$ is transversely affine in a neighborhood of L_∞ minus the set of local separatrices $\mathrm{Sep}(\mathcal{F})$. Indeed, given a polynomial differential equation $Pdy - Qdx = 0$ that defines $\mathcal{F}$ in the affine space $\mathbb{C}^2$ there is (cf. [23] Proposition 1.1 page 172) a meromorphic one-form η defined in a neighborhood of L_∞ minus $\mathrm{Sep}(\mathcal{F})$ such that $d\eta = 0$ and $d\Omega = \eta \wedge \Omega$ for $\Omega = Pdy - Qdx$ (in the terminology of [23] this one-form η is adapted to Ω along L_∞). Because the singularities in L_∞ are hyperbolic and thanks to the extension lemma ([23, Lemma 3.2, p. 178]) the one-form η admits a meromorphic extension to a neighborhood of L_∞ in $\mathbf{CP}^2$. Applying now Levi's extension theorem (see Remark 4.1 page 170 in [23]) we conclude that η extends to $\mathbf{CP}^2$ as a closed meromorphic (rational) one-form. This implies in particular that $\mathcal{F}$ admits an affine transverse structure in the complement of an invariant algebraic set (given by the invariant part of the polar set of η) in $\mathbf{CP}^2$. By Theorem 4.3 page 183 in [23] we conclude that the foliation $\mathcal{F}$ is a Darboux (logarithmic) foliation in the projective plane. Notice that at this point we need to use the hypothesis that the singular points in $\mathbb{C}^2$ are irreducible. □

7. A Noether's lemma for foliations

In this section we prove Theorem C from the Introduction, which reads as follows:

THEOREM (7.1) (Noether's lemma for foliations). *Let $\{\mathcal{F}_t\}_{t\in\mathbb{D}}$ be a holomorphic integrable deformation of a foliation $\mathcal{F}_0$ of degree n on $\mathbf{CP}^2$. Assume that for each singularity $p \in \mathrm{sing}(\mathcal{F}_0)$ the germ of integrable deformation at p is analytically trivial. Then there exists $\epsilon > 0$ such that $\{\mathcal{F}_t\}_{|t|<\epsilon}$ is analytically trivial.*

Proof. Denote by $\tilde{\mathcal{F}}$ the codimension one holomorphic foliation on $\mathbb{C} \times \mathbb{D}$ defined by $\mathcal{F}\big|_{\mathbb{C}\times\{t\}} = \mathcal{F}_t$, $\forall t \in \mathbb{D}$. Let also $\pi : \mathbb{C}^3 \backslash \{0\} \to \mathbf{CP}^2$ be the canonical projection and $\Pi : (\mathbb{C}^3 \backslash \{0\}) \times \mathbb{D} \to \mathbf{CP}^2 \times \mathbb{D}$ the map $\Pi(p,t) := (\pi(p), t)$. By Hartogs' extension theorem the foliation $\Pi^*(\tilde{\mathcal{F}})$ originally defined on $\mathbb{C}^3 \setminus \{0\} \times \mathbb{D}$ extends to a holomorphic foliation on $\mathbb{C}^3 \times \mathbb{D}$. Denote by $\tilde{\mathcal{F}}^*$ this extension. Since $H^1(\mathbb{C}^3 \times \mathbb{D}, \mathcal{O}^*) = 1$, the well-known solution of the Cousin problem [12] implies that there exists a holomorphic integrable 1-form Ω which defines $\tilde{\mathcal{F}}^*$.

$$\Omega = A(x,t)dt + \sum_{i=1}^{3} B_j(x,t)dx_j,$$

where A, B_j are holomorphic in $(x,t) \in \mathbb{C}^3 \times \mathbb{D}$, homogeneous polynomial in x of degree $n+2$, $n+1$; $\sum_{i=1}^{3} x_j B_j = 0$. The foliation $\pi^*(\mathcal{F}_t)$ extends to $\mathbb{C}^3$ and this extension $\mathcal{F}_t^*$ is given by $\Omega_t = 0$ for $\Omega_t := \sum_{i=1}^{3} B_j dx_j$.

CLAIM (7.2). *Given point $q \in \mathbb{C}^3 \times \mathbb{D}_\epsilon$, $q \notin \{0\} \times \mathbb{D}$, there exist a neighborhood $U(q)$ of q in $\mathbb{C}^3 \times \mathbb{D}_\epsilon$ and local holomorphic vector field $X_q \in \mathfrak{X}(U(q))$ such that $A = \Omega \circ X_q$ in $U(q)$, for ϵ small enough.*

Proof of Claim (7.2). If $q=(x_1,t_1)$ with $x_1\notin \mathrm{sing}(\mathcal{F}_0)$ then $x_1\notin\mathrm{sing}(\mathcal{F}_t)$ for $|t|$ small enough and in particular $x_1\notin\mathrm{sing}(\mathcal{F}_{t_1})$. Thus the existence of $X_q\in\mathfrak{X}(U(q))$ is assured in this case. On the other hand if $x_1\in\mathrm{sing}(\mathcal{F}_0)$ then we still have the existence of $X_q\in\mathfrak{X}(U(q))$ because of the local analytical triviality hypothesis for the integrable deformation at x_1. □

Using the claim we obtain an open cover $\{U_\alpha\}_{\alpha\in\mathbb{Q}}$ of $M:=\mathbb{C}^3\backslash\{0\}\times\mathbb{D}$ with U_α connected and $X_\alpha\in\mathfrak{X}(U_\alpha)$ such that $A=\Omega\circ X_\alpha$ in U_α, $\forall\alpha\in\mathbb{Q}$. Let $U_\alpha\cap U_\beta\neq\varnothing$ then we put $X_{\alpha\beta}:=(X_\alpha-X_\beta)|_{U_\alpha\cap U_\beta}$ to obtain $X_{\alpha\beta}\in\mathfrak{X}(U_\alpha\cap U_\beta)$ such that $\Omega\circ X_{\alpha\beta}=0$. Take now the rotational vector field

$$\begin{aligned} Y &= \mathrm{Rot}(B_1,B_2,B_3)\\ &= \Big(\frac{\partial B_3}{\partial x_2}-\frac{\partial B_2}{\partial x_3}\Big)\frac{\partial}{\partial x_1}+\Big(\frac{\partial B_1}{\partial x_3}-\frac{\partial B_3}{\partial x_1}\Big)\frac{\partial}{\partial x_2}+\Big(\frac{\partial B_2}{\partial x_1}-\frac{\partial B_1}{\partial x_2}\Big)\frac{\partial}{\partial x_3}. \end{aligned}$$

$Y\in\mathfrak{X}(\mathbb{C}^3\times\mathbb{D})$ and for each $t\in\mathbb{D}$ we have $i_Y(\mathrm{Vol})=d\Omega_t$ where $\mathrm{Vol}=dx_1\wedge dx_2\wedge dx_3$ is the volume element of $\mathbb{C}^3$ in the x-coordinates. Fixed now $q=(x_1,t_1)\notin\mathrm{sing}(\Omega_{t_1})$ then the leaf of $\mathcal{F}^*_{t_1}$ through q is spanned by $Y(q)$ and the radial vector field $R(q)$, as a consequence of the remark above: actually, we have $i_R i_Y(\mathrm{Vol})=i_R(d\Omega_t)=(n+1)\Omega_t$.

Given thus $U_{\alpha\beta}:=U_\alpha\cap U_\beta\neq\varnothing$, since $\Omega_t(X_{\alpha\beta})=0$ we have that $X_{\alpha\beta}$ is tangent to $\mathcal{F}^*_t$ outside the points $(x,t)\in\mathrm{sing}(\Omega_t)$ so that we can write $X_{\alpha\beta}=g_{\alpha\beta}R+h_{\alpha\beta}Y$ for some holomorphic functions $g_{\alpha\beta}$, $h_{\alpha\beta}\in\mathcal{O}(U_{\alpha\beta}\backslash\mathrm{sing}(\Omega_t))$. Since $\mathrm{sing}(\Omega_t)$ is an analytic set of codimension ≥ 2, Hartogs extension Theorem [17] implies that $g_{\alpha\beta}$, $h_{\alpha\beta}$ extend holomorphically to $U_{\alpha\beta}$. Now if $U_\alpha\cap U_\beta\cap U_\gamma\neq\varnothing$ then

$$0=X_{\alpha\beta}+X_{\beta\gamma}+X_{\gamma\alpha}=(g_{\alpha\beta}+g_{\beta\gamma}+g_{\gamma\alpha})R+(h_{\alpha\beta}+h_{\beta\gamma}+h_{\gamma\alpha})Y$$

and since R and Y are linearly independent outside $\mathrm{sing}(\Omega_t)$ we obtain: $g_{\alpha\beta}+g_{\beta\gamma}+g_{\gamma\alpha}=0$, $h_{\alpha\beta}+h_{\beta\gamma}+h_{\gamma\alpha}=0$.

Thus $(g_{\alpha\beta})$, $(h_{\alpha\beta})$ are additive cocycles in M and by Cartan's Theorem (for $\mathbb{C}^{n+1}\backslash\{0\}$, $n\geq 2$) these cocycles are trivial, that is, $\exists g_\alpha,h_\alpha\in\mathcal{O}(U_\alpha)$ such that if $U_\alpha\cap U_\beta\neq\phi$ then $g_{\alpha\beta}=g_\alpha-g_\beta$, $h_{\alpha\beta}=h_\alpha-h_\beta$ in $U_\alpha\cap U_\beta$. This gives $X_\alpha-X_\beta=X_{\alpha\beta}=g_{\alpha\beta}R+h_{\alpha\beta}Y=(g_\alpha R+h_\alpha Y)-(g_\beta R+h_\beta Y)$ in $U_\alpha\cap U_\beta\neq\phi$. Thus, in $U_\alpha\cap U_\beta\neq\phi$ we obtain $X_\alpha-g_\alpha R-h_\alpha Y=X_\beta-g_\beta R-h_\beta Y$ and this gives a global vector field $\tilde{X}\in\mathfrak{X}(M)$ such that $\tilde{X}|_{U_\alpha}:=X_\alpha-g_\alpha R-h_\alpha Y$. This vector field extends holomorphically to $\mathbb{C}^3\times\mathbb{D}$ and we have $(\Omega_t\circ\tilde{X})|_{U_\alpha}=\Omega_t\circ X_\alpha-g_\alpha\Omega_t\circ R-h_\alpha\Omega_t\circ Y=A$ so that $\Omega_t\circ\tilde{X}=A$.

It remains to prove that we may choose $\tilde{X}$ polynomial in the variable x. Indeed, we write $\tilde{X}=\sum_{k=0}^{\infty}\tilde{X}_k$ for the Taylor expansion of $\tilde{X}$ around the origin, in the variable x.

Then $\tilde{X}_k$ is holomorphic in (x,t) and homogeneous polynomial of degree k in the variable x. We have $A=\Omega_t\circ\tilde{X}=\sum_{k=0}^{+\infty}\Omega_t(\tilde{X}_k)$ and since it is polynomial homogeneous of degree $n+2$ in x it follows that $k\neq 1\Rightarrow\Omega_t(\tilde{X}_k)=0$ and $\Omega_t(\tilde{X}_1)=A$. Since $\tilde{X}_1$ is linear, the flow of $\tilde{X}_1$ gives an analytic trivialization for $\{\mathcal{F}_t\}_{t\in\mathbb{D}_\epsilon}$. □

REMARK (7.3). Theorem (7.1) can also be proved in the more analytical way as follows: Denote by $\tilde{\mathcal{F}}$ the product foliation on $\mathbf{CP}^2\times\mathbb{D}$, that is, $\tilde{\mathcal{F}}\big|_{\mathbf{CP}^2\times\{t\}}=\mathcal{F}_t,\forall t\in\mathbb{D}$. Let $\pi\colon\mathbf{CP}^2\times\mathbb{D}\to\mathbb{D}$ be the projection $\pi(p,t)=t$. By the local triviality hypothesis there exist a family of indexes J, holomorphic vector fields Y_j defined on open subsets $U_j\subset\mathbf{CP}^2\times\mathbb{D}$, such that $\{U_j\}_{j\in J}$ is an open cover of $\mathbf{CP}^2\times\mathbb{D}$, Y_j

is tangent to $\tilde{\mathcal{F}}$ on U_j and is projected by π to the field $\frac{d}{dt}$. In each nonempty intersection $U_i \cap U_j \neq \emptyset$ $Z_{ij} = Y_i - Y_j$ is a holomorphic vector field projected by π into the trivial field (Z_{ij} is *vertical*) and tangent to $\tilde{\mathcal{F}}$ in $U_i \cap U_j$. Therefore $\{Z_{ij}, U_i \cap U_j\}$ defines a cocycle of the considered cover with values in the sheaf $\mathfrak{X}^v_{\tilde{\mathcal{F}}}$ of vertical holomorphic vector fields tangent to $\tilde{\mathcal{F}}$. According to Grauert Theorem of Direct Images [9] the first cohomology group $H^1(\mathbf{CP}^2 \times \mathbb{D}, \mathfrak{X}^v_{\tilde{\mathcal{F}}})$ is a $\mathcal{O}_{\mathbb{D}}$-module of finite type and also $H^1(\mathbf{CP}^2 \times \mathbb{D}, \mathfrak{X}^v_{\tilde{\mathcal{F}}}) \otimes_{\mathcal{O}_{\mathbb{D}}} (\mathcal{O}_{\mathbb{D}}/(z)) = H^1(\mathbf{CP}^2, \mathfrak{X}_{\mathcal{F}_o})$ where $\mathfrak{X}_{\mathcal{F}_o}$ is the sheaf over $\mathbf{CP}^2$ of tangent holomorphic vector fields, tangent to $\mathcal{F}_o$. The sheaf $\mathfrak{X}_{\mathcal{F}_o}$ is isomorphic to the sheaf of sections of the line bundle associated with $\mathcal{F}_o$. It is well-known that $H^1(\mathbf{CP}^2, \mathfrak{X}_{\mathcal{F}_o}) = 0$ so that by Nakayama's Lemma we have $H^1(\mathbf{CP}^2 \times \mathbb{D}, \mathfrak{X}^v_{\tilde{\mathcal{F}}}) = 0$. It follows that there are vertical vector fields Z_j in the U_j tangent to $\tilde{\mathcal{F}}$ and such that in each $U_i \cap U_j \neq \emptyset$ we have $Z_{ij} = Z_i - Z_j$ and therefore $Y_j - Z_j = Y_i - Z_i$. This gives therefore a global holomorphic vector field Z over $\mathbf{CP}^2 \times \mathbb{D}$ which is tangent to $\tilde{\mathcal{F}}$ and projects by π on $\frac{d}{dz}$. Clearly the flow maps of Z define an analytical trivialization of the integrable deformation $\{\mathcal{F}_t\}_{t \in \mathbb{D}}$ (see [18], Lemma 6 and the Proof of Theorem 1.1 in page 400).

REMARK (7.4). It seems very reasonable to think that, in face of Theorem (7.1) and of the proof of Theorem B, one may obtain that: *a foliation on the complex projective plane, leaving invariant the infinite line and with singularities of first order type without too many resonances, is either* $\mathbb{C}^2$*-topologically rigid or it is a Darboux foliation.* Such a result would enlarge the list of topologically rigid foliations commenced by Ilyashenko.

Received October 3, 2012

Final version received February 1, 2013

M. TEYMURI GARAKANI
INSTITUTO DE MATEMATICA
UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
CP. 68530
RIO DE JANEIRO
RJ 21945-970 BRAZIL

A. MAFRA
INSTITUTO DE MATEMATICA
UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
CP. 68530
RIO DE JANEIRO
RJ 21945-970 BRAZIL

B. SCÁRDUA
INSTITUTO DE MATEMATICA
UNIVERSIDADE FEDERAL DO RIO DE JANEIRO
CP. 68530
RIO DE JANEIRO
RJ 21945-970 BRAZIL
bruno.scardua@gmail.com

REFERENCES

[1] M. BELLIART, I. LIOUSSE, AND F. LORAY, *Sur l'existence de points fixes attractifs pour les sous-groupes de* $Aut(\mathbb{C},0)$, C. R. Acad. Sci.,Paris, Ser. I, Math. **324**, (1997), 443–446.

[2] C. CAMACHO, A. LINS NETO, AND P. SAD, *Topological invariants and equidesingularization for holomorphic vector fields*, J. Differential Geom. **20** 1, (1984), 143–174.

[3] C. CAMACHO, A. LINS NETO, AND P. SAD, *Foliations with algebraic limit sets*, Ann. of Math. **136** (2), (1992), 429–446.

[4] C. CAMACHO AND P. SAD, *Topological classification and bifurcations of holomorphic flows with resonances in* $\mathbb{C}^2$, Inventiones Mathematicae **67** (3), October, 1982.

[5] C. CAMACHO AND B. SCÁRDUA, *Holomorphic foliations with Liouvillian first integrals*, Ergodic Theory Dynam. Systems **21** 3, (2001), 717–756.

[6] D. CERVEAU AND J.-F. MATTEI, *Formes intégrables holomorphes singulières*, With an English summary, Astérisque, Société Mathématique de France, Paris **97**, (1982) 193.

[7] D. CERVEAU AND P. SAD, *Problèmes de modules pour les formes différentielles singulières dans le plan complexe*, Comment. Math. Helv. **61** 2, (1986), 222–253.

[8] X. GÓMEZ-MONT, *Unfoldings of holomorphic foliations*, Publ. Mat. **33** 3, (1989), 501–515.

[9] H. GRAUERT AND R. REMMERT, Theory of Stein Spaces, Springer-Verlarg.

[10] I. GUCKENHEIMER, *Hartman's theorem for complex flows in the Poincaré domain*, Comp. Math. **24**, (1972), 75–82.

[11] R. C. GUNNING, Introduction to holomorphic functions of several complex variables, Vol. I, Wandsworth & Brooks/Cle, USA, 1990.

[12] R. C. GUNNING AND H. ROSSI, Analytic Functions of Several Complex Variables, Prentice-Hall Inc., Englewood CliQs, NJ, 1965.

[13] Y. ILYASHENKO, *Topology of phase portraits of analytic differential equations on a complex projective plane*, Trudy Sem. Petrovsk. **4**, (1978), 83–136.

[14] Y. ILYASHENKO, *Total Rigidity of Polynomial Foliations on the Complex Projective Plane* in Proceedings of the Steklov Institute of Mathematics **259**, (2007), 60–72.

[15] Y. ILYASHENKO, AND V. MOLDAVSKIS, *Total rigidity of generic quadratic vector fields*, Mosc. Math. J. **11** 3, (2011), 521–530, 630.

[16] Y. ILYASHENKO AND S. YAKOVENKO, Lectures on analytic differential equations. Graduate Studies in Mathematics, **86**. American Mathematical Society, Providence, RI, 2008. Reviewed

[17] A. LINS NETO, Algebraic solutions of polynomial differential equations and foliations in dimension two, Holomorphic dynamics (Mexico, 1986), 192–232, Lecture Notes in Math., 1345, Springer, Berlin, 1988.

[18] A. LINS NETO, P. SAD, AND B. SCÁRDUA, *On topological rigidity of projective foliations*, Bull. Soc. Math. France **126** 3, (1998), 381–406.

[19] R. MAÑÉ, R. SAD, AND D. SULLIVAN, *On the dynamics of rational maps*, Ann. Sci. Ec. Norm. Sup. 4 **16** (2), (1983), 193–217.

[20] I. NAKAI, *Separatrices for non-solvable dynamics on* $(\mathbb{C},0)$, Ann. Inst. Fourier (Grenoble) **44** 2, (1994), 569–599.

[21] I. PETROVSKI AND E. LANDIS, *On the number of limit cycles of the equation* $dy/dx = P(x,y)/Q(x,y)$ *where P and Q are polynomials of degree 2*, Amer. Math. Society Trans., Series 2, **10**, (1963), 125–176.

[22] K. SAITO, *On a generalization of de-Rham lemma*, Ann. Inst.Fourier (Grenoble) **26** 2 vii, (1976), 165–170.

[23] B. SCÁRDUA, *Transversely affine and transversely projective foliations*, Ann. Sci. Ec. Norm. Sup., 4^e série, t. **30**, (1997), 169–204

[24] A. A. SCHERBAKOV, *On the density of an orbit of a pseudogroup of conformal mappings and a generalization of the Hudai-Verenov theorem*, Vestnik Moskov. Univ. Ser. I Mat. Mekh. **31**, (1982), 10–15.

