ELSEVIER
Theoretical Computer Science 273 (2002) 263–269
Theoretical Computer Science
www.elsevier.com/locate/tcs

Quantum evolution of words

V. Malyshev

I.N.R.I.A., Rocquencourt, B.P. 105, 78153, Le Chesnay, Cedex, France

Abstract

Quantum evolution of words is defined. The existence of the evolution semigroup is proved. The spectrum of its generator is calculated for right-linear grammars and expansion–contraction evolutions. © 2002 Elsevier Science B.V. All rights reserved.

Keywords: Words; Quantum grammars; Quantum evolution; C*-algebras; Quantum computing

1. Introduction

We define quantum evolution of words. It is similar in many aspects to the random evolution of words which were called random, probabilistic and stochastic grammars, see [1, 2, 7]. The quantum evolution of words has two main motivations. One comes from the computer science and is related to quantum computation problems. The other comes from modern physics, more exactly quantum gravity. Theoretical studies in both fields become more and more popular.

However, quantum evolution of words can be interesting by itself because it relates two classical mathematical domains: combinatorics of words gives many new models of linear operators in the Hilbert space, see also [3, 4].

In this paper we define the evolution itself and prove its existence, that is the self-adjointness of the generator of the evolution semigroup. For two simplest examples we calculate the spectral decomposition of the generator. It is interesting how famous Fock spaces appear in the context of quantum evolution of words.

2. Hilbert space and Hamiltonian

Let $\Sigma = \{1,\dots,r\}$ be a finite set (the alphabet), Σ^* the set of all finite words (including the empty one) $\alpha = x_1 \dots x_n, x_i \in \Sigma$, in this alphabet. Length n of the word

E-mail addresses: malyshev@tws05.math.msu.su, vadim.malyshev@inria.fr (V. Malyshev).

PII: S0304-3975(00)00444-8

α is denoted by $|\alpha|$. Concatenation of two words $\alpha = x_1 \dots x_n$ and $\beta = y_1 \dots y_m$ is defined by

$$\alpha\beta = x_1 \dots x_n y_1 \dots y_m.$$

The word β is a factor (we use standard notation, see [5, 6]) of α if there exist words δ and γ such that $\alpha = \delta\beta\gamma$. A grammar over Σ is defined by a finite set *Sub* of substitutions (productions), that is the pairs $\delta_i \to \gamma_i, i = 1, \dots, k = |Sub|, \delta_i, \gamma_i \in \Sigma^*$. Further on we assume that all δ_i, γ_i are not empty.

Let $H = l_2(\Sigma^*)$ be the Hilbert space with the orthonormal basis $e_\alpha, \alpha \in \Sigma^*$: $(e_\alpha, e_\beta) = \delta_{\alpha\beta}$ where the function $e_\alpha(\beta) = \delta_{\alpha\beta}$. Each vector ϕ of H is a function on the set of words and can be written as

$$\phi = \sum \phi(\alpha) e_\alpha \in \mathscr{H}, \qquad \|\phi\|^2 = \sum |\phi(\alpha)|^2.$$

States of the system are wave functions, that is vectors ϕ with the unit norm $\|\phi\|^2 = 1$. We shall define dynamics in the form

$$\phi(t) = \exp(itH)\phi(0).$$

The Hamiltonian H will be written in terms of operators, which resemble creation–annihilation operators (or q-creation–annihilation, see [4]) in quantum field theory. For each $i = 1, \dots, k$ and each integer $j \geqslant 1$ we define quantum substitutions, that is linear bounded operators $a_i(j)$. If $\alpha = \tau\delta_i\rho$ for some words $\tau, \rho, |\tau| = j - 1$, we put

$$a_i(j) e_\alpha = e_\beta,$$

where $\beta = \tau\gamma_i\rho$. Otherwise we put $a_i(j)e_\alpha = 0$. Adjoint operators $a_i^*(j)$ are then defined by

$$a_i^*(j) e_\beta = e_\alpha$$

for $\beta = \tau\gamma_i\rho$ and 0 otherwise. Define the formal Hamiltonian by

$$H = \sum_{i=1}^{|Sub|} \sum_{j=1}^{\infty} (\lambda_i a_i(j) + \bar{\lambda}_i a_i^*(j))$$

for some complex numbers $\lambda_i, i = 1, \dots, |Sub|$.