[25] A. A. SHCHERBAKOV, *Dynamics of Local Groups of Conformal Mappings and Generic Properties of Differential Equations on* $\mathbb{C}^2$, Tr. Mat. Inst. im. V.A. Steklova, Ross. Akad. Nauk **254**, (2006), 111–129 [Proc. Steklov Inst. Math. 254, 103–120 (2006)].

[26] A. SEIDENBERG, *Reduction of singularities of the differential equation* $Ady = Bdx.$, Amer. J. Math. **90**, (1968), 248–269.

[27] B. WIRTZ, *Entropy and fixed points for non-solvable dynamics of* $(\mathbb{C},0)$, Bol. Soc. Brasil. Mat. (N.S.) **29** 1, (1998), 53–74.

Bol. Soc. Mat. Mexicana (3) Vol. 19, 2013

A WEAK ORLICZ-PETTIS THEOREM

CHARLES SWARTZ

ABSTRACT. Kalton has shown that if E is a separable F space and M is a subspace of the dual of E which separates the points of E, then any series in E which is subseries convergent with respect to the weak topology, $\sigma(E,M)$, on E generated by M is subseries convergent in E. In this note we generalize Kalton's result to multiplier convergent series when the space of multipliers has the signed weak gliding hump property and the space E is a separable infra-Pták space. Applications to sequence and function spaces are given.

Kalton ([7] Corollary to Theorem 3) has shown that if E is a separable F space and M is a subspace of the dual of E which separates the points of E, then any series in E which is subseries convergent with respect to the weak topology, $\sigma(E,M)$, on E generated by M is subseries convergent in E. In this note we generalize Kalton's result to multiplier convergent series when the space of multipliers has the signed weak gliding hump property and the space E is a separable infra-Pták space. If (Z,τ) is a topological vector space and λ is a scalar sequence space, a series $\sum z_j$ in Z is λ multiplier convergent if the series $\sum_{j=1}^{\infty} t_j z_j$ is τ convergent for every $t = \{t_j\} \in \lambda$; for example, a series is m_0 multiplier convergent, where m_0 is the sequence space of sequences with finite dimensional range, iff the series is subseries convergent. The elements of λ are called multipliers. An Orlicz-Pettis Theorem is a result which asserts that a series which is λ multiplier convergent in some weak topology is λ multiplier convergent in a stronger topology. The original Orlicz-Pettis Theorem asserts that a series in a Banach space which is subseries convergent in the weak topology of the space is subseries convergent in the norm topology ([12, 13]; see [6, 4] for historic discussions of the theorem).

If E is a separable infra-Pták space and λ has the weak gliding hump property, we show that if $\sum_j x_j$ is a series in E which is λ multiplier convergent in the weak topology $\sigma(E,M)$ on E generated by M, then the series is λ multiplier convergent in E. The result contains the case where the series are subseries convergent as in the original Orlicz-Pettis Theorem. Applications to sequence and function spaces are given.

We begin by establishing the notation and assumptions necessary to describe and establish the result. Throughout E will denote a Hausdorff locally convex space with dual E'; if X,Y is a dual pair, the weak topology on X from Y will be denoted by $\sigma(X,Y)$. Also, λ will denote a scalar sequence space which contains c_{00}, the space of sequences which are eventually 0.

Our result requires a gliding hump assumption on the space of multipliers λ. If $s = \{s_j\}$ and $t = \{t_j\}$ are scalar sequences, the coordinate product of s and t is denoted by $st = \{s_j t_j\}$ and if $A \subset \mathbb{N}$, the characteristic function of A is denoted by χ_A. A sequence of signs is a sequence $s = \{s_j\}$, where $s_j = \pm 1$ for every j. An

2010 *Mathematics Subject Classification*: 46A99.

Keywords and phrases: Orlicz-Pettis, multiplier convergent series.

interval in $\mathbb{N}$ is a subset of the form $I = [m,n] = \{j : m \le j \le n\}$, $m,n \in \mathbb{N}, m \le n$; a sequence of intervals $\{I_j\}$ is increasing if $\max I_j < \min I_{j+1}$. The space λ has the signed weak gliding hump property (signed WGHP) if whenever $t \in \lambda$ and $\{I_j\}$ is an increasing sequence of intervals, there exist a sequence of signs $\{s_j\}$ and a subsequence $\{n_j\}$ such that the coordinate sum $\sum_{j=1}^{\infty} s_j \chi_{I_{n_j}} t \in \lambda$; if the signs can always be chosen equal to 1, then λ has the weak gliding hump property (WGHP). A space λ is monotone if $\chi_A t \in \lambda$ whenever $A \subset \mathbb{N}$ and $t \in \lambda$; for example, c_0, l^p $(0 < p \le \infty)$. A monotone space has WGHP; the space cs of convergent series is not monotone but has WGHP and bs, the space of bounded series, has signed WGHP but not WGHP (see [17] for further examples). The WGHP was introduced by Noll ([11]) and the signed WGHP by Stuart ([15]) to treat weak sequential completeness in β duals (see [17] for such results).

The locally convex spaces which we consider are the infra-Pták spaces. The locally convex space E is an infra-Pták space if a $\sigma(E',E)$ dense subspace $M \subset E'$ is $\sigma(E',E)$ closed whenever $M \cap U^0$ is $\sigma(E',E)$ closed for every neighborhood of 0, U, in E, where U^0 is the polar of U ([5, 10, 1]). For example, any complete metrizable locally convex space is an infra-Pták space (this is essentially a result of Krein-Smulian ([5, 3.10.2],[10, 34.3(5)],[16, 23.8]).

We also require a Hahn-Schur result in the proof of the main theorem. Suppose for every i the series $\sum_{j=1}^{\infty} x_{ij}$ is λ multiplier convergent in E and for every $t = \{t_j\} \in \lambda$ the limit, $\lim_i \sum_{j=1}^{\infty} t_j x_{ij}$, exists. If λ has signed WGHP and $x_j = \lim_i x_{ij}$, then the series $\sum_j x_j$ is λ multiplier convergent and $\sum_{j=1}^{\infty} t_j x_j = \lim_i \sum_{j=1}^{\infty} t_j x_{ij}$ for every $t \in \lambda$ (see [17, Theorem 7.6] for the result).

THEOREM (1). *Assume that E is a separable infra-Pták space with $M \subset E'$ a subspace which separates the points of E and that λ has signed WGHP. If the series $\sum_j x_j$ is $\sigma(E,M)$ λ multiplier convergent, then the series $\sum_j x_j$ is λ multiplier convergent in the topology of E.*

Proof. We claim that the series $\sum_j x_j$ is $\sigma(E,E')$ λ multiplier convergent and then the result will follow from the Orlicz-Pettis Theorem for locally convex spaces ([6],[17, 4.10]). In what follows if $t = \{t_j\} \in \lambda$,then $\sum_{j=1}^{\infty} t_j x_j$ will denote the $\sigma(E,M)$ sum of the series. Set

$$M_1 = \{x' \in X' : x'(\sum_{j=1}^{\infty} t_j x_j) = \sum_{j=1}^{\infty} t_j x'(x_j) \ for\ all\ t = \{t_j\} \in \lambda\}.$$

Now M is $\sigma(E',E)$ dense in E' since M separates points and $M \subset M_1$ so if M_1 is $\sigma(E',E)$ closed, we have $M_1 = E'$ and we are finished. By the infra-Pták assumption it suffices to show $M_1 \cap U^0$ is $\sigma(E',E)$ closed when U is a neighborhood of 0 in E. Since E is separable, $(U^0, \sigma(E',E))$ is metrizable ([9, 21.3(4)]) so it suffices to show $M_1 \cap U^0$ is sequentially $\sigma(E',E)$ closed. Suppose $\{x_i'\} \subset M_1$ and $\sigma(E',E) - \lim x_i' = x'$. For $t \in \lambda$,

$$x_i'(\sum_{j=1}^{\infty} t_j x_j) = \sum_{j=1}^{\infty} t_j x_i'(x_j) \to x'(\sum_{j=1}^{\infty} t_j x_j)$$

as $i \to \infty$. Now $\lim_i x_i'(x_j) = x'(x_j)$ for every j so by the Hahn-Schur result for λ discussed above we have that $\sum_{j=1}^{\infty} x'(x_j)$ is λ multiplier convergent and for every

$t \in \lambda$,

$$x'_i(\sum_{j=1}^{\infty} t_j x_j) = \sum_{j=1}^{\infty} t_j x'_i(x_j) \to x'(\sum_{j=1}^{\infty} t_j x_j) = \sum_{j=1}^{\infty} t_j x'(x_j)$$

as $i \to \infty$. Thus, $x' \in M_1$ as desired. □

Since the space m_0 is monotone and, therefore, has WGHP, the result above applies to subseries convergent series in the spirit of the original Orlicz-Pettis Theorem. It is also the case that the series converges in topologies which may be stronger than the original topology of E ([17, 4.10]). The theorem contains the result of Kalton for locally convex spaces and subseries convergent series since any locally convex F space is an infra-Pták space; it should be pointed out that Kalton's result applies to non-locally convex spaces. There are stronger results for subseries convergent series in Banach spaces; if X is a Banach space which does not contain a copy of l^∞ and if M is a subspace of X' which separates the points of X, then any series $\sum_j x_j$ which is subseries convergent in $\sigma(X,M)$ is subseries convergent in X ([3, I.4.7],[6, 2.2]). The proofs of this stronger result or Kalton' theorem for subseries convergent series use results for vector measures or closed graph theorems and do not carry forward to the case of multiplier convergent series.

As the following example shows the separability assumption on E is important.

EXAMPLE (2). *Let $E = l^\infty$ with the sup-norm and $M = span\{e^j : j \in \mathbb{N}\}$, where e^j is the sequence with 1 in the j^{th} coordinate and 0 in the other coordinates. Then the series $\sum_j e^j$ is subseries convergent with respect to $\sigma(l^\infty, M)$ but is not subseries convergent with respect to the sup-norm.*

We indicate several applications of the theorem to sequence and function spaces. In the examples below we assume the multiplier space λ has signed WGHP.

EXAMPLE (3). *Let E be a complete metrizable locally convex space with a Schauder basis $\{b_j\}$ and coordinate functionals $\{f_j\}$: that is, each $x \in E$ has a unique series expansion, $x = \sum_{j=1}^{\infty} t_j b_j$, with the coordinate functionals defined by $f_j(x) = t_j$. Such a space is separable. The coordinate functionals are continuous in this instance and we set $M = span\{f_j : j \in \mathbb{N}\}$. Then any series $\sum_{j=1}^{\infty} x_j$ in E which is subseries convergent in $\sigma(E,M)$ is λ multiplier convergent in E. A theorem of this type for non-locally convex, complete metric linear spaces was established for subseries convergent series by Stiles ([14]). A version for multiplier convergent series can be found in* [17, 4.74, 9.10].

EXAMPLE (4). *Let E be a separable, complete, metrizable scalar sequence space which is a K-space, that is, the coordinate functionals $f_j : \{t_j\} \to t_j$ are continuous from E to the scalar field for every j. Then $M = span\{f_j : j \in \mathbb{N}\} \subset E'$ separates the points of E so the theorem applies and any series in E which is λ multiplier convergent in $\sigma(E,M)$ (that is, coordinatewise convergent) is λ multiplier convergent in E. This result applies to any AK-space, i.e., each sequence $t = \{t_j\} \in E$ is represented by the series $t = \sum_{j=1}^{\infty} t_j e^j$, where e^j is the sequence with 1 in the j^{th} coordinate and 0 in the other coordinates, since any AK-space is separable. For example, the spaces c_0, l^p $(1 \le p < \infty), cs, bv, bv_0$ are AK-spaces (see* [2, 17] *for lists* *of sequence spaces). The result also applies to the space c which is not an AK-space and to certain domain spaces such as c_A (see* [2, 8.1.6]*).*

EXAMPLE (5). *Let S be a compact, metric space and $C(S)$ the space of continuous functions with the sup-norm. Such a space is separable (*[8, p. 245]*). Let δ_t be the Dirac measure concentrated at t and set $M = span\{\delta_t : t \in D\}$, where D is a dense subset of S. Then M is a subset of the dual of $C(S)$ which separates points. Thus, a series $\sum_j f_j$ in $C(S)$ which is pointwise λ multiplier convergent on D is λ multiplier convergent with respect to the sup-norm. Results of this type for subseries convergent series were established by Thomas (*[18]*); see* [17, 4.68] *for multiplier convergent versions. When $S = D$ the method employed in* [16, 10.4.7] *can be used to remove the metrizability assumption on S.*

EXAMPLE (6). *Let $C(\mathbb{R}^n)$ be the space of all continuous real valued functions on $\mathbb{R}^n$ with the topology of uniform convergence on compact subsets. With this topology $C(\mathbb{R}^n)$ is a separable F-space and if D is a dense subset of $\mathbb{R}^n$, the set $M = \{\delta_x : x \in D\}$ separates the points of $C(\mathbb{R}^n)$. Therefore, the theorem applies and any series which is pointwise λ multiplier convergent on D is λ multiplier convergent in $C(\mathbb{R}^n)$. The same remarks apply to $\mathcal{E}$, the space of entire functions, $f : \mathbb{C} \to \mathbb{C}$, with the topology of uniform convergence on compact subsets of $\mathbb{C}$.*

There are also similar applications to vector valued sequence and function spaces.

EXAMPLE (7). *Let X be a separable, complete, metrizable locally convex space and E a vector space of X valued sequences which is a complete, metrizable locally convex space. We assume that E is an AK-space in the sense that every $x = \{x_j\} \in E$ has a series expansion $x = \sum_{j=1}^{\infty} e^j \otimes x_j$, where $e^j \otimes x_j$ is the sequence with x_j in the j^{th} coordinate and 0 in the other coordinates. Then E is a separable space. Let*

$$M = span\{e^j \otimes x' : j \in \mathbb{N}, x' \in X'\} \subset E'.$$

Then M separates the points of E and a sequence in E converges to 0 with respect to $\sigma(E,M)$ iff the sequence converges coordinatewise with respect to the weak topology on X. Thus, a series which is λ multiplier convergent with respect to $\sigma(E,M)$ is λ multiplier convergent in E. A similar result is established in 4.76 of [17].

The example applies to such sequence spaces as $c_0(X), l^p(X)$ $(1 \le p < \infty)$ when X is a complete metrizable space.

EXAMPLE (8). *Let S be a complete metric space and X a separable Banach space. Let $C_X(S)$ be the space of all X valued continuous functions defined on S with the sup-norm. Then $C_X(S)$ is separable. Let*

$$M = span\{\delta_t \otimes x' : t \in D \subset S, x' \in X'\},$$

where $D \subset S$ is dense and $\delta_t \otimes x' \in C_X(S)'$ is defined by $\delta_t \otimes x'(f) = x'(f(t))$. A sequence in $C_X(S)$ converges with respect to $\sigma(C_X(S),M)$ iff the sequence is pointwise convergent on D with respect to the weak topology of X. Then M separates the points of $C_X(S)$ so the theorem applies.

EXAMPLE (9). *Suppose X,Y are Banach spaces and $L(X,Y)$ is the space of all continuous linear operators from X into Y with the operator norm. Let*

$$M = span\{y' \otimes x : y' \in Y', x \in X\},$$

where $y' \otimes x \in L(X,Y)'$ is defined by $y' \otimes x(T) = y'(Tx)$. Then M separates the points of $L(X,Y)$ and convergence in $\sigma(L(X,Y),M)$ is just convergence in the weak operator topology. Thus, If E is any separable subspace of $L(X,Y)$, the theorem applies and any series in E which is λ multiplier convergent in the weak operator topology is λ multiplier convergent in the operator norm.

Received December 17, 2012

Final version received April 19, 2013

CHARLES SWARTZ
MATHEMATICS DEPARTMENT
NEW MEXICO STATE UNIVERSITY
LAS CRUCES, NM 88003, USA
cswartz@nmsu.edu

REFERENCES

[1] E. BECKENSTEIN AND L. NARICI, Topological Vector Spaces, Marcel Dekker, NY, 1985.
[2] J. BOOS, Classical and Modern Methods in Summability, Oxford Univ. Press, 2000.
[3] J. DIESTEL AND J. UHL, Vector Measures, Amer. Math. Soc., Providence, 1977.
[4] W. FILTER AND L. LABUDA, *Essays on the Orlicz-Pettis Theorem I*, Real Anal. Exchange **16**, (1990/1991), 393–403.
[5] J. HORVATH, Topological Vector Spaces and Distributions, Addison-Wesley, Reading, 1966.
[6] N. KALTON, The Orlicz-Pettis Theorem, Contemporary Math., Amer. Math. Soc., Providence, 1980.
[7] N. KALTON, *Subseries Convergence in Topological Groups and Vector Spaces*, Israel J. Math. **10**, (1971), 402–412.
[8] J. KELLEY, General Topology, Van Nostrand, Princeton, 1955.
[9] G. KÖTHE, Topological Vector Spaces I, Springer-Verlag, Berlin, 1983.
[10] G. KÖTHE, Topological Vector Spaces II, Springer-Verlag, Berlin, 1979.
[11] D. NOLL, *Sequential Completeness and Spaces with the Gliding Hump Property*, Manuscripta Math. **66**, (1990), 237–252.
[12] W. ORLICZ, *Beitrage zur Theoreie der Orthogonalent Wichlungen II*, Studia Math. **1**, (1929), 241–255.
[13] B. J. PETTIS, *On Integration in Vector Spaces*, Trans. Amer. Math. Soc. **44**, (1938), 277–304.
[14] W. J. STILES, *On Subseries Convergence in F-Spaces*, Israel J. Math. **8** (1970), 53–56.
[15] C. STUART, *Weak Sequential Completeness of β-duals*, Rocky Mt. J. Math. **26**, (1996), 1559–1568.
[16] C. SWARTZ, An Introduction to Functional Analysis, Marcel Dekker, NY, 1992.
[17] C. SWARTZ, Multiplier Convergent Series, World Sci. Publ., Singapore, 2009.
[18] G. E. F. THOMAS, *L'integration par rapport a une mesure de Radon vectorielle*, Ann. Inst. Fourier **20**, (1970), 55–191.

Bol. Soc. Mat. Mexicana (3) Vol. 19, 2013

IDEALS WITH AT MOST COUNTABLE HULL IN CERTAIN ALGEBRAS OF FUNCTIONS ANALYTIC ON THE HALF-PLANE

ANDRZEJ SOŁTYSIAK AND ANTONI WAWRZYŃCZYK

ABSTRACT. We describe all closed ideals with at most countable hull in the algebras $\mathcal{A}^{(\alpha)}(\mathbb{C}^+)$ ($\alpha > 0$) of analytic functions on the complex half-plane.

1. Introduction

In the sequel by a Banach algebra we mean a Banach space endowed with a jointly continuous multiplication. Hence if $(\mathcal{B}, \|\cdot\|)$ is a Banach algebra, then a constant $C > 0$ in the inequality $\|ab\| \le C\|a\|\|b\|$ $(a, b \in \mathcal{B})$ need not be equal to one. However, it is always possible to introduce a submultiplicative norm equivalent to the given one and such that it takes value one at a unit, if $\mathcal{B}$ is unital (see e.g. [8], p. 10).

The space $\mathcal{A}^{(n)}(\mathbb{C}^+)$ $(n \in \mathbb{N})$ is the set of functions F analytic on the right half-plane $\mathbb{C}^+ = \{z \in \mathbb{C} \colon \operatorname{Re} z > 0\}$ continuously extendable to $i\mathbb{R}$, whose derivatives $F^{(k)}$ are continuous on $\overline{\mathbb{C}^+} \setminus \{0\}$ and satisfy $\lim_{z\to 0} z^k F^{(k)}(z) = 0$ for $1 \le k \le n$, while $\lim_{z\to\infty} z^k F^{(k)}(z) = 0$, $0 \le k \le n$ (we denote $F^{(0)} = F$).