We can equally assume that together with the substitution $\delta_i \to \gamma_i$ also its "inverse" substitution $\gamma_i \to \delta_i$ belongs to *Sub*. The Hamiltonian, then, can be written simply as

$$H = \sum_{i=1}^{|Sub|} \sum_{j=1}^{\infty} \lambda_i a_i(j).$$

We always assume that $H = H^*$, that is $\lambda_i = \bar{\lambda}_j$ in case $\delta_j = \gamma_i, \gamma_j = \delta_i$. We shall use only this representation further on.

H is well defined and symmetric on the set $D(\Sigma^*)$ of finite linear combinations of e_α. These vectors are C^∞-vectors (see [8]) for H, that is $He_\alpha \in D(\Sigma^*)$. Note that H is unbounded in general.

Theorem 1. *H is essentially self-adjoint on $D(\Sigma^*)$.*

Proof. We shall prove that each vector $\phi \in D(\Sigma^*)$ is an analytic vector of H, that is

$$\sum_{k=0}^{\infty} \|H^k \phi\| k! t^k < \infty$$

for some $t > 0$. It is sufficient to take $\phi = e_\alpha$ for some α. In this case, the number of pairs (i,j) such that $a_i(j)e_\alpha \neq 0$ is not greater than $nk, n = |\alpha|, k = |Sub|$.

Write the decomposition of H as

$$H = \sum_a V_a,$$

where V_a equals one of $\lambda_i a_i(j)$. Then

$$H^n e_\alpha = \sum_{a_n,\dots,a_1} V_{a_n} \dots V_{a_1} e_\alpha = \sum C_\beta e_\beta. \tag{1}$$

The maximal length of the words β in the expansion of $V_{a_n} \dots V_{a_1} e_\alpha$ does not exceed $|\alpha| + C_1 n, C_1 = \max(|\gamma_i| - |\delta_i|)$. Then, for given $e_\alpha, a_1, \dots, a_n$, the number of operators $V_{a_{n+1}}$ giving a nonzero contribution to $V_{a_{n+1}} V_{a_n} \dots V_{a_1} e_\alpha$. It does not exceed $k(|\alpha| + C_1 n), k = |Sub|$. Thus, the number of nonzero terms $V_{a_n} \dots V_{a_1} e_\alpha$ does not exceed

$$k^n \prod_{j=1}^{n} (|\alpha| + C_1 j) = (kC_1)^n (|\alpha| C_1 + n)! n! (|\alpha| C_1)! \leqslant (kC_1)^n n^{|\alpha| C_1}$$

and the norm of each term is bounded by $(\max \lambda_i)^n$. This gives the convergence of the series for $|t| < t_0$, where t_0 does not depend on α. □

3. Spectrum

Due to our substitution rules $e_\emptyset$ is the zero eigenvector of H. There can be other invariant subspaces of $\exp(itH)$ of the special type: $\mathscr{H}(L') = l_2(L')$, generated by $e_\alpha, \alpha \in L'$, where L' is a subset of the set Σ^* of all words. We have

$$\mathscr{H} = \oplus_{k=0}^{\infty} \mathscr{H}(L_k)$$

if $\mathscr{H}(L_k)$ are invariant and $\Sigma^* = \cup_{k=0}^{\infty} L_k$.

3.1. Right-linear grammars

Consider the case when $\Sigma = \{a, w\}$ consists of two letters and the only substitutions are $aw \to w, w \to aw$. We put $\lambda(aw \to w) = \lambda(w \to aw) = \lambda$. Then the subspaces $\mathscr{H}(L_k)$, where L_k is the set of words with exactly k symbols w, are invariant. For example, $L_1 = \{a^m w = aa \dots aw, m = 0, 1, 2, \dots\}$. We shall call $\mathscr{H}(L_k)$ the k-particle space. Let H_k be the restriction of H onto $\mathscr{H}(L_k)$.