For a bounded function F on $\mathbb{C}^+$ let $\|F\|_\infty = \sup_{z\in\mathbb{C}^+} |F(z)|$. Provided with the norm $\|F\|_{(n)} = \sum_{j=0}^{n} \|\zeta^j F^{(j)}\|_\infty$ (ζ stands for the identity function $z \mapsto z$) and the pointwise multiplication the space $\mathcal{A}^{(n)}(\mathbb{C}^+)$ is a Banach algebra in the sense described above. Notice that the norm $\|F\|_{(n)}$ is equivalent to the norm $\|F\|_n = \|F\|_\infty + \|\zeta^n F^{(n)}\|_\infty$ (see [4], Prop. 3.3 and Rem. 3.6).

The space $\mathcal{A}^{(0)}(\mathbb{C}^+)$ is the set of continuous functions on $\overline{\mathbb{C}^+}$ vanishing at infinity and analytic on the half-plane $\mathbb{C}^+$. It is a Banach algebra with the pointwise multiplication and the norm $\|F\|_\infty = \sup_{z\in\mathbb{C}^+} |F(z)|$.

The algebras $\mathcal{A}^{(n)}(\mathbb{C}^+)$ $(n \in \mathbb{N}_0)$ as well as the algebras $\mathcal{A}^{(\alpha)}(\mathbb{C}^+)$ for $\alpha > 0$ (see the next section) appeared in the paper [4] as the spaces of Gelfand transforms of "fractional convolution algebras" of functions on the half-line $\mathbb{R}^+ = (0, +\infty)$.

Taking into account that the functions F and $\zeta^n F^{(n)}$ are analytic on $\mathbb{C}^+$, vanish at infinity, and are continuous at zero, it follows by the maximum principle that the norm $\|F\|_n$ is equal to $\sup_{x\in\mathbb{R}} |F(ix)| + \sup_{x\in\mathbb{R}} |x|^n |F^{(n)}(ix)|$.

Closed ideals of the algebra $\mathcal{A}^{(n)}(\mathbb{C}^+)$ were described in [5]. In the present paper we study the ideals of the algebras $\mathcal{A}^{(\alpha)}(\mathbb{C}^+)$ $(\alpha > 0)$.

2. Algebras $\mathcal{A}^{(\alpha)}(\mathbb{C}^+)$

The spaces $\mathcal{A}^{(\alpha)}(\mathbb{C}^+)$, $\alpha > 0$, are defined by means of the fractional complex derivation introduced in [4].

2010 *Mathematics Subject Classification*: Primary: 46J15; Secondary: 26A33, 30H50.

Keywords and phrases: Banach algebras, closed ideals, holomorphic functions, fractional complex derivation.

For $F \in \mathcal{A}^{(n)}(\mathbb{C}^+)$, $\alpha > 0$, $n = [\alpha]+1$, and $z = re^{i\theta} \in \overline{\mathbb{C}^+}$ the complex α-derivative of F is given by the formula

$$W^\alpha F(z) = \frac{(-1)^n}{\Gamma(n-\alpha)} \int_0^\infty t^{n-\alpha-1} F^{(n)}(z + te^{i\theta})\,dt.$$

It should be mentioned that the integral in the above formula is independent of θ, $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ ([4], Lemma 3.1). Therefore in the definition of $W^\alpha F(z)$ the integration can be performed along an arbitrary ray starting at z and contained in $\overline{\mathbb{C}^+}$.

In particular, for $z = ix$ ($x \in \mathbb{R}$) we can use the formulas

$$W^\alpha F(ix) = \frac{(-1)^n}{\Gamma(n-\alpha)} \int_0^\infty t^{n-\alpha-1} F^{(n)}(i(x+t))\,dt$$

when $x \geq 0$ and

$$\begin{aligned} W^\alpha F(ix) &= \frac{(-1)^n}{\Gamma(n-\alpha)} \int_0^\infty t^{n-\alpha-1} F^{(n)}(i(x-t))\,dt \\ &= \frac{1}{\Gamma(n-\alpha)} \int_0^\infty t^{n-\alpha-1} \check{F}^{(n)}(i(x+t))\,dt \end{aligned}$$

when $x < 0$. We have denoted $\check{F}(x) = F(-x)$. The latter formulas are useful because the fractional derivation on the half-line was extensively studied.

For an arbitrary $F \in \mathcal{A}^{(n)}(\mathbb{C}^+)$ let

$$\|F\|_\alpha = \|F\|_\infty + \sup_{z \in \mathbb{C}^+} |z|^\alpha |W^\alpha F(z)|.$$

By arguments used in the case of the norm $\|\cdot\|_n$ we can represent

$$\begin{aligned} \|F\|_\alpha &= \sup_{x \in \mathbb{R}} |F(ix)| + \sup_{x \in \mathbb{R}} |x|^\alpha |W^\alpha F(ix)| \\ &= \sup_{x \in \mathbb{R}} |F(ix)| + \max\{\sup_{x \geq 0} x^\alpha |W^\alpha F(ix)|, \sup_{x \geq 0} x^\alpha |W^\alpha \check{F}(ix)|\}. \end{aligned}$$

If a function G is defined on $i\mathbb{R}$ and the right-hand side of the latter formula makes sense for G we shall use the notation

$$\|G\|_{\alpha,\mathbb{R}} = \sup_{x \in \mathbb{R}} |G(ix)| + \sup_{x \in \mathbb{R}} |x|^\alpha |W^\alpha G(ix)|.$$

For functions on $\mathbb{R}^+$ we introduce

$$\|G\|_{\alpha,\mathbb{R}^+} = \sup_{x > 0} |G(x)| + \sup_{x > 0} x^\alpha |W^\alpha G(x)|.$$

The space $\mathcal{A}^{(\alpha)}(\mathbb{C}^+)$ is defined as the completion of the space $\mathcal{A}^{(n)}(\mathbb{C}^+)$ in the norm $\|\cdot\|_\alpha$ ($n = [\alpha]+1$).

Propositions 3.5 and 3.8 from [4] provide the following properties of the family of spaces $\mathcal{A}^{(\alpha)}(\mathbb{C}^+)$, $\alpha > 0$.

THEOREM (2.1). [4] (i) *For every $\alpha > 0$ the space $\mathcal{A}^{(\alpha)}(\mathbb{C}^+)$ is a Banach algebra under the pointwise multiplication.*

(ii) *For $\beta \geq \alpha \geq 0$ there is a constant $C_{\alpha\beta} > 0$ such that*

$$\|F\|_\alpha \leq C_{\alpha\beta} \|F\|_\beta.$$

Consequently, there is a natural continuous embedding $\mathcal{A}^{(\beta)}(\mathbb{C}^+) \hookrightarrow \mathcal{A}^{(\alpha)}(\mathbb{C}^+)$ with a dense range.

(iii) *The norm* $\|\cdot\|_\alpha$ *is equivalent to the norm given by the formula*

$$\|F\|'_{(\alpha)} = \|F\|_\infty + \sup_{\beta\le\alpha}\sup_{z\in\mathbb{C}^+} |z|^\beta |W^\beta F(z)|.$$

By Theorem (2.1) (iii) the norm $\|F\|_n = \|F\|_\infty + \|\zeta^n F^{(n)}\|_\infty$ is equivalent to the norm $\|F\|'_n = \max_{0\le k\le n}\|\zeta^k F^{(k)}\|_\infty$.

The algebras $\mathcal{A}^{(\alpha)}(\mathbb{C}^+)$ are not unital, however the unit can be attached in the standard way. The norm in the algebra $\mathcal{A}^{(\alpha)}(\mathbb{C}^+)\oplus\mathbb{C}$ is defined by $\|F+c\|_\alpha = \|F\|_\alpha + |c|$, where $F\in\mathcal{A}^{(\alpha)}(\mathbb{C}^+)$ and c is a constant.

The algebra $\mathcal{A}^{(\alpha)}(\mathbb{C}^+)\oplus\mathbb{C}$ will be denoted by $\mathcal{A}^{(\alpha)}_u(\mathbb{C}^+)$. It is a unital Banach algebra in the sense mentioned at the very beginning of the paper.

The space $\mathfrak{M}_\alpha$ of multiplicative linear functionals on $\mathcal{A}^{(\alpha)}(\mathbb{C}^+)$ consists of functionals of the point evaluations, hence it can be identified with $\overline{\mathbb{C}^+}$. The space $\mathfrak{M}_{\alpha,u}$ of multiplicative linear functionals on the algebra $\mathcal{A}^{(\alpha)}_u(\mathbb{C}^+)$ contains an additional functional $\phi_\infty(F) = \lim_{z\to\infty}F(z)$.

PROPOSITION (2.2). *Let* $\alpha>0$. *There exists a constant* $C>0$ *such that for every* $F\in\mathcal{A}^{(\alpha)}(\mathbb{C}^+)$ *and every rational function* G *bounded on* $i\mathbb{R}$ *and such that* $FG\in\mathcal{A}^{(\alpha)}(\mathbb{C}^+)$ *the following inequality is valid*:

$$\|FG\|_\alpha \le C\|F\|_\alpha\|G\|_{\alpha,\mathbb{R}}.$$

Proof. As mentioned above, similar properties were studied in the case of functions on $\mathbb{R}^+$, hence our purpose is to reduce the proof to that case.

In [2] it was introduced a Banach space $\mathcal{M}^{(\alpha)}_\infty$ as the completion of the bounded functions of class $C^{(\infty)}$ on $\mathbb{R}^+$ in the space of the so-called functions of weak bounded variation. $\mathcal{M}^{(\alpha)}_\infty$ is called the Mikhlin space. The norm in $\mathcal{M}^{(\alpha)}_\infty$ is just $\|\cdot\|_{\alpha,\mathbb{R}^+}$. It was proved that the Mikhlin space is a Banach algebra.

For a function F defined on $i\mathbb{R}$ we denote by $\widetilde{F}$ the function $\mathbb{R}^+\ni x\mapsto F(ix)$. If G is a rational function bounded on $i\mathbb{R}$, then the function $\widetilde{G}$ belongs to $\mathcal{M}^{(\alpha)}_\infty$.

There exists a constant C_1 depending only of α, such that

$$\|\widetilde{FG}\|_{\alpha,\mathbb{R}^+} \le C_1\|\widetilde{F}\|_{\alpha,\mathbb{R}^+}\|\widetilde{G}\|_{\alpha,\mathbb{R}^+}.$$

Taking into account that F and $FG\in\mathcal{A}^{(\alpha)}(\mathbb{C}^+)$ we obtain

$$\begin{aligned}
\|FG\|_\alpha = \|FG\|_{\alpha,\mathbb{R}} &= \sup_{x\in\mathbb{R}}|FG(ix)| + \sup_{x\in\mathbb{R}}|x|^\alpha|W^\alpha(FG)(ix)| \\
&\le \sup_{x\in\mathbb{R}^+}|FG(ix)| + \sup_{x\in\mathbb{R}^+}x^\alpha|W^\alpha(FG)(ix)| + \sup_{x\in\mathbb{R}^+}|\check{FG}(ix)| \\
&+ \sup_{x\in\mathbb{R}^+}x^\alpha|W^\alpha(\check{FG})(ix)| = \|\widetilde{FG}\|_{\alpha,\mathbb{R}^+} + \|\widetilde{\check{FG}}\|_{\alpha,\mathbb{R}^+} \\
&\le C_1(\|\widetilde{F}\|_{\alpha,\mathbb{R}^+}\|\widetilde{G}\|_{\alpha,\mathbb{R}^+} + \|\widetilde{\check{F}}\|_{\alpha,\mathbb{R}^+}\|\widetilde{\check{G}}\|_{\alpha,\mathbb{R}^+}) \\
&\le C_1(\|\widetilde{F}\|_{\alpha,\mathbb{R}^+} + \|\widetilde{\check{F}}\|_{\alpha,\mathbb{R}^+})(\|\widetilde{G}\|_{\alpha,\mathbb{R}^+} + \|\widetilde{\check{G}}\|_{\alpha,\mathbb{R}^+}) \\
&\le 4C_1\|F\|_{\alpha,\mathbb{R}}\|G\|_{\alpha,\mathbb{R}} = 4C_1\|F\|_\alpha\|G\|_{\alpha,\mathbb{R}}. \quad \square
\end{aligned}$$

3. Algebras $\mathcal{A}_u^{(\alpha)}(\mathbb{D})$

The papers [1] and [7] provide the characterization of closed ideals with at most countable hull of subalgebras of the disc algebra $A(\mathbb{D})$ satisfying certain conditions. We use these results for the description of ideals with at most countable hull in the algebras $\mathcal{A}_u^{(\alpha)}(\mathbb{C}^+)$.

The Möbius transform $m(w) = \dfrac{1+w}{1-w}$ carries the disc $\mathbb{D}$ onto the half-plane $\mathbb{C}^+$ and the circle $\mathbb{T}$ onto $i\mathbb{R}\cup\{\infty\}$. The inverse of m is the function $z \mapsto \dfrac{z-1}{z+1}$.

If F is a function on $\overline{\mathbb{C}^+}$ which has a limit at ∞, then the formulas $f = F\circ m$, $f(1) = \lim_{z\to\infty} F(z)$ define a function on $\overline{\mathbb{D}}$.

From the formulas obtained in the proof of Lemma 2.2 in [4] we have the following.

PROPOSITION (3.1). *There exist positive constants $\{c_j\}$ and $\{d_j\}$ such that for every $f \in A(\mathbb{D})\cap C^{(k)}(\overline{\mathbb{D}}\setminus\{-1,1\})$ and $F = f\circ m^{-1}$ we have*

$$\left|(1-w^2)^k f^{(k)}(w)\right| \le \sum_{j=0}^{k} c_j \left|F^{(j)}\left(\frac{1+w}{1-w}\right)\right| \left|\frac{1+w}{1-w}\right|^j$$

and

$$\left|F^{(k)}\left(\frac{1+w}{1-w}\right)\right| \left|\frac{1+w}{1-w}\right|^k \le \sum_{j=0}^{k} d_j \left|(1-w^2)^j f^{(j)}(w)\right|,$$

$w \in \overline{\mathbb{D}}\setminus\{-1,1\}$.

Let us define

$$\mathcal{A}_u^{(\alpha)}(\mathbb{D}) = \{f \in C(\mathbb{D}) : f\circ m^{-1} \in \mathcal{A}_u^{(\alpha)}(\mathbb{C}^+)\}$$

and

$$\mathcal{A}^{(\alpha)}(\mathbb{D}) = \{f \in C(\mathbb{D}) :\ f\circ m^{-1} \in \mathcal{A}^{(\alpha)}(\mathbb{C}^+)\}.$$

The space $\mathcal{A}_u^{(\alpha)}(\mathbb{D})$ is obviously a unital algebra under the pointwise multiplication and a Banach algebra with the norm given by

$$\|f\|_{\alpha,\mathbb{D}} = \|f\circ m^{-1}\|_{\alpha}.$$

The algebra $\mathcal{A}^{(\alpha)}(\mathbb{D})$ is a maximal ideal in $\mathcal{A}_u^{(\alpha)}(\mathbb{D})$.

More explicit description of the space $\mathcal{A}_u^{(\alpha)}(\mathbb{D})$ as well as its norm can be obtained easily in the case of the natural α. In this case $W^n F(z) = (-1)^n F^{(n)}(z)$ and the corresponding space $\mathcal{A}^{(\alpha)}(\mathbb{C}^+)$ coincides with the space $\mathcal{A}^{(n)}(\mathbb{C}^+)$, which was defined in the introduction.

By Proposition (3.1) we obtain the following description of the algebra $\mathcal{A}^{(n)}(\mathbb{D})$.

PROPOSITION (3.2). [4] *For a positive integer n the algebra $\mathcal{A}^{(n)}(\mathbb{D})$ is isomorphic to the subalgebra of the disc algebra $A(\mathbb{D})$ of functions f such that $f(1) = 0$, $(\zeta^2-1)^k f^{(k)} \in A(\mathbb{D})$ and*

$$\lim_{z\to\pm1} (z^2-1)^k f^{(k)}(z) = 0, \quad 1 \le k \le n,$$

endowed with the norm $[f]_{(n),\mathbb{D}} = \sum_{k=0}^{n} \|(\zeta^2-1)^k f^{(k)}\|_\infty$.

As before the norms $[f]_{(n),\mathbb{D}}$ and $[f]_{n,\mathbb{D}} = \|f\|_\infty + \|(\zeta^2-1)^n f^{(n)}\|_\infty$ are equivalent.

The results obtained in the paper [7] concern subalgebras $\mathcal{B}$ of the disc algebra $A(\mathbb{D})$ that satisfy the following conditions.

(H1) The space of polynomials is a dense subset of $\mathcal{B}$.

(H2) $\lim_{k\to\infty} \|\zeta^k\|_{\mathcal{B}}^{\frac{1}{k}} = 1$.

(H3) There exist $k \geq 0$ and $C > 0$ such that

$$|1-|\lambda||^k \|f\|_{\mathcal{B}} \leq C\|(\zeta-\lambda)f\|_{\mathcal{B}}, \quad f \in \mathcal{B},\ |\lambda| < 2.$$

(D) For every $z_0 \in \mathbb{T}$ there exists $N(z_0) \in \mathbb{N}_0$ such that the functionals $\mathcal{B} \ni f \mapsto f^{(j)}(z_0)$ $(0 \leq j \leq N(z_0))$ are well-defined and continuous, and there exists a sequence (σ_n) in the algebra $\mathcal{B}$ such that $\sigma_n(z_0) = 0$ for all n and

$$\lim_{n\to\infty} \|(\zeta - z_0)^{N(z_0)+1}\sigma_n - (\zeta - z_0)^{N(z_0)+1}\|_{\mathcal{B}} = 0. \tag{3.3}$$

THEOREM (3.4). *For every* $\alpha > 0$ *the algebra* $\mathcal{A}_u^{(\alpha)}(\mathbb{D})$ *satisfies conditions* (H1)–(H3) *and* (D).

Proof. It is sufficient to prove the properties in question for an arbitrary norm equivalent to the one introduced in the space $\mathcal{A}_u^{(\alpha)}(\mathbb{D})$. Hence in each of the consecutive calculations we will use the most convenient norm.

Proposition 2.1 from [4] states that the polynomials in $(w-1)$ without a constant term are dense in the space $\mathcal{A}^{(n)}(\mathbb{D})$. Therefore the space of all polynomials is dense in the unital algebra $\mathcal{A}_u^{(n)}(\mathbb{D})$. On the other hand the latter space is by definition dense in $\mathcal{A}_u^{(\alpha)}(\mathbb{D})$ for $n-1 < \alpha < n$. This proves (H1).

For every natural n and $k \geq n$ we have

$$1 + k(k-1)\ldots(k-n+1) \leq [\zeta^k]_{n,\mathbb{D}} \leq 1 + 2^n k(k-1)\ldots(k-n+1).$$

Thus $\lim_{k\to\infty} [\zeta^k]_{n,\mathbb{D}}^{\frac{1}{k}} = 1$. If $n-1 < \alpha < n$ the norm $\|\cdot\|_{\alpha,\mathbb{D}}$ is dominated by $\|\cdot\|_{n,\mathbb{D}}$ and dominates $\|\cdot\|_{n-1,\mathbb{D}}$. This proves (H2).

To the operator of multiplication by $(\zeta - \lambda)$ of a function f on the disc, there corresponds the operator of multiplication by $(\zeta \circ m^{-1} - \lambda)$ of the function $f \circ m^{-1}$ on $\mathbb{C}^+$.

For $f \in \mathcal{A}^{(\alpha)}(\mathbb{D})$ we obtain by applying Proposition (2.2)

$$\begin{aligned}\|f\|_{\alpha,\mathbb{D}} &= \|f\circ m^{-1}\|_{\alpha} = \|f\circ m^{-1}(\zeta\circ m^{-1}-\lambda)(\zeta\circ m^{-1}-\lambda)^{-1}\|_{\alpha,\mathbb{R}} \\ &\leq \|(\zeta\circ m^{-1}-\lambda)^{-1}\|_{\alpha,\mathbb{R}}\|(\zeta\circ m^{-1}-\lambda)f\circ m^{-1}\|_{\alpha} \\ &\leq \|(\zeta\circ m^{-1}-\lambda)^{-1}\|_{n,\mathbb{R}}\|(\zeta-\lambda)f\|_{\alpha,\mathbb{D}},\end{aligned}$$

where $n = [\alpha]+1$.

By the definition of the norm $\|\cdot\|_{n,\mathbb{R}}$ and the second formula of Proposition (3.1) it follows that there exists $C > 0$ such that for $|\lambda| < 2$, $|\lambda| \neq 1$, we have

$$\begin{aligned}\|(\zeta\circ m^{-1}-\lambda)^{-1}\|_{n,\mathbb{R}} &= \sup_{z\in i\mathbb{R}} |(\zeta\circ m^{-1}-\lambda)^{-1}| + \sup_{z\in i\mathbb{R}} |z|^n (|(\zeta\circ m^{-1}-\lambda)^{-1})^{(n)}| \\ &\leq \sup_{w\in\mathbb{T}} |w-\lambda|^{-1} + \sum_{j=0}^{n} d_j \sup_{w\in\mathbb{T}} |w^2-1|^j \left|\frac{d^j}{dw^j}(w-\lambda)^{-1}\right| \\ &\leq C|1-|\lambda||^{-n-1}.\end{aligned}$$

Condition (H3) is satisfied in the form

$$(1-|\lambda|)^{[\alpha]+2}\|f\|_{\alpha,\mathbb{D}} \leq C\|(\zeta-\lambda)f\|_{\alpha,\mathbb{D}}.$$

Elements of the algebra $\mathcal{A}_u^{(n)}(\mathbb{D})$ are continuous on the closed disc and since the functions $g(z) = (z+1)^{\frac{1}{2}}$ and $h(z) = (1-z)^{\frac{1}{2}}$, which belong to these spaces, are not derivable at -1 and 1 respectively, we have $N(-1) = N(1) = 0$.

At all points $z_0 \in \mathbb{T} \setminus \{-1, 1\}$ the functions $f \in \mathcal{A}_u^{(n)}(\mathbb{D})$, $n \in \mathbb{N}$, have derivatives up to the order n and by the definition of the norm the functionals $f \mapsto f^{(k)}(z_0)$, $0 \le k \le n$, are continuous. Therefore the numbers $N(z_0)$ are equal to n for $z_0 \in \mathbb{T} \setminus \{-1, 1\}$ in the spaces $\mathcal{A}_u^{(n)}(\mathbb{D})$.

For general $\alpha > 0$ the algebra $\mathcal{A}_u^{(\alpha)}(\mathbb{D})$ is embedded continuously into $\mathcal{A}_u^{([\alpha])}(\mathbb{D})$ hence $N(1) = N(-1) = 0$ and for other points of the circle $N(z_0) = [\alpha]$.

Now, we proceed to construct the sequences (σ_m) satisfying (3.3). For natural $\alpha = n$ and $z_0 = 1$ define $\sigma_m(z) = (1-z)^{\frac{1}{m}}$. By simple calculations it is easy to verify that σ_m belongs to $\mathcal{A}_u^{(n)}(\mathbb{D})$. The condition $\sigma_m(1) = 0$ is satisfied. Since $N(1) = 0$, it suffices to show that $[(\zeta-1)\sigma_m - (\zeta-1)]_{n,\mathbb{D}} \to 0$ as $m \to \infty$. First, let us investigate the uniform convergence of $(\zeta-1)\sigma_m$. For every $0 < \varepsilon < 1$ and $|1-z| < \varepsilon$ the inequality $|(z-1)\sigma_m(z) - (z-1)| < 2\varepsilon$ is obvious. There exists $N \in \mathbb{N}$ such that $|\sigma_m(z) - 1| < \varepsilon$ for $|z-1| \ge \varepsilon$ and for all $m > N$. In this way we obtain $\|(\zeta-1)\sigma_m - (\zeta-1)\|_\infty < 2\varepsilon$ for $m > N$, that is

$$\lim_{m\to\infty} \|(\zeta-1)\sigma_m - (\zeta-1)\|_\infty = 0. \tag{3.5}$$

Since $((\zeta-1)\sigma_m - (\zeta-1))' = \left(1 + \frac{1}{m}\right)\sigma_m - 1$ we get

$$(\zeta^2-1)((\zeta-1)\sigma_m - (\zeta-1))' = (\zeta+1)(\zeta-1)\left(\left(1+\frac{1}{m}\right)\sigma_m - 1\right).$$

These formulas and (3.5) imply

$$\lim_{m\to\infty} \|(\zeta^2-1)((\zeta-1)\sigma_m - (\zeta-1))'\|_\infty = 0$$

and therefore

$$\lim_{m\to\infty} [(\zeta-1)\sigma_m - (\zeta-1)]_{1,\mathbb{D}} = 0.$$

For $n > 1$ we have

$$((\zeta-1)\sigma_m - (\zeta-1))^{(n)} = c_{nm}\sigma_m(\zeta-1)^{-n+1},$$

where the constant $c_{nm} \to 0$ as $m \to \infty$. Thus we have

$$(\zeta^2-1)^n \left((\zeta-1)\sigma_m - (\zeta-1)\right)^{(n)} = c_{nm}(\zeta+1)^n(\zeta-1)\sigma_m.$$

By (3.5) $(\zeta-1)\sigma_m$ tends uniformly to $\zeta - 1$, and since $c_{nm} \to 0$ as $m \to \infty$, we get

$$\lim_{m\to\infty} \|(\zeta^2-1)^n((\zeta-1)\sigma_m - (\zeta-1))^{(n)}\|_\infty = 0$$

and so

$$\lim_{m\to\infty} [(\zeta-1)\sigma_m - (\zeta-1)]_{n,\mathbb{D}} = 0.$$

Condition (D) is satisfied at the point 1 in all spaces $\mathcal{A}_u^{(n)}(\mathbb{D})$. In a similar way it follows that this condition is also satisfied in each of the spaces $\mathcal{A}_u^{(n)}(\mathbb{D})$ at the point $z_0 = -1$ if we define $\sigma_m(z) = (z+1)^{\frac{1}{m}}$. For arbitrary $\alpha > 0$ the norm $\|\cdot\|_\alpha$ is dominated by $\|\cdot\|_{[\alpha]+1}$, so

$$\lim_{m\to\infty} \|(\zeta \pm 1)\sigma_m - (\zeta \pm 1)\|_{\alpha,\mathbb{D}} = 0.$$

Hence condition (D) holds true in both points 1 and -1 in all spaces $\mathcal{A}_u^{(\alpha)}(\mathbb{D})$, $\alpha > 0$.

We proceed to study validity of this condition at the points $z_0 \in \mathbb{T} \setminus \{-1, 1\}$ in the case of the spaces $\mathcal{A}_u^{(n)}(\mathbb{D})$, $n \in \mathbb{N}$. Then $N(z_0) = n$.

For $r>0$ let $\sigma_r(z)=\frac{z-z_0}{z-z_0-rz_0}$. We have

$$(z-z_0)^{n+1}\sigma_r(z)-(z-z_0)^{n+1}=\frac{rz_0(z-z_0)^{n+1}}{z-z_0-rz_0}.$$

By the configuration existing between the points $z-z_0$ and rz_0 the inequality $|z-z_0-rz_0|\geq\max\{|z-z_0|,r|z_0|\}$ holds true. It is clear that

$$\lim_{r\to 0}\|(\zeta-z_0)^{n+1}\sigma_r-(\zeta-z_0)^{n+1}\|_\infty=0.$$

By the Leibniz formula we get

$$\left((z-z_0)^{n+1}\sigma_r(z)-(z-z_0)^{n+1}\right)^{(n)}=rz_0\sum_{m=0}^{n}c_{nm}(z-z_0)^{n-m+1}(z-z_0-rz_0)^{m-n-1}$$

for some constants c_{nm}. Therefore

$$\sup_{z\in\mathbb{D}}\left|(z^2-1)^n\left((z-z_0)^{n+1}\sigma_r(z)-(z-z_0)^{n+1}\right)^{(n)}\right|\leq r\sup_{z\in\mathbb{D}}|(z^2-1)^n|\sum_{m=0}^{n}|c_{nm}|.$$

It follows that

$$\lim_{r\to 0}[(\zeta-z_0)^{n+1}\sigma_r(z)-(\zeta-z_0)^{n+1}]_{n,\mathbb{D}}=0$$

which means that condition (D) is satisfied in all algebras $\mathcal{A}_u^{(n)}(\mathbb{D})$, $n\in\mathbb{N}$.

The function $\frac{w-w_0}{w+1}$ on the half-plane $\mathbb{C}^+$ with $w_0\in i\mathbb{R}$ corresponds up to a constant coefficient to the function $\zeta-z_0$, $z_0\in\mathbb{T}$, on the disc under the Möbius map m.

Let $\sigma_r(w)=\frac{w-w_0}{w-w_0+r}$ for $r>0$. Notice that $|w-w_0+r|\geq\max\{r,|w-w_0|)$ and $|w+1|>1$. For $w_0\in i\mathbb{R}$ and $\alpha\notin\mathbb{N}$ we have $N(w_0)=[\alpha]$. Let $N(w_0)+1=n$ and let

$$F_r(w)=\left(\frac{w-w_0}{w+1}\right)^n\sigma_r(w)-\left(\frac{w-w_0}{w+1}\right)^n=\frac{-r(w-w_0)^n}{(w-w_0+r)(w+1)^n}.$$

We must prove that this function tends to zero when $r\to 0$ in the space $\mathcal{A}_u^{(\alpha)}(\mathbb{C}^+)$. The uniform convergence is obvious, so it remains to prove that $|w|^\alpha|W^\alpha F_r(w)|$ also converges to zero uniformly on $\mathbb{C}^+$.

To simplify calculations of the derivative of this function we introduce $G(h)=\frac{h^n}{(h+r)(h+u)^n}$, where $h=w-w_0$ and $u=1+w_0$. Then by the Leibniz formula we obtain

$$G^{(n)}(h)=\sum_{k=0}^{n}\sum_{j=0}^{n-k}b_{nkj}(h+r)^{-j-1}(h+u)^{-2n+j+k}h^{n-k},$$

where b_{nkj} are appropiate constants. The corresponding terms of this sum behave at infinity as h^{-n-1}.

The proof of Proposition 3.3 (ii) in [4] provides the following estimate for $|w|^\alpha|W^\alpha F(w)|$, $\alpha>0$, $n=[\alpha]+1$, $F\in\mathcal{A}^{(n)}(\mathbb{C}^+)$:

$$|w|^\alpha|W^\alpha F(w)|\leq\frac{1}{\Gamma(n-\alpha)}\int_1^\infty(t-1)^{n-\alpha-1}|w|^n|F^{(n)}(tw)|\,dt. \tag{3.6}$$

For $F_r(w)=\frac{-r(w-w_0)^n}{(w-w_0+r)(w+1)^n}$ we obtain

$$|w|^\alpha|W^\alpha F_r(w)|\leq r\sum_{k=0}^{n}\sum_{j=0}^{n-k}|c_{nkj}|\frac{1}{\Gamma(n-\alpha)}\int_1^\infty(t-1)^{n-\alpha-1}t^{-n}|\Phi(t,w,w_0)|\,dt,$$

where

$$\Phi(t,w,w_0)=(tw)^n(tw-w_0+r)^{-j-1}(tw+1)^{-2n+j+k}(tw-w_0)^{n-k}.$$

By the previous observation the function $\Phi(t,w,w_0)$ behaves at infinity as $\frac{1}{wt}$. It follows that each term of the sum is less or equal to $C\int_1^\infty (t-1)^{n-\alpha-1}t^{-n}\,dt = \widetilde{C}$, where the constant C depends only on w_0, n, k, j. The convergence

$$\lim_{r\to 0}\left\|\left(\frac{\zeta-w_0}{\zeta+1}\right)^n\sigma_r-\left(\frac{\zeta-w_0}{\zeta+1}\right)^n\right\|_\alpha=0$$

is proved. The condition (D) is satisfied in every space $\mathcal{A}_u^{(\alpha)}(\mathbb{D})$, $\alpha>0$. □

4. Ideals with at most countable hull

Let $\mathcal{B}$ be a semi-simple Banach algebra which is a unital subalgebra of the disc algebra $\mathcal{A}(\mathbb{D})$ and let $f\in\mathcal{B}$. The set $h(f)=\{z\in\overline{\mathbb{D}}: f(z)=0\}$ is called the hull of f. If I is an ideal of $\mathcal{B}$ then the hull of I is the set

$$h(I)=\bigcap_{f\in I}h(f).$$

Obviously $h(I)$ is a closed set. For $j\in\mathbb{N}_0$ let us define

$$h^j(I)=\{z\in\mathbb{T}:\ N(z)\geq j \text{ and } f(z)=f'(z)=\ldots=f^{(j)}(z)=0 \text{ for all } f\in I\}.$$

Suppose that the algebra $\mathcal{B}$ satisfies condition (D). A function U from the Hardy space H^∞ is called inner if its boundary function equals to one almost everywhere on $\mathbb{T}$. The inner function U divides $f\in H^\infty$ (denoted by $U\,|\,f$) if f/U is a bounded function.

Let $\mathcal{H}=\{H^0,H^1,\ldots\}$ be a descending family of sets $H^j\subset\mathbb{T}$ such that $H^j\subset\{z\in\mathbb{T}: N(z)\geq j\}$. For a given inner function U we define

$$I(U;\mathcal{H})=\{f\in\mathcal{B}:\ U|f \text{ and } f^{(j)}(z)=0 \text{ for } z\in H^j,\ j\in\mathbb{N}_0\}.$$

Ideals of $\mathcal{B}$ which are of this form are called *standard*. By U_I we denote the greatest common inner divisor of all nonzero elements of the ideal I.

The following theorem was proved in [7].

THEOREM (4.1). *Let $\mathcal{B}$ be a semi-simple Banach algebra which is a unital subalgebra of the disc algebra $A(\mathbb{D})$. Suppose that $\mathcal{B}$ satisfies conditions* (H1), (H2), (H3), *and* (D). *If $I\subset\mathcal{B}$ is a closed ideal and $h(I)$ is at most countable, then I is a standard ideal*:

$$I=I(U_I;\mathcal{H}),$$

where $\mathcal{H}=\{h^j(I)\}$.

By Theorem (3.4) we obtain

COROLLARY (4.2). *For every $\alpha>0$ all closed ideals with at most countable hull of the algebra $\mathcal{A}_u^{(\alpha)}(\mathbb{D})$ are standard.*

In the case of the algebra $\mathcal{A}_u^{(\alpha)}(\mathbb{D})$ the number $N(z)$ is equal to 0 for $z=\pm1$ and is equal to $[\alpha]$ for other points of the circle. For a closed ideal I the points ±1 can appear only in the set $h^0(I)$. For $0<j\leq[\alpha]$ the sets $h^j(I)$ are relatively closed subsets of $\mathbb{T}\setminus\{-1,1\}$. The following theorem gives a more explicite description of closed ideals with at most countable hull in $\mathcal{A}_u^{(\alpha)}(\mathbb{D})$.

THEOREM (4.3). *Let I be a closed ideal in $\mathcal{A}_u^{(\alpha)}(\mathbb{D})$ such that $h(I)$ is at most countable. Then the ideal I is standard, i.e.*

$$I=I(U_I;h^0(I),h^1(I),\ldots,h^{[\alpha]}(I)).$$

In particular the ideal of $\mathcal{A}_u^{(\alpha)}(\mathbb{D})$ of the form $I(1;\{1\},\emptyset,\ldots,\emptyset)$ is just the algebra $\mathcal{A}^{(\alpha)}(\mathbb{D})$.

The principal object of our interest is the algebra $\mathcal{A}^{(\alpha)}(\mathbb{C}^+)$ which is a maximal ideal in the algebra $\mathcal{A}_u^{(\alpha)}(\mathbb{C}^+)$. Taking into account the isomorphism $\mathcal{A}_u^{(\alpha)}(\mathbb{C}^+) \cong \mathcal{A}_u^{(\alpha)}(\mathbb{D})$ we can obtain immediately the description of all closed ideals with at most countable hull in the algebra $\mathcal{A}^{(\alpha)}(\mathbb{C}^+)$. These ideals are in one to one correspondence with the ideals of $\mathcal{A}_u^{(\alpha)}(\mathbb{D})$ of the form $I(U;H^0,H^1,\ldots,H^{[\alpha]})$, where H^0 is at most countable and $1 \in H^0$.

The description will be more concrete if we include information about the form of inner functions on the half-plane.

If U is an inner function on the unit disc $\mathbb{D}$, then the function $\mathcal{U} = U \circ m^{-1}$ on $\mathbb{C}^+$ can be represented as a product $U \circ m^{-1} = BS$, where B is of the form

$$B(z) = \left(\frac{z-1}{z+1}\right)^k \prod_n \frac{|1-z_n^2|}{1-z_n^2} \cdot \frac{z-z_n}{z+\overline{z}_n},$$

while S is uniquely representable as

$$S(z) = e^{-\rho z} \exp\left(-\int_{\mathbb{R}} \frac{tz+i}{t+iz}\, d\mu(t)\right)$$

for some positive measure μ on $\mathbb{R}$ singular with respect to the Lebesgue measure and $\rho \ge 0$. Hence, the set of zeros of the function B coincides with the set of zeros of $U \circ m^{-1}$ in the open half-plane $\mathbb{C}^+$ and the set of zeros of the factor S consists of zeros of $U \circ m^{-1}$ on the imaginary axis $i\mathbb{R}$ and is equal to the support of μ (see [6], p. 132).

Functions on $\mathbb{C}^+$ of the form $\mathcal{U} = BS$ are called inner functions on the half-plane.

Now, let $\mathcal{U}$ be an inner functions on the half-plane. Let $\mathcal{H} = \{H^j\}_{j=0}^{[\alpha]}$ be an arbitrary descending family of closed subsets of $i\mathbb{R}$ such that $0 \notin H^j$ for $j > 0$. Let us denote

$$\mathcal{I}(\mathcal{U};\mathcal{H}) = \{F \in \mathcal{A}^{(\alpha)}(\mathbb{C}^+) \colon \mathcal{U} \,|\, F \text{ and } F^{(j)} \text{ vanishes on } H^j \text{ for } j = 0,1,\ldots,[\alpha]\}.$$

If $\mathcal{I}$ is a closed ideal in $\mathcal{A}^{(\alpha)}(\mathbb{C}^+)$, then we denote

$$H^0(\mathcal{I}) = \{z \in i\mathbb{R} \colon F(z) = 0 \text{ for all } F \in \mathcal{I}\},$$

and for $0 < j \le [\alpha]$

$$H^j(\mathcal{I}) = \{z \in i\mathbb{R}\setminus\{0\} \colon F(z) = F'(z) = \ldots = F^{(j)}(z) = 0 \text{ for all } F \in \mathcal{I}\}.$$

THEOREM (4.4). *Let $\mathcal{I}$ be a closed ideal in $\mathcal{A}^{(\alpha)}(\mathbb{C}^+)$ such that $H^0(\mathcal{I})$ is at most countable. Let $\mathcal{U}_{\mathcal{I}}$ be the greatest common inner divisor of all nonzero elements of $\mathcal{I}$. Then*

$$\mathcal{I} = \mathcal{I}(\mathcal{U}_{\mathcal{I}}; H^0(\mathcal{I}),\ldots,H^{[\alpha]}(\mathcal{I})).$$

In particular, we have a characterization of dense ideals in $\mathcal{A}^{(\alpha)}(\mathbb{C}^+)$ which can be called a Nyman-type theorem (see [3]).

COROLLARY (4.5). *An ideal $\mathcal{J} \subset \mathcal{A}^{(\alpha)}(\mathbb{C}^+)$ is dense if and only if the following two conditions are satisfied*:

1. $H^0(\mathcal{J}) = \emptyset$,
2. *for every $a > 0$ there exists $f \in \mathcal{J}$ such that $f e^{a\zeta}$ is not bounded on $\mathbb{C}^+$.*

Proof. The first condition is obviously necessary. The second one says that the inner function $e^{-a\zeta}$ is not a common inner divisor of elements of $\mathcal{J}$, so it is also a neccessary condition for the density of $\mathcal{J}$.

On the other hand condition 1 implies that the greatest common inner divisor $\mathcal{U}_{\mathcal{J}}$ nowhere wanishes in $\mathbb{C}^+$. The only inner functions without zeros in $\mathbb{C}^+$ are the exponentials $e^{a\zeta}$, $a \le 0$.