Theorem 2 (One-particle spectrum). *H_1 on $\mathscr{H}(L_1)$ is unitary equivalent to the multiplication operator $\lambda(z + z^{-1})$ in $L_2(S^1, dv)$ where dv is the Lebesgue measure on the unit circle S^1 in the complex plane C.*

$\mathscr{H}(L_1)$ is isomorphic to $l_2(Z_+)$ via $aa\dots aw \to m = l(a^m w)$. Then H becomes a Toeplitz operator in Z_+ equal to

$$\lambda b + \lambda b^*,$$

where b is the left shift in l_2

$$(bf)(m) = f(m-1), \quad m = 1,2,\dots, \quad (bf)(0) = 0.$$

Denote W_1 the C^*-algebra of Toeplitz operators, that is the C^*-algebra of operators in $l_2(Z_+)$ generated by b. Note that the commutator $[b,b^*]$ has finite rank. More generally, the following exact sequence of algebras holds:

$$0 \to K \to^j W_1 \to^\pi W \to 0,$$

where K is the closed two-sided ideal in W_1, consisting of all compact operators in $l_2(Z_+)$, j is the embedding, π the natural projection of W_1 onto $W_1/K \sim W$. In fact, the commutator of any two elements from W_1 is compact. This can be easily checked for monomials $b_1 b_2 \dots b_k, b_i = b$ or b^*. By limiting procedure this can be proved for any elements. It follows that W_1/K is commutative. Substitute

$$y = \pi b, \quad y^* = y^{-1} = \pi b^*.$$

One can prove that the elements

$$\sum_{i=-k}^{k} c_i y^i$$

are all different and that their norm equals $\sum |c_i|$. It is sufficient to consider its action of the function $f(m)$ equal to 1 for some m sufficiently large and zero otherwise. It follows that $W_1/K \sim W$. We proved that the continuous part of the spectrum is the same as the spectrum of the multiplication operator.

To prove the absence of discrete spectrum, we rewrite the equation

$$Af - \lambda f = \phi$$

as a well-known difference equation

$$\phi_n = -\lambda f_n + f_{n-1} + f_{n+1}, \quad n \geqslant 1,$$
$$\phi_0 = -\lambda f_0 + f_1.$$

With $\phi_n \equiv 0$, there are no proper l_2-solution.

Another way to get the spectrum could be to use generating functions $F(z) = \sum_{k=0}^{\infty} f_k z^k, \phi(z) = \sum_{k=0}^{\infty} \phi_k z^k$. We have

$$\phi(z) + \frac{f_0}{z} = \left(-z + \frac{1}{z} - \lambda\right) F(z).$$

The quadratic equation has the roots $z = \lambda/2 \pm \sqrt{\lambda^2/4 - 1}$. For $|\lambda| > 2$ one root is inside the unit circle and the other outside, thus the equation has a unique solution. This solution is obtained by putting z equal to root inside the unit circle, from this we get f_0. If $|\lambda| < 2$ and λ is not real, then there are two roots: one inside and the other outside the unit circle (take λ imaginary to check it), thus the equation has a unique solution. For $|\lambda| < 2$ with λ real, there are two conjugate complex roots on the unit circle. The operator is not invertible in this case.