If condition 2 is satisfied, then $\mathcal{U}_{\mathcal{J}} = 1$, hence $\overline{\mathcal{J}} = \mathcal{I}(1;\emptyset,\dots,\emptyset) = \mathcal{A}^{(\alpha)}(\mathbb{C}^+)$. $\square$

Received June 13, 2012

Final version received November 15, 2012

Andrzej Sołtysiak
Faculty of Mathematics and Computer Science,
Adam Mickiewicz University,
Ul. Umultowska 87, 61–614 Poznań, Poland
asoltys@amu.edu.pl

Antoni Wawrzyńczyk
Departamento de Matemáticas,
Universidad Autónoma Metropolitana-Iztapalapa
AP 55-534, 09340
México D. F., México
awaw@xanum.uam.mx

References

[1] C. Agrafeuil and M. Zarrabi, *Closed ideals with countable hull in algebras of analytic functions smooth up to the boundary*, Publ. Mat. **52** (2008), 19–56.

[2] J. E. Galé and P. J. Miana, *H^∞ functional calculus and Mikhlin-type multiplier conditions*, Canad. J. Math. **60** (5) (2008), 1010–1027.

[3] J. E. Galé, P. J. Miana, and J. J. Royo, *Nyman type theorem in convolution Sobolev algebras*, Rev. Mat. Complutense **25** (1) (2012), 1–19, available online, doi: 10.1007/s13163-010-0051-6.

[4] J. E. Galé, P. J. Miana, and J. J. Royo, *Estimates of the Laplace transform on convolution Sobolev algebras*, J. Approx. Theory **164** (2012), 162–178.

[5] J. E. Galé and A. Wawrzyńczyk, *Standard ideals in weighted algebras of Korenblyum and Wiener types*, Math. Scand. **108** (2) (2011), 291–319.

[6] K. Hoffman, Banach Spaces of Analytic Functions, Dover Publications, Inc., New York 1962.

[7] A. Sołtysiak and A. Wawrzyńczyk, *Ditkin's condition and ideals with at most countable hull in algebras of functions analytic in the unit disc*, Comment. Math. **52** (1) (2012), 101–112.

[8] W. Żelazko, Banach Algebras, Elsevier Publ. Co. and PWN, Amsterdam, Warszawa, 1973.

Bol. Soc. Mat. Mexicana (3) Vol. 19, 2013

RUIN PROBABILITIES AND THE RUIN TIME DISTRIBUTION

CARLOS G. PACHECO GONZÁLEZ

ABSTRACT. We analyze the renewal properties of a discounted risk process to derive integral equations that help to characterize quantities such as: ruin probability, survival probability, first passage time and time of ruin.

1. Introduction

Let $\delta > 0$ be a continuous time interest rate. We consider the following discounted risk process for an insurance company,

$$U_t := u + r(t) - Z_t^{(\delta)}, t \geq 0. \tag{1.1}$$

Here,

$$r(t) := \int_0^t \rho e^{-\delta s} ds$$

is the present value of the incomes received by the company up to time t, which is determined with the premium rate $\rho > 0$. Variable u is the initial capital of the company and

$$Z_t^{(\delta)} := \sum_{i=1}^{N_t} X_i e^{-\delta T_i},\ t \geq 0,$$

where N_t is a renewal process with interarrival times $\tau_1, \tau_2, \ldots$ which are independent identically distributed positive random variables (i.i.d. positive r.v.s.). Process N is defined by

$$N_t := \max\left\{k : \sum_{i=1}^{k} \tau_i \leq t\right\},$$

and $T_i := \sum_{j=1}^{i} \tau_j, i = 1, 2, \ldots$ represent the arrival times.

Variables $X_1, X_2, \ldots$ (named the claim size) are i.i.d. positive r.v.s. Throughout this paper we assume that the interarrival time and claim size are independent, and we denote by τ and X the generic random variables, such that $\tau \stackrel{d}{=} \tau_1$ and $X \stackrel{d}{=} X_1$, with distributions F_τ and F_X, respectively. To avoid technical problems we assume that $P(\tau > 0) = 1$ and that $P(X < \infty) = 1$.

Given the initial capital u, the *ruin probability* is given as

$$\psi(u) := P\left(\chi_u < \infty\right) \text{ where } \chi_u := inf\{s : U_s < 0\}.$$

Together with the ruin probability one is also interested on the *time of ruin*, that is

$$\inf\{s \geq 0 : U_s \leq 0\},$$

2010 *Mathematics Subject Classification*: 60K05, 65R20.

Keywords and phrases: discounted risk process, ruin and survival probabilities, ruin time distribution, Volterra integral equations.

This research was partially supported by a CONACYT program (México).

which is related to the first passage time of process $Z^{(\delta)}$, i.e.

$$\inf\{s \geq 0 : Z_s^{(\delta)} \geq h(s)\},$$

for some function h.

Our aim is to study $\psi(u)$ (or equivalently the survival probability $\phi(u) := 1 - \psi(u)$) and also the distribution of the time of ruin. In turn, we derive integral equations that help to characterize these quantities.

It is known that one can find equations to characterize the ruin or the survival probability, see [1] or [10], for instance. In fact, when the claims X and τ are exponentially distributed an explicit expression of $\psi(u)$ is given in Harrison [7]. There are also papers (see [12, 4, 5]) where, using stochastic calculus, it is derived an integro-differential equation for the survival probability, even when the discounted factor is random given by $\delta + \sigma B(t)$ where B is a Brownian motion. In this paper we give more elementary arguments to find such integro-differential equation for model (1.1), where δ or ρ are not perturbed. We do so using methods as for the classical Cramér-Lundberg model, see Grandell [6] for instance. The reader should notice that despite it has not been in general possible to solved explicitly these equations, one may use numerical procedures to approximate solutions.

The paper is organized in the following way. Next section presents integral equations in the general case when the interarrival time τ is not necessarily exponential. In Section 3 we derive the integro-differential equation and the Volterra integral equation for the survival probability $\phi(u)$ when τ is exponential r.v.; we also give some consequences of these equations. Finally, in last section we carry out some analysis to derive equations for the first passage time and the time of ruin distributions.

2. Basic equation for the ruin probability

It is well known that integral equations in renewal theory comes from the renewal properties of the processes; in the following result we use a renewal argument (similarly as in [8] for studying perpetuities).

LEMMA (2.1). *i) The ruin probability satisfies*

$$\begin{aligned}\psi(u) &= P\left(u + \frac{\rho}{\delta} - e^{-\delta\tau_1}(\frac{\rho}{\delta} + X_1) < 0\right) \\ &+ \int_0^\infty \int_0^{e^{\delta t}(u+\frac{\rho}{\delta})-\frac{\rho}{\delta}} \psi\left(e^{\delta t}(u + \frac{\rho}{\delta}) - (\frac{\rho}{\delta} + x)\right) F_X(dx) F_\tau(dt),\end{aligned}$$

where F_X and F_τ are the distributions of X and τ, respectively.

ii) The survival probability $\phi(u)$ satisfies

$$\phi(u) = \int_0^\infty \int_0^{e^{\delta t}(u+\rho/\delta)-\rho/\delta} \phi\left(e^{\delta t}(u + \frac{\rho}{\delta}) - (\frac{\rho}{\delta} + x)\right) F_X(dx) F_\tau(dt). \tag{2.2}$$

Proof. i) Notice that

$$\psi(u) = P\left(\chi_u < \infty | A\right) P(A) + P\left(\chi_u < \infty | A^c\right) P(A^c),$$

where $A = \left\{u + \frac{\rho}{\delta} - e^{-\delta\tau_1}(\frac{\rho}{\delta} + X_1) < 0\right\}$. Of course $P\left(\chi_u < \infty | A\right) = 1$.

Now, to calculate $P\left(\{\chi_u < \infty\}\bigcap A^c\right)$ consider that on A^c and when $t \geq \tau_1$

$$\begin{aligned}
U_t &= u + \frac{\rho}{\delta} - \frac{\rho}{\delta}e^{-\delta t} - e^{\delta\tau_1}\sum_{i=1}^{N_t} X_i e^{-\delta(T_i-\tau_1)} \\
&= u + \frac{\rho}{\delta} - \frac{\rho}{\delta}e^{-\delta\tau_1}e^{-\delta(t-\tau_1)} - X_1 e^{-\delta\tau_1} - e^{-\delta\tau_1}\sum_{i=2}^{N_t} X_i e^{-\delta(T_i-\tau_1)} \\
&= u + \frac{\rho}{\delta} - X_1 e^{-\delta\tau_1} + e^{-\delta\tau_1}\left(-\frac{\rho}{\delta}e^{-\delta(t-\tau_1)} - \sum_{i=2}^{N_t} X_i e^{-\delta(T_i-\tau_1)}\right) \\
&= e^{-\delta\tau_1}\left(\underbrace{e^{\delta\tau_1}(u+\frac{\rho}{\delta}) - (\frac{\rho}{\delta} + X_1)}_{\geq 0} + \underbrace{\frac{\rho}{\delta} - \frac{\rho}{\delta}e^{-\delta(t-\tau_1)}}_{=r(t-\tau_1)} - \underbrace{\sum_{i=2}^{N_t} X_i e^{-\delta(T_i-\tau_1)}}_{\stackrel{d}{=}\sum_{i=1}^{N_{t-\tau_1}}}\right).
\end{aligned}$$

Hence

$$P\left(\{\chi_u < \infty\}\bigcap A^c\right) = \int_0^\infty \int_0^{e^{\delta t}(u+\frac{\rho}{\delta})-\frac{\rho}{\delta}} \psi\left(e^{\delta t}(u+\frac{\rho}{\delta}) - (\frac{\rho}{\delta}+x)\right) F_X(dx) F_\tau(dt).$$

ii) The proof follows similar reasoning, or one might substitute $\psi(u) = 1 - \phi(u)$ in i). □

3. Equations for exponential interarrivals

In this section, assuming that τ is exponential, we derive an integro-differential equation for the survival probability.

(3.1) An integro-differential equation.

THEOREM (3.1). *If $\tau \sim \exp(\lambda)$, then ϕ is differentiable and it satisfies*

$$\phi'(u) = \frac{\lambda}{\delta(u+\rho/\delta)}\phi(u) - \frac{\lambda}{\delta(u+\rho/\delta)}\int_0^u \phi(u-x) F_X(dx). \tag{3.2}$$

Proof. Under the assumption, equation (2.2) of Lemma 2.1 is

$$\phi(u) = \int_0^\infty \int_0^{e^{\delta t}(u+\rho/\delta)-\rho/\delta} \phi\left(e^{\delta t}(u+\frac{\rho}{\delta}) - (\frac{\rho}{\delta}+x)\right) F_X(dx) \lambda e^{-\lambda t} dt.$$

Taking the change of variable $z = e^{\delta t}(u+\rho/\delta) - \rho/\delta$, we have that

$$e^{-\lambda t} = \left(\frac{u+\rho/\delta}{z+\rho/\delta}\right)^{\lambda/\delta} \text{ and } dt = \frac{dz}{\delta(z+\rho/\delta)}.$$

This turns equation (2.2) into

$$\begin{aligned}
\phi(u) &= \int_u^\infty \int_0^z \phi(z-x) F_X(dx) \frac{\lambda(u+\rho/\delta)^{\lambda/\delta}}{\delta(z+\rho/\delta)^{1+\lambda/\delta}} dz \\
&= \lambda(u+\rho/\delta)^{\lambda/\delta} \int_u^\infty \int_0^z \phi(z-x) F_X(dx) \frac{dz}{\delta(z+\rho/\delta)^{1+\lambda/\delta}},
\end{aligned}$$

which shows that ϕ is differentiable. Since we assume that $P(X < \infty) = 1$ and we take $\delta > 0$, then $\phi(u)$ is an increasing function of u and $\phi(u) < 1$ for all $u < \infty$. Thus, taking the derivative $\frac{d}{du}$ in previous display gives the result. □

COROLLARY (3.3). *Function ϕ satisfies*

$$\alpha = g(u)\phi(u) + \int_0^u K(u,s)\phi(s)ds, \ u \geq 0, \tag{3.4}$$

where $\alpha := -\rho\phi(0)$, $g(u) := -\rho - \delta u$ and $K(u,s) := \lambda + \delta - \lambda F_X(u-s)$.

Proof. We write equation (3.2) as

$$(\delta u + \rho)\phi'(u) = \lambda\phi(u) - \lambda \int_0^u \phi(u-x)F_X(dx).$$

Integrating on $[0,u]$ we obtain

$$\delta\left(u\phi(u) - \int_0^u \phi(s)ds\right) + \rho(\phi(u) - \phi(0)) = \lambda \int_0^u \phi(s)ds - \lambda \int_0^u \phi(s)F_X(u-s)ds.$$

For the last integral we used the indentity

$$\int_0^u \int_0^s \phi(s-x)F_X(dx)ds = \int_0^u \phi(s)F_X(u-s)ds.$$

(see for instance [11], page 194). This gives rise to the Volterra integral equation (3.4). □

REMARK (3.5). *Equation (3.4) can be written in the form*

$$\phi(u) = \frac{\rho\phi(0)}{\rho + \delta u} + \int_0^u \frac{(\lambda+\delta) - \lambda F_X(u-s)}{\rho + \delta u}\phi(s)ds.$$

Appealing to Theorem 5 in [2], p. 183, we know that previous Volterra linear equation admits a unique solution in the space of continuous function with compact support.

REMARK (3.6). *Equation (3.4) has been previously derived in [1] (see Proposition 1.8 of chapter VIII) using different methods. When $\delta = 0$, previous equation becomes the well known integral equation for non-discounted risk process (consult for example [6]).*

REMARK (3.7). *In Harrison [7] it is found explicitly the ruin probability $\psi(u)$ when $X \sim \exp(1/m)$; such solution is*

$$\psi(u) = c\int_{\rho/\delta+u}^{\infty} x^{\lambda/\delta - 1} e^{-x/m} dx,$$

where

$$c := \frac{m(\lambda/\delta)}{\int_{\rho/\delta}^{\infty} x^{\lambda/\delta} e^{-x/m} dx}.$$

Thus, under this circumstance, $1-\psi(u)$ solves equation (3.4).

Although not very explicit, the theory of successive approximations for integral equations (see e.g. section 1.3 of Corduneanu [3]) gives the following from Corollary (3.3).

PROPOSITION (3.8). *The unique solution of equation (3.4) is given by*

$$\phi(u) = \frac{\rho\phi(0)}{\rho + \delta u} + \rho\phi(0)\int_0^u \frac{R(u,s)}{\rho + \delta s} ds, \tag{3.9}$$

where $R(t,s)$ solves equation

$$R(t,s) = \tilde{K}(t,s) + \int_s^t R(t,r)\tilde{K}(r,s)dr, \ 0 \leq s \leq t, \tag{3.10}$$

with

$$\tilde{K}(t,s) := \frac{\lambda + \delta - \lambda F_X(t-s)}{\rho + \delta t}$$

(3.2) The Laplace transform of ϕ. Now we study the Laplace transform of $\{\phi(u), u \geq 0\}$, solution of (3.4). For this purpose, the Laplace transform of a function $\{h(u), u \geq 0\}$ evaluated at $z \in \mathbb{R}$ is denoted by

$$L[h(u)](z) := \int_0^\infty e^{-zu} h(u) du,$$

or simply $L[h](z)$. Let us then set $\Phi(z) := L[\phi](z)$.

From Widder [13] (see pages 446, 453 and 454, respectively) we have the following identities:

$$\begin{aligned} L[u\phi(u)](z) &= -\Phi'(z), \\ zL\left[\int_0^u \phi(s)ds\right](z) &= \Phi(z), \\ L\left[\int_0^u \phi(s)F_X(u-s)ds\right](z) &= \Phi(z)L[F_X](z). \end{aligned}$$

Thus, the Laplace transform of (3.4) yields

$$\frac{\alpha}{z} = -\rho\Phi(z) + \delta\Phi'(z) + (\lambda+\delta)\frac{\Phi(z)}{z} - \lambda\Phi(z)G(z),$$

where $G(z) := L[F_X](z)$.

This is an ordinary differential equation (ODE) of the form

$$\Phi'(z) = F(z)\Phi(z) + \frac{\alpha}{\delta z},$$

with

$$F(z) := \frac{1}{\delta}\left(\rho + G(z) - \frac{\lambda+\delta}{z}\right).$$

Hence, if $z_0 > 0$ we can calculate

$$\int_0^\infty \exp(-z_0 u)\phi(u) du,$$

and we can use theory of ODE to have the following expression for the Laplace transform.

PROPOSITION (3.11). *The Laplace transform of $\phi(u)$ is given by*

$$\Phi(z) = e^{\int_{z_0}^z F(r)dr}\left[\int_{z_0}^z \frac{\alpha}{\delta v} e^{-\int_{z_0}^v F(r)dr} dv + c\right],$$

where $c := \int_0^\infty \exp(-z_0 u)\phi(u)du$.

To have a more explicit expression for $\int_{z_0}^v F(r)dr$ above, we can use the Theorem of Fubini to write

$$\int_{z_0}^v F(r)dr = \frac{(v-z_0)\rho}{\delta} - \frac{(\lambda+\delta)}{\delta}\ln\left(\frac{v}{z_0}\right) + \frac{1}{\delta}\int_0^\infty \left(\frac{e^{-z_0 u} - e^{-vu}}{u}\right) F_X(u) du.$$

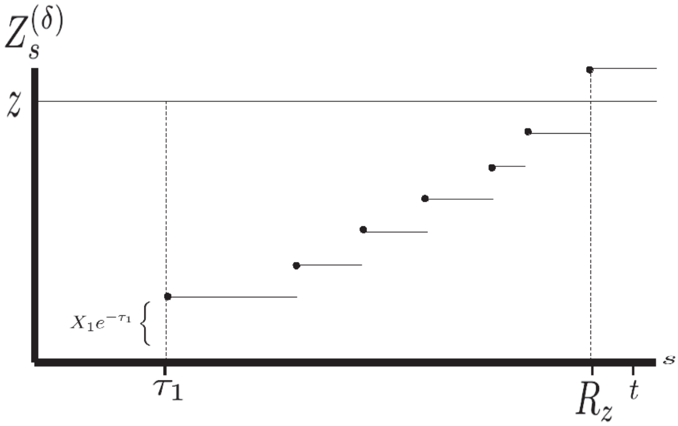

Figure 1. Draw of the first passage time.

4. Ruin time distribution

In this section, we want to exploit the renewal properties of the risk process U_t to derive a two-dimensional Volterra integral equation to describe the distribution of the time of ruin; the basic idea is taken from [9] where it was obtained equations to describe $P(Z_t^{(\delta)} \leq z)$. One may find useful to look at Figure 1.

Let us first analyze the first passage time of process $Z_t^{(\delta)}$. Define

$$R_z := \inf\{s \geq 0 : Z_s^{(\delta)} > z\},$$

and let $G_1(z,t) := P(R_z \leq t)$. Then we have

PROPOSITION (4.1). *The following integral equation holds for $G_1(z,t)$:*

$$G_1(t,z) = P(\tau \leq t, Xe^{-\delta\tau} \geq z) + \int_0^t \int_0^{ze^{\delta s}} G_1\left(ze^{\delta s} - x, t-s\right) F_X(dx) F_\tau(ds)$$

Proof. If $\tau_1 > t$, *a fortiori* $R_z > t$, for $z > 0$. Thus

$$G_1(z,t) = P(R_z \leq t | E_1)P(E_1) + P(R_z \leq t | E_2)P(E_2),$$

where

$$E_1 := \{\tau_1 \leq t,\ X_1 e^{-\delta\tau_1} \geq z\}$$

and

$$E_2 := \{\tau_1 \leq t,\ X_1 e^{-\delta\tau_1} < z\}.$$

Notice that if $\tau_1 \leq t$ and $X_1 e^{-\delta\tau_1} \geq z$, then $R_z \leq t$ always, i.e.

$$P(R_z \leq t | E_1) = 1.$$

Let us now concentrate on $P(R_z \leq t, E_2)$. From the equality of events

$$\{R_z \leq t\} = \{Z_t^{(\delta)} > z\}$$

we draw the following,

$$P(R_z \leq t, E_2) = P(Z_t^{(\delta)} > z, E_2)$$

$$= P\left(X_1 e^{-\delta\tau_1} + X_2 e^{-\delta(\tau_1+\tau_2)} + X_3 e^{-\delta(\tau_1+\tau_2+\tau_3)} + \ldots + X_{N_t} e^{-\delta(\tau_1+\ldots+\tau_{N_t})} > z, E_2\right)$$

$$= P\left(X_2 e^{-\delta\tau_2} + X_3 e^{-\delta(\tau_2+\tau_3)} + \ldots + X_{N_t} e^{-\delta(\tau_2+\ldots+\tau_{N_t})} > z e^{\delta\tau_1} - X_1, E_2\right),$$

which, due to the set E_2, becomes

$$= P\left(Z^{(\delta)}_{t-\tau_1} > z e^{\delta\tau_1} - X_1, E_2\right) = P\left(R_{z e^{\delta\tau_1} - X_1} \le t - \tau_1, E_2\right)$$

Thus, we have

$$\begin{aligned} P(R_z \le t, E_2) &= E\left(P\left(R_{z e^{\delta\tau_1} - X_1} \le t - \tau_1\right) I_{E_2}\right) \\ &= \int_0^t \int_0^{z e^{\delta s}} G_1\left(z e^{\delta s} - x, t - s\right) F_X(dx) F_\tau(ds). \end{aligned}$$

We can now gather terms to write down the equation. □

To study the time of ruin, we consider the following expression of U_t:

$$U_t = v - \frac{\rho}{\delta} e^{-\delta t} - Z^{(\delta)}_t,$$

where $v := u + \rho/\delta$. The ruin time is defined by

$$R(v) := \inf\{s \ge 0 : U_s \le 0\} = \inf\left\{s \ge 0 : Z_s \ge v - \frac{\rho}{\delta} e^{\delta s}\right\},$$

and its distribution is denoted by $G(v,t) := P(R(v) \le t)$. We can now characterize function G. In the proof, it might be useful to look at Figure 2.

THEOREM (4.2). *G is the unique solution of*

$$G(v,t) = P\left(\tau \le t,\ X \ge v e^{\delta\tau} - \rho/\delta\right) + \int_0^t \int_0^{v e^{\delta t} - \rho/\delta} G\left(v e^{\delta s} - x, t - s\right) F_X(dx) F_\tau(ds)$$

Proof. Notice that $P(R(v) \le t | \tau_1 > t) = 0$, then

$$G(v,t) = P(R(v) \le t | E_1) P(E_1) + P(R(v) \le t | E_2) P(E_2),$$

where

$$E_1 := \left\{X_1 e^{-\delta\tau_1} \ge v - \frac{\rho}{\delta} e^{-\delta\tau_1},\ \tau_1 \le t\right\}$$

and

$$E_2 := \left\{X_1 e^{-\delta\tau_1} < v - \frac{\rho}{\delta} e^{-\delta\tau_1},\ \tau_1 \le t\right\}.$$

However $P(R(v) \le t | E_1) = 1$. For the next term, we use now the renewal properties of process $Z^{(\delta)}_t$ to yield

$$\begin{aligned} & P(R(v) \le t,\ E_2) \\ = \ & P\left(\inf\left\{s \ge 0 : \sum_{i=2}^{N_s} X_i e^{-\delta T_i} \ge v - \frac{\rho}{\delta} e^{-\delta(\tau_1+s)} - X_1 e^{-\delta\tau_1}\right\} \le t - \tau_1,\ E_2\right) \\ & \text{(here multiplying both sides by } e^{\delta\tau_1}\text{)} \\ = \ & P\left(\inf\left\{s \ge 0 : \sum_{i=2}^{N_s} X_i e^{-\delta\sum_{j=2}^{i}\tau_j} \ge v e^{\delta\tau_1} - X_1 - \frac{\rho}{\delta} e^{-\delta s}\right\} \le t - \tau_1,\ E_2\right) \\ = \ & \int_0^t \int_0^{v e^{\delta t} - \rho/\delta} G\left(v e^{\delta s} - x, t - s\right) F_X(dx) F_\tau(ds); \end{aligned}$$

for the last equality we used the fact that

$$\left\{\sum_{i=2}^{N_s} X_i e^{-\delta\sum_{j=2}^{i}\tau_j},\ s \ge 0\right\} \stackrel{(d)}{=} \{Z^{(\delta)}_s,\ s \ge 0\}.$$

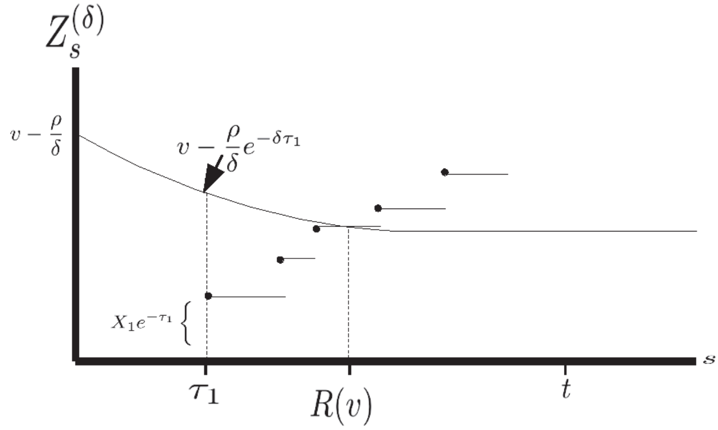

Figure 2. Draw of the time of ruin.

Joining pieces we end up with the relation. □

Acknowledgements

We would like to thanks the anonymous referee, who gave several useful corrections to improve the paper.

Received September 30, 2012

Final version received March 5, 2013

Carlos G. Pacheco González
Departamento de Matemáticas, CINVESTAV-IPN,
A. Postal 14-740, México D.F.
07000, MÉXICO
cpacheco@math.cinvestav.mx

References

[1] S. Asmussen and H. Albrecher, Ruin Probabilities (2nd edition), World Scientific, 2010.
[2] W. Cheney, Analysis for Applied Mathematics, Springer, 2001.
[3] C. Corduneanu, Integral Equations and Applications, Cambridge University Press, 1991.
[4] J. Gaier, *A Karamata-type and ruin probabilities for an insurer investing proportionally in the stock market*, Insurance Math. Economics **34** (2004), 297–305.
[5] J. Gaier and P. Grandits, *Ruin probabilities and investment under interest force in the presence of regular varying tails*, Scand. Actuarial J. **4** (2004), 256–278.
[6] J. Grandell, Aspects of Risk Theory, Springer, 1991.
[7] J. M. Harrison, *Ruin problems with compounding assets*, Stochastic Process. Appl. **5** (1977), 67–79.
[8] G. Hernandez-del-Valle and C. G. Pacheco-González, *Embedded Markov chains in discounted processes and the present value distribution*, Morfismos **14 (2)** (2010).
[9] C. G. Pacheco-González, *Integral equations for the aggregate claim amount*, Revista de Matemática: Teoría y Aplicaciones **18 (1)** (2011), 1–7.
[10] J. Paulsen, *Ruin models with investment income*, Probab. Surv. **5** (2008), 416–434.
[11] Š. Schwabik, M. Tvrdý, and O. Vejvoda, Differential and Integral Equations, D. Reidel Publishing Company, 1979.

[12] G. WANG AND R. WU, *Distribution for the risk process with a stochastic return on investments*, Stochastic Process. Appl. **95** (2001), 329–341.
[13] D. V. WIDDER, Advanced Calculus (2nd edition), Dover, 1989.

Bol. Soc. Mat. Mexicana (3) Vol. 19, 2013

OPTIMAL SOLUTIONS OF CONSTRAINED DISCOUNTED SEMI-MARKOV CONTROL PROBLEMS

JUAN GONZÁLEZ-HERNÁNDEZ AND CÉSAR EMILIO VILLARREAL-RODRÍGUEZ

ABSTRACT. We give conditions for the existence of optimal solutions to the constrained semi-Markov decision problem on Borel spaces, with possibly unbounded costs and discounted performance index. We also demonstrate the existence of optimal solutions which are given by a convex combinations of $N+1$ measurable selectors, where N is the number of constraints.

1. Introduction

There are not many articles that work with constrained semi-Markov decision processes in Borel spaces and performance index of expected discounted cost. Luque Vásquez and Robles-Alcaraz [23] worked on unbounded and unconstrained semi-Markov decision processes. Abdel-Hameed [1] gave applications to reliability. Love et al [20] analyzed the optimal repair/replacement policy in a machine system modeled as a discrete semi-Markov decision process. Cao [4] made a sensitivity analysis in a finite state semi-Markov process. Hu and Yue [14] showed the existence of optimal control limit policies of a semi-Markov decision process whose objective function is an expected discounted cost and the environment changes according to a semi-Markov process. Hudak and Nollau [17] considered an approximation procedure for calculating the optimal value of a discounted semi-Markov decision process with countable state space and finite action space. Hao, Hongsheng and Baoqun [10] proposed an optimal robust control policy for uncertain semi-Markov control processes. Dai et al [5] studied performance optimization algorithms for a class of semi-Markov control processes with compact action set. Tang, Xi and Yin [28] extend techniques based on performance potential for Markov control problems to uncertain semi-Markov control processes. Liu [18] proved the existence of ε-optimal in a weighted semi-Markov decision processes. Luque-Vásquez and Minjárez-Sosa [21] studied a discounted semi-Markov control process with Borel state space, unbounded cost function and unknown holding time distribution. Singh, Tadić and Doucet [27] implemented an algorithm using the gradient method for semi-Markov decision processes with application to call admission control. Hao, Baoqun and Hongsheng [9] established error bounds for potential in optimization algorithms for semi-Markov decision processes. Yin et al [32] studied the problems of discounted-cost performance optimization for a class of semi-Markov decision processes. Luque-Vásquez, Minjárez-Sosa and Rosas-Rosas [22] boarded the problem of semi-Markov control processes with unknown and partially unknown holding times distribution under average cost and

2010 *Mathematics Subject Classification*: 60K15, 47N10, 47N30.

Keywords and phrases: semi-Markov decision process, performance index, stochastic kernel, Polish space.

discounted cost criteria. Huang and Guo in [15] studied a first passage model for discounted semi-Markov decision processes with denumerable states and non-negative costs, and in [16] the finite horizont semi-Markov decision processes with denumerable states.

For a deeper study on Markov Decision Processes see [11] or [25]. The cases with constraints can be seen in [7, 6, 8].

So far there are no known works that address to the constrained semi-Markov control problem whose state space and control space are Borel spaces. The main contributions of this article is the focus by occupation measures.

Section 2 contains preliminary notions of semi-Markov control model with constraints. Section 3 works with a generalization of the occupation measures for semi-Markov decision processes with discounted cost. This allows to give an equivalent problem in terms of a family of measures and to use the direct approach. In Section 4 we prove the existence of solutions to the control problem. We conclude the article with Section 5, where the existence of stationary optimal solutions which are convex combinations of at most $N+1$ measurable selectors is demonstrated.

Notation (1.1). A Polish space Z is a complete separable metric space and a Borel space is a measurable subset of a Polish space. We denote by $\mathcal{B}(Z)$ its Borel σ-algebra. Measurable always means Borel-measurable. If Z and W are Borel spaces, a stochastic kernel P on Z given W is a function $(w,B)\mapsto \mathrm{P}(B|w)$ such that $\mathrm{P}(\cdot|w)$ is a probability measure on $\mathcal{B}(Z)$ for each $w\in W$, and $\mathrm{P}(B|\cdot)$ is a measurable function on W for each $B\in\mathcal{B}(Z)$. We also shall denote by $\mathcal{P}(Z)$ the set of all probability measures on $(Z,\mathcal{B}(Z))$.

2. The semi-Markov control model

Definition (2.1). A *constrained semi-Markov decision process* (CSMDP)

$$(\mathbb{X},\mathbb{A},A,Q,F,c_j,d_j;\ j\in\{0,1,\dots,N\})$$

consists of:

(a) A Borel space space $\mathbb{X}$, called the *state space*.
(b) A Borel space $\mathbb{A}$, the *control* (or action) space.
(c) A function $A:\mathbb{X}\to\{B:B \text{ is a measurable subset of } \mathbb{A}\}$. For each x we have that $A(x)\neq\varnothing$ and it is the set of *admissible controls* (or actions) at the state x. Moreover we assume that the set

$$\mathbb{K}:=\{(x,a):x\in\mathbb{X},\ a\in A(x)\}$$

is a Borel subset of $\mathbb{X}\times\mathbb{A}$ and contains the graph of a measurable map from $\mathbb{X}$ to $\mathbb{A}$.
(d) A stochastic kernel Q on $\mathbb{X}$ given $\mathbb{K}$ called the *transition law*.
(e) A continuous function $t\mapsto F(t|x,a,y)$ which is a probability distribution function, for each $(x,a,y)\in\mathbb{K}\times\mathbb{X}$, and we assume $F(t|\cdot)$ is jointly measurable for each real number t.
(f) The nonnegative measurable real functions on $\mathbb{X}\times\mathbb{A}$, c_j and d_j for $j\in\{0,1,\dots,N\}$ are the so called *cost functions*.

The CSMDP represents a stochastic system that evolves in the next way: At stage i the system is in the state $x_i \in \mathbb{X}$ and a control $a_i \in A(x_i)$ is applied, then the following things happen: The immediate costs $c_j(x_i,a_i)$ for $j \in \{0,1,\ldots,N\}$ are incurred. The system moves to the next state $x_{i+1} \in \mathbb{X}$ according to the probability measure $Q(\cdot|x_i,a_i)$. Conditional to (x_i,a_i,x_{i+1}) the time t_{i+1} from the transition i occurs until the transition $i+1$ occurs has the distribution function $F(\cdot|x_i,a_i,x_{i+1})$. The costs $d_j(x_i,a_i)$ for $j \in \{0,1,\ldots,N\}$ are imposed until the transition $i+1$ occurs. After the transition $i+1$ occurs, a control $a_{i+1} \in A(x_{i+1})$ is chosen and the process continues in this way.

For each $i \in \mathbb{N}\cup\{0\}$, define the space of admissible histories up to stage i by $H_0 := \mathbb{X}$ and $H_i := \mathbb{K}^{i-1} \times \mathbb{X} = \mathbb{K} \times H_{i-1}$. A generic element $h_i \in H_i$ is a vector, or history, of the form $h_i = (x_0,a_0,\ldots,x_{i-1},a_{i-1},x_i)$, where $(x_j,a_j) \in \mathbb{K}$ for $j \in \{0,1,2,\ldots,i-1\}$ and $x_i \in \mathbb{X}$.

Definition (2.2). (a) A *control policy* is a sequence $\pi = (\pi_i)$ of stochastic kernels π_i on $\mathbb{A}$ given H_i, satisfying the constraint $\pi_i(A(x_i)|h_i) = 1$ for all $h_i \in H_i$ and $i \in \mathbb{N}\cup\{0\}$. We denote by Π the class of all policies.

(b) A control policy is said to be *randomized stationary*, if there exists a stochastic kernel φ on $\mathbb{A}$ given $\mathbb{X}$, satisfying the constraint $\varphi(A(x)|x) = 1$ such that $\pi_i(\cdot|h_i) = \varphi(\cdot|x_i)$ for all $h_i \in H_i$ and $i \in \mathbb{N}\cup\{0\}$. We identify the policy π with φ and denote by Φ the set of all such policies.

(c) A randomized stationary policy φ is said to be *stationary deterministic* if there exists a function f from $\mathbb{X}$ to $\mathbb{A}$, satisfying the constraint $f(x) \in A(x)$ such that $\varphi(\cdot|x)$ is concentrated at $f(x)$. These functions are called measurable selectors. We identify the policy φ with f and we denote by $\mathbb{F}$ the set of all such policies.

(d) A randomized stationary policy is said to be *N-randomized policy* if it is a convex combination of at most N stationary deteministic policies. We denote the set of all such policies by Φ_N.

Given an initial distribution $\nu \in \mathcal{P}(\mathbb{X})$ and $\pi = (\pi_i) \in \Pi$, by Ionescu-Tulcea Theorem ([3, Th. 2.7.2], [13, Sec. 11] or [19, pp. 137-139]), there exists a probability space $(\Omega,\mathcal{A},\mathrm{P})$ such that

1. $\mathrm{P}^\pi_\nu(x_0 \in B) = \nu(B)$ for $B \in \mathcal{B}(\mathbb{X})$;
2. $\mathrm{P}^\pi_\nu(x_{i+1} \in B|h_i,a_i) = Q(B|x_i,a_i)$ for all $B \in \mathcal{B}(\mathbb{X})$, $h_i \in H_i$ and $a_i \in A(x_i)$, $i \in \mathbb{N}\cup\{0\}$;
3. $\mathrm{P}^\pi_\nu(a_i \in C|h_i) = \pi_i(C|h_i)$ for all $C \in \mathcal{B}(\mathbb{A})$ and $h_i \in H_i$, $i \in \mathbb{N}\cup\{0\}$;
4. $\mathrm{P}^\pi_\nu(t_i \le r|h_{i+1}) = F(r|x_i,a_i,x_{i+1})$ for all $r \in \mathbb{R}$, $h_{i+1} \in H_{i+1}$ and $a_i \in A(x_i)$, $i \in \mathbb{N}\cup\{0\}$;

From the dynamic of the process given by Definitions (2.1) (d) and (e) and the nature of a policy, Definition (2.2) (a), we can see that:

REMARK (2.3). *The random variables $t_1,t_2,\ldots$ are conditionally independent given the process $(x_0,a_0,x_1,a_1,\ldots)$.*

We denote by E^π_ν the expectation with respect to P^π_ν , and we denote by P^π_x and E^π_x respectively when ν is concentrated at x.

In order to assure that an infinite number of transitions $t_1,t_2,\ldots$ does not occur in a finite interval, we need to impose a condition: (see Vega-Amaya [31]). To do this, we introduce the following notations: for the distibution function of the holding time t_i conditional to $(x,a)\in\mathbb{K}$ we put

$$G(t|x,a):=\int_{\mathbb{X}} F(t|x,a,y)Q(\mathrm{d}y|x,a) \quad \text{for } t\geq 0; \tag{2.4}$$

for the conditional expected value of t_i given (x,a) we put

$$\begin{aligned} \tau(x,a) &:= \int_0^{+\infty} (1-G(t|x,a))\mathrm{d}t \\ &= \int_{\mathbb{X}} \int_0^{+\infty} (1-F(t|x,a,y))\mathrm{d}t Q(\mathrm{d}y|x,a)); \end{aligned} \tag{2.5}$$

and we introduce the auxiliary functions τ_α such that $\tau_\alpha(x,a)$ is the conditional expected value of $\int_0^{t_i} \mathrm{e}^{-\alpha t}\mathrm{d}s$ given (x,a) puting

$$\begin{aligned} \chi_\alpha(x,a) &:= \int_0^{+\infty} \mathrm{e}^{-\alpha t} G(\mathrm{d}t|x,a) \text{ and} \\ \tau_\alpha(x,a) &:= (1-\chi_\alpha(x,a))/\alpha, \end{aligned} \tag{2.6}$$

for $\alpha\in(0,1)$ (the discount rate) and $(x,a)\in\mathbb{K}$.

Condition (2.7). There exist $\varepsilon>0$ and $\bar{t}>0$ such that $\mathrm{P}(t>\bar{t}|x,a) = 1-G(\bar{t}|x,a)\geq \varepsilon$ for all $(x,a)\in\mathbb{K}$.

So we have the next lemma (Proposition 2.4 of [31]):

LEMMA (2.8). *If Condition (2.7) holds, then*

1. $\inf_{\mathbb{K}} \tau(x,a)\geq \varepsilon\bar{t}$;
2. $\bar{\tau}_\alpha := \sup_{\mathbb{K}} \tau_\alpha(x,a) < 1$;
3. $\mathrm{P}_x^\pi\left(\sum_{i=1}^{\infty} t_i = +\infty\right) = 1$ *for every* $x\in\mathbb{X}$ *and* $\pi\in\Pi$.