Now, we consider H on all the Hilbert space. Each word can be represented like this

$$\alpha = a^{m_1}wa^{m_2}w\ldots a^{m_k}w, \quad m_1, m_2, \ldots, m_k = 0, 1, 2, \ldots .$$

Correspondingly the vectors of the basis in $\mathscr{H}(L_k)$ can be written as a tensor product as

$$e_\alpha = e_{\alpha(1)} \otimes e_{\alpha(1)} \otimes \cdots \otimes e_{\alpha(1)}, \quad \alpha(i) = a^{m_i}w.$$

Otherwise speaking $\mathscr{H}(L_k)$ is the kth tensor power of $\mathscr{H}(L_1)$ and the evolution is (about tensor products see, for example, [8])

$$\exp(\mathrm{i}tH_1) \otimes \cdots \otimes \exp(\mathrm{i}tH_1)$$

or

$$H_k = H_1 \otimes 1 \otimes \cdots \otimes 1 + \cdots + 1 \otimes 1 \otimes \cdots \otimes H_1.$$

Then we have the following:

Theorem 3 (Multi-particle spectrum). *The Hilbert space has the following decomposition on invariant subspaces:*

$$\mathscr{H}(\Sigma^*) = \oplus_{k=0}^{\infty}\mathscr{H}(L_k), \quad \mathscr{H}(L_k) = \mathscr{H}(L_1) \otimes \cdots \otimes \mathscr{H}(L_1).$$

The Hamiltonian H_k on $\mathscr{H}(L_k)$ is unitary equivalent to the multiplication on the function

$$\sum_{j=1}^{k} \lambda(z_j + z_j^{-1})$$

in the space $L_2(S^k, dv)$ of functions $f(z_1, \ldots, z_k)$.

3.2. Expansion–contraction evolution of words

Take $\Sigma = \{a, b\}$ and substitutions

$$a \to aa, \quad aa \to a, \quad b \to bb, \quad bb \to b$$

with

$$\lambda(a \to aa) = \lambda(aa \to a) = \lambda, \quad \lambda(b \to bb) = \lambda(bb \to b) = \mu.$$

Here there are two invariant one-particle spaces $\mathscr{H}_a, \mathscr{H}_b$ with cyclic vectors e_a, e_b correspondingly. For example, $\mathscr{H}_a$ is generated by words $a, a^2 = aa, \ldots, a^n, \ldots$. We shall see now the structure of the Hamiltonian on this subspaces.

Consider one symbol alphabet $\Sigma = \{a\}$ and with the following substitutions:

$$a \to aa, \quad aa \to a.$$

We take $\lambda(a \to aa) = \lambda(aa \to a) = \lambda$. If the Hamiltonian

$$H_a = \lambda \sum_{j=1}^{\infty} (a_1(j) + a_2(j))$$

with real λ. Then the Hilbert space $\mathscr{H}$ is isomorphic to $l_2(Z_+)$, because the word $aa\ldots a$ can be identified with its length minus 1. The Hamiltonian is unitarily equivalent to the Jacobi operator H_1 on $\{1,2,\ldots\}$ defined by

$$Hf(n) = \lambda(nf(n+1) + (n-1)f(n-1)).$$

Consider now the equation

$$(H_1 - \lambda)f = \phi$$

or

$$\phi_n = nf_{n+1} + (n-1)f_{n-1} - \lambda f_n.$$

Introducing the generating functions

$$F(z) = \sum_{n=1}^{\infty} f_n z^n, \quad \Phi(z) = \sum_{n=1}^{\infty} \phi_n z^n,$$

we get the following equations for them:

$$\begin{aligned}\Phi(z) &= -\lambda F(z) + \sum_{n=1}^{\infty} nf_{n+1}z^n + \sum_{n=2}^{\infty}(n-1)f_{n-1}z^n \\ &= \sum_{n=1}^{\infty}(n+1)f_{n+1}z^n - \sum_{n=1}^{\infty} f_{n+1}z^n + z^2\sum_{n=1}^{\infty}(n-1)f_{n-1}z^{n-2} - \lambda F \\ &= (F - f_1 z)' - \frac{1}{z}(F - f_1 z) + z^2 F' = F'(z)(1+z^2) - \left(\lambda + \frac{1}{z}\right)F(z)\end{aligned}$$

or

$$F' - \frac{\lambda + \frac{1}{z}}{1+z^2}F = \frac{\phi}{1+z^2}.$$

Lemma 4. *H does not have discrete spectrum.*

Proof. Consider the homogeneous equation. It has the following solution:

$$\exp\left(\int \frac{\lambda + \frac{1}{z}}{1+z^2}\,\mathrm{d}z\right) = z\left(\frac{z+i}{z-i}\right)^{\frac{1}{2}+\frac{i\lambda}{2}}$$

For any λ, the solution has a nonalgebraic singularity in at least one of the points $\pm i$. This singularity cannot get l_2-solution $f(n)$.