Performance index. Let us define the sum of the transition times as $T_0 := t_0 := 0$, and $T_i := T_{i-1}+t_i$ for $i\in\mathbb{N}$. If we consider $V_j^\alpha(\pi,\nu)$ for $\pi\in\Pi$ and ν a fixed initial distribution on $\mathbb{X}$ as:

$$V_j^\alpha(\pi,\nu) := \mathrm{E}_\nu^\pi\left(\sum_{i=1}^{\infty} \mathrm{e}^{-\alpha T_i}\left(c_j(x,a)+d_j(x,a)\int_0^{t_{i+1}} \mathrm{e}^{-\alpha t}\mathrm{d}t\right)\right),$$

for $j\in\{0,1,\ldots,N\}$. Now we define new current costs when the process is in state x and an action a is chosen as:

$$C_j^\alpha(x,a) := c_j(x,a)+d_j(x,a)\tau_\alpha(x,a), \tag{2.9}$$

where τ_α is given in (2.6). So, by properties of conditional probability, we can express $V_j^\alpha(\pi,\nu)$ as:

$$V_j^\alpha(\pi,\nu) := \mathrm{E}_\nu^\pi\left(\sum_{i=0}^{\infty} \mathrm{e}^{-\alpha T_i} C_j^\alpha(x_i,a_i)\right), \tag{2.10}$$

for $j\in\{0,1,\ldots,N\}$.

Definition (2.11). For $\alpha \in (0,1)$, an initial distribution ν on $\mathbb{X}$ and a policy $\pi \in \Pi$, the values of $V_j^\alpha(\pi,\nu)$ given in (2.10) are called the *α-discounted expected costs*. When ν is concentrated at some $x \in \mathbb{X}$ we write $V_j^\alpha(\pi,x)$ instead.

Let $k_j > 0$ for $j \in \{1,\ldots,N\}$ given. The discounted control problems is:

$$\begin{array}{ll} \textbf{DCP} & \min V_0^\alpha(\pi,\nu), \\ \text{subject to} & V_j^\alpha(\pi,\nu) \le k_j \ \text{ for } \ j \in \{1,\ldots,N\}, \\ & \pi \in \Pi. \end{array}$$

If a policy π^* reaches this minimum, it is said that it is an optimal solution for **DCP** and it is called an *alpha-discounted optimal policy*.

3. Discounted occupation measures

Notation (3.1). For an arbitrary Borel space Z we shall denote by $\mathcal{M}(Z)$ the family of finite (signed) measures on $(Z,\mathcal{B}(Z))$. We shall denote by $\mathcal{M}(Z|Z')$ the family of all conditional finite measures on Z given Z'. That is, an element of $\mathcal{M}(Z|Z')$ is a function $(z',B) \mapsto m(B|z')$, such that $m(B|\cdot)$ is a measurable function on Z' for each $B \in \mathcal{B}(Z)$ and $m(\cdot|z') \in \mathcal{M}(Z)$ for each $z' \in Z'$. Also when $C \subset Z$, $\mathcal{M}(C)$ shall denote the family of all finite measures with support on C. Similarly we denote by $\mathcal{M}_+(Z)$, $\mathcal{M}_+(Z|Z')$ and $\mathcal{M}_+(C)$ the corresponding families for nonnegative finite measures.

Let us define the conditional measure $H_\alpha \in \mathcal{M}(\mathbb{X}|\mathbb{X} \times \mathbb{A})$ as

$$\begin{aligned} H_\alpha(C|x,a) &:= \int_C \mathrm{e}^{-\alpha t} G(\mathrm{d}t|x,a) \\ &= \mathrm{E}_\pi^\nu \left(\mathrm{e}^{-\alpha t_i} \mathbb{1}_C(x_i) | x_{i-1} = x, a_{i-1} = a \right), \end{aligned}$$

for all $i \in \mathbb{N}$, where $\mathbb{1}_C$ is the indicator function of the set C, that is $\mathbb{1}_C(x) = 1$, if $x \in C$, and $\mathbb{1}_C(x) = 0$ otherwise.

Definition (3.2). Given $\alpha \in (0,1)$, ν an initial distribution on $\mathcal{P}(\mathbb{X})$ and a policy $\pi \in \Pi$, the *α -discounted occupation measure* $m(\cdot : \pi,\nu,\alpha)$ is defined as

$$m(B : \pi,\nu,\alpha) := \mathrm{E}_\nu^\pi \left(\sum_{i=0}^\infty \mathrm{e}^{-\alpha T_i} \mathbb{1}_B(x_i,a_i) \right), \tag{3.3}$$

for $B \in \mathcal{B}(\mathbb{X} \times \mathbb{A})$.

REMARK (3.4). *Observe that $m(\cdot : \pi,\nu,\alpha) \in \mathcal{M}_+(\mathbb{X} \times \mathbb{A})$. In fact*

$$m(B : \pi,\nu,\alpha) := \mathrm{E}_\nu^\pi \left(\sum_{i=0}^\infty \mathrm{e}^{-\alpha T_i} \mathbb{1}_B(x_i,a_i) \right) \le \sum_{i=0}^\infty (\bar{\tau}_\alpha)^i,$$

REMARK (3.5). (a) *If $\mu \in \mathcal{M}(\mathbb{X} \times \mathbb{A})$, there is a randomized control $\varphi \in \Phi$ and a signed measure $\hat{\mu} \in \mathcal{M}(\mathbb{X})$ such that*

$$\mu(B \times C) = (\hat{\mu} \otimes \varphi)(B \times C) := \int_B \varphi(C|x)\hat{\mu}(\mathrm{d}x) \tag{3.6}$$

for $B \in \mathcal{B}(\mathbb{X})$ and $C \in \mathcal{B}(\mathbb{A})$. The measure $\hat{\mu}$ is called the marginal measure of μ on $\mathbb{X}$ and it is obtained by mean of $\hat{\mu} := \mu(\cdot \times \mathbb{A})$. Observe that for each $C \in \mathcal{B}(\mathbb{A})$, $\hat{\mu}(B) = 0 \implies \mu(B \times C) = 0$ and the function $\varphi(C|\cdot)$ is the Radon-Nikodým derivative of $\mu(\cdot \times C)$ with respect to $\hat{\mu}$.

(b) *Conversely, if* φ *is a randomized control and* $\hat{\mu} \in \mathcal{M}(\mathbb{X})$*, there is an unique signed measure* $\mu \in \mathcal{M}(\mathbb{X} \times \mathbb{A})$ *such that* (3.6) *is satisfied.*

LEMMA (3.7). *Let* f *be a nonnegative measurable function on* $\mathcal{B}(\mathbb{X} \times \mathbb{A})$*, and let* $\alpha \in (0,1)$*,* ν *an initial distribution on* $\mathcal{P}(\mathbb{X})$ *and a policy* $\pi \in \Pi$*. Set* $m(B) = m(B : \pi, \nu, \alpha)$*, thus*

$$\int_{\mathbb{X}\times\mathbb{A}} f(x,a)m(\mathrm{d}(x,a)) = \mathrm{E}_\nu^\pi\left(\sum_{i=0}^{\infty} \mathrm{e}^{-\alpha T_i} f(x_i,a_i)\right).$$

Proof. This property can be proved by following the classic way of supposing first the case when the function f is an indicator function, then simple function, then increasing limit of simple functions. □

Notation (3.8). Given a measurable function f on $\mathbb{K}$, a stochastic kernel P on $\mathbb{X}$ given $\mathbb{K}$, a stochastic kernel φ on $\mathbb{A}$ given $\mathbb{X}$ and a Borel subset of $\mathbb{X}$, we denote by

$$f(x,\varphi) := \int_{\mathbb{A}} f(x,a)\varphi(\mathrm{d}a|x)$$

and

$$\mathrm{P}(B|x,\varphi) = \int_{\mathbb{A}} \mathrm{P}(B|x,a)\varphi(\mathrm{d}a|x).$$

THEOREM (3.9). *A measure* $m \in \mathcal{M}_+(\mathbb{K})$ *is an* α*-discounted occupation measure if and only if it satisfies*

$$\hat{m}(B) = \nu(B) + \int_{\mathbb{X}\times\mathbb{A}} H_\alpha(B|x,a)m(\mathrm{d}(x,a)), \tag{3.10}$$

for every Borel subset B *of* $\mathbb{X}$*, where* $\hat{m}$ *is the marginal measure of* m *on* $\mathbb{X}$*, that is* $\hat{m}(B) := m(B \times \mathbb{A})$ *for* $B \in \mathcal{B}(\mathbb{X})$*.*

Proof. Let us take $m(B) = m(B : \pi, \nu, \alpha)$. We have

$$\begin{aligned}
\hat{m}(B) = m(B \times \mathbb{A}) &= \mathrm{E}_\nu^\pi\left(\sum_{i=0}^{\infty} \mathrm{e}^{-\alpha T_i} \mathbb{1}_{B\times\mathbb{A}}(x_i,a_i)\right) \\
&= \nu(B) + \mathrm{E}_\nu^\pi\left(\sum_{i=1}^{\infty} \mathrm{e}^{-\alpha T_i} \mathbb{1}_{B\times\mathbb{A}}(x_i,a_i)\right) \\
&= \nu(B) + \mathrm{E}_\nu^\pi\left(\sum_{i=1}^{\infty} \mathrm{E}_\nu^\pi(\mathrm{e}^{-\alpha T_i} \mathbb{1}_{B\times\mathbb{A}}(x_i,a_i)|h_{i-1},a_{i-1})\right) \\
&= \nu(B) + \mathrm{E}_\nu^\pi\left(\sum_{i=1}^{\infty} \mathrm{e}^{-\alpha T_{i-1}} \mathrm{E}_\nu^\pi(\mathrm{e}^{-\alpha t_i} \mathbb{1}_{B\times\mathbb{A}}(x_i,a_i))|x_{i-1},a_{i-1})\right) \\
&= \nu(B) + \mathrm{E}_\nu^\pi\left(\sum_{i=1}^{\infty} \mathrm{e}^{-\alpha T_{i-1}} H_\alpha(B|x_{i-1},a_{i-1})\right) \\
&= \nu(B) + \int_{\mathbb{X}\times\mathbb{A}} H_\alpha(B|x,a)m(\mathrm{d}(x,a)).
\end{aligned}$$

Conversely, let us consider a measure $m \in \mathcal{M}_+(\mathbb{K})$ such that satisfies (3.10). Let us disintegrate $m = \hat{m} \otimes \varphi$, then iterations of this equation produce

$$\begin{aligned}
\hat{m}(B) &= \nu(B) + \int_{\mathbb{X}} H_\alpha(B|x_0,\varphi)\hat{m}(\mathrm{d}x_0) \\
&= \nu(B) + \int_{\mathbb{X}} H_\alpha(B|x_0,\varphi)\nu(\mathrm{d}x_0) \\
&+ \int_{\mathbb{X}}\int_{\mathbb{X}} H_\alpha(B|x_1,\varphi)H_\alpha(\mathrm{d}x_1|x_0,\varphi)\hat{m}(\mathrm{d}x_0) = \cdots = \nu(B) \\
&+ \sum_{i=1}^{M-1} \int_{\mathbb{X}}\int_{\mathbb{X}}\cdots\int_{\mathbb{X}} H_\alpha(B|x_i,\varphi)\prod_{k=1}^{i} H_\alpha(\mathrm{d}x_k|x_{k-1},\varphi)\nu(\mathrm{d}x_0) \\
&+ \int_{\mathbb{X}}\int_{\mathbb{X}}\cdots\int_{\mathbb{X}} H_\alpha(B|x_M,\varphi)\prod_{k=1}^{M} H_\alpha(\mathrm{d}x_k|x_{k-1},\varphi))\hat{m}(\mathrm{d}x_0),
\end{aligned}$$

for all $M \in \mathbb{N}$. In this last expression we consider an empty product as equal to 1. The last sumand tends to zero, in fact

$$\int_{\mathbb{X}}\int_{\mathbb{X}}\cdots\int_{\mathbb{X}} H_\alpha(B|x_M,\varphi)\prod_{k=1}^{M} H_\alpha(\mathrm{d}x_k|x_{k-1},\varphi))\hat{m}(\mathrm{d}x_0) \leq (\bar{\tau}_\alpha)^M \to 0,$$

as $M \to \infty$. Now, for $i \in \mathbb{N}$ we have

$$\begin{aligned}
&\int_{\mathbb{X}}\int_{\mathbb{X}}\cdots\int_{X} H_\alpha(B|x_i,\varphi)\prod_{k=1}^{i} H_\alpha(\mathrm{d}x_{k-1}|x_{k-1},\varphi)\nu(\mathrm{d}x_0) \\
&= \mathrm{E}_\nu^\varphi(\mathrm{e}^{-\alpha T_i}\mathbb{1}_B(x_i)),
\end{aligned}$$

and

$$\nu(B) = \mathrm{E}_\nu^\varphi\left(\mathrm{e}^{-\alpha T_0}\mathbb{1}_B(x_0)\right).$$

Hence

$$\hat{m}(B) = \sum_{i=0}^{\infty} \mathrm{E}_\nu^\varphi(\mathrm{e}^{-\alpha T_i}\mathbb{1}_B(x_i)) = m(B \times \mathbb{A} : \pi,\nu,\alpha) = \hat{m}(B:\pi,\nu,\alpha),$$

therefore

$$m = \hat{m}\otimes\varphi = \hat{m}\otimes\varphi(\cdot:\pi,\nu,\alpha) = m(\cdot:\pi,\nu,\alpha). \qquad \square$$

COROLLARY (3.11). *The family of stationary policies is sufficient for the control problems.*

Let us denote

$$\langle m, f\rangle := \int f(x,a)m(\mathrm{d}(x,a)).$$

Let $k_j \geq 0$ for $j \in \{0,1,\ldots,N\}$ given.
The control problems now has the form

$$\begin{array}{rl}
\textbf{MDCP} & \min\left\langle m, C_0^\alpha\right\rangle \\
\text{subject to} & \left\langle m, C_j^\alpha\right\rangle \leq k_j \ \text{ for } j \in \{1,2,\ldots,N\} \\
\text{and} & \hat{m} = \nu + (m \otimes H_\alpha) \\
 & m \in \mathcal{M}_+(\mathbb{K}).
\end{array}$$

Actually.

COROLLARY (3.12). *The problems* ***DCP*** *and* ***MDCP*** *are equivalent.*

4. Existence of solution

Remember that $\mathbb{X}$ and $\mathbb{A}$ are Borel spaces. The functions c_j and d_j are nonnegative for $j \in \{1,2,\dots,N\}$ and F is a continuous function. By using the former corollary we shall prove that **MDCP** is solvable in Theorem (4.4) below.

Condition (4.1). Let us suppose:

(a) There is a policy $\pi \in \Pi$ such that $V_j^{\alpha}(\pi,\nu) \le k_j$ for $j \in \{1,2,\dots,N\}$ and $V_0^{\alpha}(\pi,\nu) < +\infty$.
(b) The function c_0 is inf-compact or is a moment and lower semicontinuous.
(c) The functions c_k and d_j are lower semicontinuous functions for $k \in \{1,2,\dots,N\}$ $j \in \{0,1,\dots,N\}$.
(d) The stochastic kernel Q is weakly continuous.
(e) There is a density function $f(t|x,a,y)$ for $F(t|x,a,y)$ such that f is an uniformly continuous function on all its variables.

REMARK (4.2). *Claus (a) is in order the problem* ***DCP*** *(or* ***MDCP****) makes sense. Claus (b) assure we have a tight family, so we can apply Prohrov's Theorem. Claus (c) implies that with limits we fulfill the constraints. Clauses (d) and (e) force the function τ_α to be continuous.*

Let us define the set $\mathcal{M}_f$ of all feasible occupation measures, that is, the set of all measures $m \in \mathcal{M}_+(\mathbb{K})$ such that

$$\begin{aligned}
&\langle m, C_0^{\alpha}\rangle < +\infty,\\
&\langle m, C_j^{\alpha}\rangle \le k_j \quad \text{for} \quad j \in \{1,2,\dots,N\},\\
&\text{and} \quad \hat{m} = \nu + (m \otimes H_\alpha)
\end{aligned}$$

and let us define the value of the program **MDCP** by

$$\bar{V} = \inf\{\langle m, C_0^{\alpha}\rangle : m \in \mathcal{M}_f\}.$$

LEMMA (4.3). *Under Condition (4.1) the function τ_α given in (2.6) is a continuous function.*

Proof. As $f(t|x,a,y)$ is an uniformly continuous function on (t,y), given $\varepsilon > 0$ and $(x,a) \in \mathbb{X} \times \mathbb{A}$, there is $\delta > 0$, such that

$$|f(t|x,a,y) - f(t|x,a,y')| < \varepsilon,$$

for all $(t,y),(t,y') \in (0,+\infty) \times \mathbb{X}$ such that $d_1(y,y') < \delta$, where d_1 is the distance in the space $\mathbb{X}$. Hence

$$\left| \int_0^{+\infty} e^{-\alpha t} f(dt|x,a,y) - \int_0^{+\infty} e^{-\alpha t} f(dt|x,a,y') \right| = \frac{\varepsilon}{\alpha}.$$

That is, for each $(x_0,a_0) \in \mathbb{X} \times \mathbb{A}$, the function

$$y \mapsto \int_0^{+\infty} e^{-\alpha t} f(dt|x_0,a_0,y)$$

is continuous on $\mathbb{X}$. Hence, by Condition (4.1) (d), the function

$$(x,a) \mapsto \int_{\mathbb{X}} \int_0^{+\infty} \mathrm{e}^{-\alpha t} f(\mathrm{d}t|x_0,a_0,y) Q(\mathrm{d}y|(x,a))$$

is a continuous function on $\mathbb{X} \times \mathbb{A}$ for each $(x_0,a_0) \in \mathbb{X} \times \mathbb{A}$. Now from the inequality

$$\begin{aligned}
&\left| \int_{\mathbb{X}} \int_0^{+\infty} \mathrm{e}^{-\alpha t} f(\mathrm{d}t|x',a',y) Q(\mathrm{d}y|x',a') \right. \\
&\left. - \int_{\mathbb{X}} \int_0^{+\infty} \mathrm{e}^{-\alpha t} f(\mathrm{d}t|x,a,y) Q(\mathrm{d}y|x,a) \right| \\
&\le \left| \int_{\mathbb{X}} \int_0^{+\infty} \mathrm{e}^{-\alpha t} f(\mathrm{d}t|x',a',y) Q(\mathrm{d}y|x',a') \right. \\
&\left. - \int_{\mathbb{X}} \int_0^{+\infty} \mathrm{e}^{-\alpha t} f(\mathrm{d}t|x',a',y) Q(\mathrm{d}y|x,a) \right| \\
&+ \int_{\mathbb{X}} \int_0^{+\infty} \mathrm{e}^{-\alpha t} |f(\mathrm{d}t|x',a',y) - f(\mathrm{d}t|x,a,y)| Q(\mathrm{d}y|x,a),
\end{aligned}$$

we can see that the function τ_α is a continuous function. □

THEOREM (4.4). *If conditions (2.7) and (4.1) holds then the* ***MDCP*** *is solvable.*

Proof. Let $(m_i)_{i=1}^{\infty}$ be a sequence of occupation measures in $\mathcal{M}_f$ such that $\langle m_i, C_0^\alpha \rangle \searrow \bar{V}$. Let $m_i(\cdot) = m_i(\cdot : \pi, \nu, \alpha)$. By Condition (2.7) the family of occupation measures is uniformly bounded by $\frac{1}{1-\bar{\tau}_\alpha}$.

Now, from Condition (4.1), (a) and (b) the family of occupation measures $\mathcal{M}_f$ is tight. Hence by Prohorov's Theorem, there is a measure m_0 and a subsequence $(m_{n_i})_{i=1}^{\infty}$ which is weakly convergent to m_0.

From this we obtain that its marginals $\hat{m}_{n_i} \to \hat{m}_0$ and by Theorem (3.9) we have $\hat{m}_0 = \nu + (m_0 \otimes \varphi_0)$, then $m_0(\cdot) = m_0(\cdot : \varphi, \nu, \alpha)$.

Now by (4.1), (b) and (c) c_j and d_j are lower semicontinuous functions for $j \in \{0,1,\ldots,N\}$, then by Lemma (2.6) C_j^α is a semicontinuous function for $j \in \{0,1,\ldots,N\}$. By Fatou Lemma and since $\hat{m}_{n_i} \to \hat{m}_0$, $\langle m_0, C_j^\alpha \rangle \le \liminf_k \langle m_k, C_j^\alpha \rangle \le k_j$ for $j \in \{1,\ldots,N\}$. Finally $\bar{V} \le \langle m_0, C_0^\alpha \rangle \le \liminf_k \langle m_k, C_0^k \rangle \le \bar{V}$. □

5. Characterization of the solutions

In this section we shall prove that if the stochastic kernel is nonatomic, then there exists an $N+1$-randomized optimal policy. For this we shall need some preliminaries definitions and lemmas.

Definition (5.1). Let μ a finite (nonnegative) measure on $\mathcal{B}(\mathbb{A})$. Then μ is said to be:

(a) *regular* if $\mu(D) = \sup\{\mu(C) : C \subset D$ and C is closed$\}$ for each Borel set $D \in \mathcal{B}(\mathbb{A})$;

(b) *τ-smooth* if for every decreasing net $(F_\eta)_\eta$ of closed subsets of S we have $\mu(\bigcap_\eta F_\eta) = \inf_\eta \mu(F_\eta)$.

REMARK (5.2). (a) *If $\mathbb{A}$ is a Hausdorff (ot T_2) space, then every Radon measure on $\mathbb{A}$ is τ-smooth, and if $\mathbb{A}$ is regular (or T_3), then every τ-smooth measure is regular (see e.g.,* [30, Proposition I.3.1]*).*

(b) *If $\mathbb{A}$ is strongly Lindelöf (which is the case e.g., if $\mathbb{A}$ is a Suslin space, see* [26, p. 104]*), then every finite measure on $\mathcal{B}(\mathbb{A})$ is τ-smooth. The latter fact and (a) yield, in particular the following.*

(c) *In particular, if $\mathbb{A}$ is a locally compact and separable metric space, the parts (a) and (b) imply that each p.m. on $\mathcal{B}(\mathbb{A})$ is both τ-smooth and regular.*

By Remark (3.5) and [6, Th. 2.6] we get immediately the next theorem.

THEOREM (5.3). *Let $\mathbb{X}$ be an arbitrary topological space, and $\mathbb{A}$ a topological space such that every p.m. on $\mathcal{B}(\mathbb{A})$ is τ-smooth and regular. Fix an arbitrary finite measure $\hat{\mu}$ on $\mathcal{B}(\mathbb{X})$, nonnegative real-valued measurable functions $C_1^\alpha, C_2^\alpha, \ldots, C_N^\alpha$ on $\mathbb{K}$, and real numbers $k_1, \ldots, k_N$. Consider the set $\Lambda \subset \Phi$ that consists of all the randomized strategies $\varphi \in \Phi$ for which*

$$\int C_j^\alpha(x, \varphi(x))\hat{\mu}(\mathrm{d}(x,a)) \le k_j \quad \text{for all} \quad j \in \{1, \ldots, N\}, \tag{5.4}$$

and let $\mathrm{ex}(\Lambda)$ *be the set of extreme points of* Λ*. Then:*

1. *Λ is convex and*

$$\mathrm{ex}(\Lambda) \subset \Phi_{N+1}^0 \tag{5.5}$$

where Φ_{N+1}^0 is the set of all the $(N+1)$-randomizations of the form $\varphi(\cdot|x) = \sum_{j=1}^{N+1} \lambda_j \mathbb{1}.(f_j(x)) \in \Phi_{N+1}$ for some $f_j \in \mathbb{F}$ and nonnegative numbers λ_j such that $\sum_{j=1}^{N+1} \lambda_j = 1$ and the vectors

$$\left(\int C_1^\alpha(x, f_j(x))\hat{\mu}(\mathrm{d}x), \ldots, \int C_N^\alpha(x, f_j(x))\hat{\mu}(\mathrm{d}x), 1\right) \in \mathbb{R}^{N+1}, \tag{5.6}$$

for $j \in \{1, \ldots, N\}$, are linearly independent.

2. *If equality holds in (5.4), then we have equality of the sets in (5.5).*

REMARK (5.7). (a) *Theorem (5.3) requires the action set $\mathbb{A}$ is a topological space such that*

$$\textit{every p.m. on } \mathcal{B}(\mathbb{A}) \text{ is } \tau\textit{-smooth and regular}. \tag{5.8}$$

This condition ensures that the set of extreme points of the space $\mathcal{P}(\mathbb{A})$ of p.m.'s on $\mathbb{A}$ coincides with the set of Dirac measures δ_a for all $a \in \mathbb{A}$ (see, for instance, [29, Th. 11.1]*).*

(b) *If a p.m. is tight, then it is τ-smooth and regular (see* [29, p. xiii]*). It follows that to obtain (5.8) it suffices to give conditions on $\mathbb{A}$ so that every p.m. on $\mathcal{B}(\mathbb{A})$ is tight. This is the case if, for instance, $\mathbb{A}$ is: (i) a σ-compact Hausdorff space; (ii) a Polish space; or (iii) a locally compact separable metric space. (See* [26, 29]*.)*

In the remainder we also consider the following sets:

- The convex cone $\mathcal{D}_+ := \mathbb{R}_+ \cdot \Phi$ of the so-called *transition measures* restricted to $\mathbb{K}$;
- the linear space $\mathcal{D} := \mathcal{D}_+ - \mathcal{D}_+$ of *signed* transition measures with the obvious definitions of sum and scalar multiplication; and
- $\mathcal{M}(\mathbb{K})$ the linear space of finite signed measures on $\mathbb{X} \times \mathbb{A}$ concentrated on $\mathbb{K}$.

Let $l : \mathcal{D} \to \mathcal{M}(\mathbb{K})$ be the linear mapping defined by $l(\varphi) := \nu \otimes \varphi$, where ν is finite measure on $\mathbb{X}$ and $\nu \otimes \varphi$ is as in Remark (3.5). We define the quotient space $\overline{\mathcal{D}} := \mathcal{D}/\ker(l)$, where $\ker(l) := \{\varphi \in \mathcal{D} : (\varphi) = 0\}$ is the kernel of l. For each $\varphi \in \mathcal{D}$, let $\overline{\varphi} := \{\varphi' \in \mathcal{D} : \nu \otimes \varphi' = \nu \otimes \varphi\}$ be the corresponding equivalence class in $\overline{\mathcal{D}}$, and the quotient sets $\overline{\mathbb{F}}, \overline{\Phi}, \overline{\Lambda}, \overline{\Phi}^0_{N+1}$ are defined similarly. For instance, $\overline{\Lambda} := \{\overline{\varphi} : \varphi \in \Lambda\}$.

In the next lemma we use the following notation. If ν is a finite measure on $\mathbb{X}$ and Φ' is a subfamily of randomized strategies in Φ, then $\nu \otimes \Phi' := \{\nu \otimes \varphi : \varphi \in \Phi'\}$. Next lemma can be proved as in [6, Th. 2.6] and [8, Th. 5.6].

LEMMA (5.9). 1. *If $\nu \cdot \varphi$ is an extreme point of $\nu \cdot \Phi$ (resp., $\nu \cdot \Lambda$), then $\overline{\varphi}$ is an extreme point of $\overline{\Phi}$ (resp., $\overline{\Lambda}$).*

2. *If $\overline{\varphi}$ is an extreme point of $\overline{\Phi}$, then $\overline{\varphi}$ has an extreme point f in $\overline{\mathbb{F}} \bigcap \overline{\varphi}$.*

3. *If $\overline{\varphi}$ is an extreme point of $\overline{\Lambda}$, with Λ as in Theorem (5.3), then $\overline{\varphi}$ has an extreme point φ^* of Λ with φ^* in $\overline{\mathcal{R}}^0_{q+1} \bigcap \overline{\varphi}$.*

THEOREM (5.10). *(Bauer's extremum principal (see [2])) If S is a compact convex subset of a locally convex Hausdorff topological vector space, then every l.s.c. concave function on S achieves its minimum at an extreme point.*

Finally we get as in [6, Th. 6.2] and [8, Th. 5.8].

THEOREM (5.11). *Suppose that $\mathbb{A}$ is locally compact separable metric space and that Q and ν are nonatomic, then there exists an $N+1$-randomized optimal policy.*

6. Example

Let $\mathbb{X} = \mathbb{A} = [0, +\infty)$, $A(x) = [0, x]$ and $\nu(B) = \int_B \mathrm{e}^{-t} \mathrm{d}t$, for $B \in \mathcal{B}(\mathbb{X})$.

Let us consider a device such that the probability of passing from a state $x \in \mathbb{X}$ with an action $a \in A(x)$ to a state in $B \in \mathcal{B}(\mathbb{X})$ is given by

$$Q(B|x,a) = \int_B \lambda_2(x+1-a)\exp(-\lambda_2(x+1-a)y)\mathrm{d}y$$

and the transition occurs in a random time whose distribution function is given by

$$F(t|x,a,y) = 1 - \exp(-\lambda_1(x+1-a)yt^2),$$

and so

$$f(t|x,a,y) = (-2\lambda_1(x+1-a)yt)\exp(-\lambda_1(x+1-a)yt^2).$$

The current operation cost functions to be minimized are $c_0(x,a) = \gamma_0 a^2 + \gamma_1 a^2$ and $d_0(x,a) = \eta_0 x^2$. The cost function for which is important to keep under some bounds are $c_1(x,a) = \gamma_2(x-a)^2$ and $d_1(x,a) = \eta_1 x$ which represent some measure of risk associated with big values. Then the distribution G and its density g are independent of (x,a). $G_0(t) = G(t|x,a) = 1 - \frac{\lambda_2}{\lambda_1 t^2 + \lambda_2}$, $g_0(t) = g(t|x,a) = \frac{2\lambda_1\lambda_2 t}{(\lambda_1 t^2 + \lambda_2)^2}$.

The expected value of the time T is $\tau_\alpha = \tau_\alpha(x,a) = \left(\frac{\lambda_2}{\lambda_1}\right)^{\frac{1}{2}} \frac{\pi}{2} = \bar{\tau}_\alpha$. The other costs are $C_0^\alpha(x,a) = \gamma_0 a^2 + \eta_0 \tau_\alpha x^2$, $C_1^\alpha(x,a) = \gamma_1 (x-a)^2 + \eta_1 \tau_\alpha x$. The kernel H_α of finite measures is $H_\alpha(C|x,a) = \int_C e^{-\alpha t} 2\lambda_1\lambda_2 \frac{t}{(\lambda_1 t^2 + \lambda_2)^2} dt$. Now we shall see that this example satisfies all the conditions. First, Condition (2.7) is fulfilled if

$$\frac{\lambda_2}{\lambda_1 \bar{t}^2 + \lambda_2} \geq \bar{\varepsilon} > 0.$$

Condition (4.1): (a) To see that there is a policy $\pi \in \Pi$ such that $V_j^\alpha(\pi,\nu) \leq k_1$ and $V_0^\alpha(\pi,\nu) < +\infty$ let us consider the deterministic stationary policy given by the function $f(x) = x$. Hence

$$V_0^\alpha(\pi,\nu) \leq 2(\gamma_0 + \gamma_1 + \eta_0\tau_\alpha) + \frac{2(\gamma_0 + \gamma_1 + \eta_0\tau_\alpha)}{\lambda_2^2} \sum_{i=1}^{\infty} (\tau_\alpha)^i$$

and

$$V_1^\alpha \leq 1 + \frac{1}{\lambda_1} \sum_{i=1}^{\infty} (\tau_\alpha)^i,$$

that is, this condition is hold if $1 + \frac{1}{\lambda_1} \sum_{i=1}^{\infty} (\tau_\alpha)^i \leq k_1$.

(b), (c) and (d) hold.

(e) The fact that the density function f is uniformly continuous is a consequence of its properties. It is nonnegative, bounded, analytic, the function itself and all its derivatives tend to zero when its argument tend to infinity and is such that the maximum of the absolute value of all its second partial derivatives are reached, hence it is uniformly continuous.

Finally the space $\mathbb{A}$ is locally compact separable metric space and Q and ν are nonatomic.

7. Conclusions and open problems

In this article, for discounted constrained semi-Markov decision processes in Borel spaces, we transform the original control problem in an optimization problems in the space of finite measures. This allowed to demonstrate the existence of solutions to the control problem, to characterize the extreme points of this family and to show there are solutions which are extreme points.

A work in process that we can mention is to find the analogous family of occupation measures for average constrained semi-Markov decision processes in Borel spaces. The target is to demonstrate existence of solutions to the control problem and to characterize the solutions.

In the model worked in this article it was considered that the actions are taken just in base of the previous states and actions, independent of sojourn times. A variant is to allow the actions may depend on the sojourn times also and that the dynamic of the system were described by a stochastic kernel on $\mathbb{X} \times [0,\infty)$ given $\mathbb{X} \times [0,\infty) \times \mathbb{A}$. The problem is to find the family of ocupation measures and to follow the scheme of this article.

Other open problem is to pose the control problem as an infinite linear programming. To do this, the first step is to characterize the family of occupation measures. Moreover this family allows to use also convex programming.

Acknowledgment

The present work was partially supported by the Programa de Apoyo a la Investigación Científica y Tecnológica of the Universidad Autónoma de Nuevo León under the grant CE396-10. The authors wish to thank the anonymous reviewers for their valuable suggestions.

Received May 23, 2012

Final version received April 2, 2013

JUAN GONZÁLEZ-HERNÁNDEZ
DEPARTAMENTO DE PROBABILIDAD Y ESTADÍSTICA
I. I. M. A. S.
UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
A.P. 20-726, 01000 MÉXICO, D.F.
MÉXICO
juan@sigma.iimas.unam.mx

CÉSAR EMILIO VILLARREAL-RODRÍGUEZ
FACULTAD DE INGENIERÍA MECÁNICA Y ELÉCTRICA
UNIVERSIDAD AUTÓNOMA DE NUEVO LEÓN
A.P. 66-F, CIUDAD UNIVERSITARIA
66455 SAN NICOLÁS DE LOS GARZA, N. L.
MÉXICO
cesar.villarrealrd@uanl.edu.mx

REFERENCES

[1] M. S. ABDEL-HAMEED, *Applications of semi-Markov processes in reliability and maintenance*, in Semi-Markov models and Applications, Janssen, Jacques and Limnios, Nikolaos (eds.), Kluwer Academic Publishers, Dordrecht, Netherlands, 1999, 57–63.

[2] C. D. ALIPRANTIS AND K. C. BORDER, Infinite Dimensional Analysis, 2nd ed., Springer-Verlag, Berlin, 1999.

[3] R. B. ASH, Real Analysis and Probability, Academic Press, London, 1972.

[4] X.-R. CAO, *Semi-Markov decision problems and performance sensitivity analysis*, IEEE Trans. Automat. Control **48** (5), (2003), 758–769.

[5] G.-P. DAI, B.-Q. YIN, Y.-J. LI, AND H.-S. XI, *Performance optimization algorithms based on potentials for semi-Markov control processes*, Internat. J. Control **78** (11), (2005), 801–812.

[6] J. GONZÁLEZ-HERNÁNDEZ AND O. HERNÁNDEZ-LERMA, *Extreme points of randomized strategies in constrained optimization and control problems*, SIAM J. Control Optim. **15**, (2004), 1085–1104.

[7] J. GONZÁLEZ-HERNÁNDEZ AND C. E. VILLARREAL, *Optimal policies in constrained average Markov decision processes in Borel spaces*, TOP **19**, (2011), 107–120.

[8] J. GONZÁLEZ-HERNÁNDEZ, R. R. LÓPEZ-MARTÍNEZ, J. A. MINJÁREZ SOSA, AND J. R. GABRIEL-ARGÜELLES, *Constrained Markov Control Processes with Randomized Discounted Cost Criteria: Occupation Measures and Extremal Points*, Risk Decis. Anal. (to appear).

[9] T. HAO, Y. BAOQUN, AND X. HONGSHENG, *Error bounds of optimization algorithms for semi-Markov decision processes*, Error bounds of optimization algorithms for semi-Markov decision processes **38** (9), (2007), 725–736.

[10] T. HAO, X. HONGSHENG, AND Y. BAOQUN, *The optimal robust control policy for uncertain semi-Markov control processes*, Internat. J. Systems Sci. **36** (13), (2005), 791–800.

[11] O. HERNÁNDEZ-LERMA AND J. B. LASSERRE, Discrete-Time Markov Control Processes: Basic Optimality Criteria, Springer-Verlag, New York, 1996.

[12] O. HERNÁNDEZ-LERMA AND J. B. LASSERRE, Further Topics on Discrete-Time Markov Control Processes, Springer-Verlag, New York, 1999.

[13] K. HINDERER, Foundations of Non-Stationary Dynamic Programming Discrete-Time Parameter, Lecture Notes Oper. Res. Syst. vol. 33, Springer, Berlin, 1970.

[14] Q. HU AND W. YUE, *Optimal replacement of a system according to a semi-Markov decision process in a semi-Markov environment*, Optim. Methods Softw. **18** (2), (2003), 181–196.

[15] Y. HUANG AND X. GUO, *Finite horizon semi-Markov decision processes with application to maintenance systems*, European J. Oper. Res. **212** (1), (2011), 131–140.

[16] Y.-H. HUANG AND G. XIAN-PING, *First passage models for denumerable semi-Markov decision processes with nonnegative discounted costs*, Acta Math. Appl. Sin., Engl. Ser. **27** (2), (2011), 177–190.

[17] D. HUDAK AND V. NOLLAU, *Approximation solution and suboptimality for discounted semi-Markov decision problems with countable state space*, Optimization **53** (4), (2004), 339–353.

[18] K. LIU, *Weighted semi-Markov decision processes and their perturbations*, Information **8** (6), (2005), 785–800.

[19] M. LOÈVE, Probability Theory, vol. 1 4th ed., Springer, New York, 1977.

[20] C. E. LOVE, Z. G. ZHANG, M. A. ZITRON, AND R. GUO, *A discrete semi-Markov decision model to determine the optimal repair/replacement policy under general repairs*, European J. Oper. Res. **125** (2), (2000), 398–409.

[21] F. LUQUE-VÁSQUEZ AND J. A. MINJÁREZ-SOSA, *Semi-Markov control processes with unknown holding times distribution under a discounted criterion*, Math. Methods Oper. Res. **61** (3), (2005), 455–468.

[22] F. LUQUE-VÁSQUEZ, J. A. MINJÁREZ-SOSA, AND L. C. ROSAS-ROSAS, *Semi-Markov control models with partially known holding times distribution: discounted and average criteria*, Acta Appl. Math. **114** (3), (2011), 135–156.

[23] F. LUQUE-VÁSQUEZ AND MA. T. ROBLES ALCARAZ, *Controlled semi-Markov models with discounted unbounded costs*, Bol. Soc. Mat. Mexicana (2) **39** (1-2), (1994), 51–68.

[24] A. B. PIUNOVSKII, *Control of random sequences in problems with constraints*, Theory Prob. Appl., **38** (4), (1993), 751–762.

[25] A. B. PIUNOVSKIY, Optimal Control of Random Sequences in Problems with Constraints, Kluwer Academic Publishers, Dordrecht, 1997.

[26] L. SCHWARTZ, Radon Measures on Arbitrary Topological Spaces and Cylindrical Measures, Oxford University Press, London, 1973.

[27] S. S. SINGH, V. B. TADIĆ, AND A. DOUCET, *A policy gradient method for semi-Markov decision processes with application to call admission control*, European J. Oper. Res. **178** (3), (2007), 808–818.

[28] H. TANG, H. XI, AND B. YIN, *The optimal robust control policy for uncertain semi-Markov control processes*, Int. J. Syst. Sci. **36** (13), (2005), 791-800.

[29] F. TOPSOE, Topology and Measure, Lecture Notes in Math. 133, Springer-Verlag, Berlin, 1975.

[30] N. N. VAKHANIA, V. I. TARIELADZE, AND S. A. CHOBANYAN, Probability Distributions in Banach Spaces, D. Reidel, Dordrecht, 1987.

[31] O. VEGA-AMAYA, *Average optimality in semi-Markov control models on Borel spaces: unbounded costs and control*, Bol. Soc. Mat. Mexicana (2), **38**, (1993), 47–60.

[32] B. YIN, Y. LI, Y. ZHOU, AND H. XI, *Performance optimization of semi-Markov decision processes with discounted-cost criteria*, Eur. J. Control, **14** (3), (2008), 213–222.