Using these techniques, Shestopal proved (see [9]) that the spectrum is absolutely continuous on $(-\infty, \infty)$.

For all Hamiltonians we can write a multi-particle representation as follows. There are two invariant two-particle spaces $\mathscr{H}_{ab} = \mathscr{H}_a \otimes \mathscr{H}_b, \mathscr{H}_{ba} = \mathscr{H}_b \otimes H_a$ with cyclic vectors e_{ab}, e_{ba} correspondingly. For example, $\mathscr{H}_{ab}$ is generated by vectors $e_{a^k b^l}, k, l > 0$. In general, for each n there are two invariant $2n$-particle spaces $\mathscr{H}_{(ab)_n} = \mathscr{H}_a \otimes \mathscr{H}_b \otimes \cdots \otimes \mathscr{H}_a \otimes \mathscr{H}_b$. It is generated by words $a^{k_1} b^{l_1} \cdots a^{k_n} b^{l_n}$, and $\mathscr{H}_{(ba)_n}$ defined similarly. For each n there are two invariant $(2n+1)$-particle spaces $\mathscr{H}_{b(ab)^n} = \mathscr{H}_b \otimes \mathscr{H}_a \otimes \mathscr{H}_b \otimes \cdots \otimes \mathscr{H}_a \otimes \mathscr{H}_b$ and $\mathscr{H}_{(ab)^n a}$. $\square$

Theorem 5. *We have the following decomposition:*

$$\mathscr{H} = \mathscr{H}_0 \oplus \mathscr{H}_a \oplus \mathscr{H}_b \oplus \mathscr{H}_{ab} \oplus \mathscr{H}_{ba} \oplus \cdots$$
$$\oplus \mathscr{H}_{(ab)_n} \oplus \mathscr{H}_{(ba)_n} \oplus \mathscr{H}_{b(ab)^n} \oplus \mathscr{H}_{(ab)^n a} \cdots.$$

On each of these invariant subspaces, the Hamiltonian is unitarily equivalent to the corresponding sum of one-particle Hamiltonians. For example, on $\mathscr{H}_{ab}$ it is equivalent to $\mathscr{H}_a \otimes 1 + 1 \otimes \mathscr{H}_b$.

Remark 6. We call this evolution expansion–contraction evolution because the lengths of maximal factors with only one symbol can expand and contract. The law of this expansion can be obtained using the spectral decomposition. We do not do it here.

References

[1] R. Bod, Beyond Grammar, CSLI Publications, 1998.
[2] E. Charniak, Statistical Language Learning, MIT Press, Cambridge, MA, 1993.
[3] G. Duchamp, J. Katriel, Ordering relations for q-boson operators, continued fraction techniques, and the q-CBH enigma, J. Phys. A 28 (1995) 7209–7225.
[4] G. Duchamp, D. Krob, B. Leclerc, J.Y. Thibon, Non commutative symmetric functions III: deformations of cauchy and convolution structures, Discrete Math. Theoret. Comput. Sci.
[5] S. Eilenberg, Automata, Languages, and Machines, Vol. A, Academic Press, New York, 1974.
[6] N. Lothaire, Combinatorics on Words, Cambridge University Press, Cambridge, 1997.
[7] V.A. Malyshev, Random grammars, Russian Math. Surveys 53 (2) (1998) 107–134.
[8] M. Reed, B. Simon, Methods of Modern Mathematical Physics, Vols. 1–4, Academic Press, New York, 1980.
[9] V.E. Shestopal, About the theory of quantum grammars, preprint 12-00, Institute of Theoretical and Experimental Physics, Moscow, 2000.