Computational and Theoretical Chemistry 1115 (2017) 80–87

Contents lists available at ScienceDirect

Computational and Theoretical Chemistry

journal homepage: www.elsevier.com/locate/comptc

From quantum superposition to orbital communication

Dariusz W. Szczepanik [a,*], Emil Zak [b], Janusz Mrozek [c]

[a] K. Guminski Department of Theoretical Chemistry, Jagiellonian University, Ingardena 3, 30-060 Krakow, Poland
[b] Department of Physics and Astronomy, University College London, WC1E 6BT, UK
[c] Department of Computational Methods in Chemistry, Jagiellonian University, Ingardena 3, 30-060 Krakow, Poland

ARTICLE INFO

Article history:
Received 25 April 2017
Received in revised form 30 May 2017
Accepted 31 May 2017

Keywords:
Superposition principle
Information theory
Shannon entropy
Markov chain
Orbital communication
Symmetrical orthogonalization

ABSTRACT

The orbital communication theory (OCT) by Nalewajski is derived step by step from first principles of quantum mechanics. It is shown that the entropy representation within the molecular orbital theory arises as a natural consequence of the probabilistic interpretation of quantum superposition. The algebra of selected types of molecular information channels is reinvestigated within the framework of the theory of Markov chains and several representative models of molecular communication systems in atomic-orbital resolution are discussed. The presented results show that the Shannon entropy alone, i.e. with no insight into its components - mutual information and conditional entropy, does not allow one to correctly identify the source of uncertainty connected with the electron probability distribution, which in some cases leads to wrong conclusions about the electron delocalization effects in a molecule.

1. Introduction

The probability is one of the most important concepts that we unwittingly and almost involuntarily use in everyday life when making decisions and assessing risk. The theory of probability underlies statistical modeling methods that are widely applied in business, insurance industry, marketing, gambling, environmental and financial regulations, entitlement analyses, and many others [1–3]. In science, it provides the conceptual and mathematical basis for statistical mechanics and quantum theory, although the roles that it plays in both are known to be fundamentally different [4,5]. In the field of chemistry, there actually would not be the entire analytical chemistry without the analysis of measurement errors [6]. The probability theory also enables one to elucidate a more precise meaning of several important but semantically not sharply defined terms like molecular similarity, open subsystem, molecular fragment and the chemical bonding concept itself [7]. It were Giambiagi and Mayer who first found that the effect of electron delocalization between two chemically bonded atoms can be evaluated using statistical tools that measure linear dependence of the electron probability distributions on each bonded atom [8,9]. A couple of years later Yamasaki and Goddard generalized this concept to a hierarchy of quantities, i.e. partial atomic charges, bond covalencies, bond-bond correlation coefficients, where each quantity is the expectation value of an operator related to the statistical covariance of the previous quantity [10,11]. In turn, Mohajeri and Dasmeh introduced and examined several different probabilistic, information-theoretic and fuzzy models to quantify covalency and electrovalency of the chemical bond [12]. But probably the most comprehensive description of the chemical bonding grounded on purely probabilistic approach is the orbital communication theory (OCT) by Nalewajski [13–22]. The theory has been developed over the last decade within the framework of studies on the application of information-theoretic tools in the molecular electronic structure theory, and so far has been largely outlined in four monographs and more than one hundred papers by its author. In OCT the standard tools used in analyzing and comparing the probability distributions and their dependencies within the statistical theory of communication are utilized in probing the chemical bonding. As demonstrated in numerous studies, description of the electronic structure of chemical bonds within the OCT framework enables one to precisely define such concepts as chemical valence, bond multiplicity, covalency and ionicity and many other general laws of chemistry [23–26]. Furthermore, adopting the representation of entropy in the description of molecular electronic structure opens up new possibilities to practical studies of dynamical electron correlation [27–29], through-space and through-bridge chemical interactions [30], many-orbitals conjugation effects [31,32], the extent of electron delocalization in aromatic and heteroaromatic species [33–36], etc.

In this paper we investigate step by step how the entropy representation in the formalism of orbital communications emerges from first principles of quantum mechanics (the 2nd

* Corresponding author.
E-mail address: dariusz.szczepanik@uj.edu.pl (D.W. Szczepanik).

http://dx.doi.org/10.1016/j.comptc.2017.05.041
2210-271X/

and 3rd Section) and provide a brief review of the original OCT formulation by Nalewajski (a part of the 3rd and 4th Section). Also, we examine some properties of the algebra of molecular information channels within the framework of the theory of Markov chains and discuss several representative probabilistic models of molecular communication systems in atomic-orbital resolution (the 5th Section).

2. Conditional probabilities of quantum states

Quantum superposition is probably one of the most fundamental postulates of quantum theory [37]. It states that for a given set of allowed quantum state vectors, $\{|\chi_i\rangle\}$, any linear combination

$$|\psi_k\rangle = \sum_{i=1}^{n} \hat{P}_i|\psi_k\rangle = \sum_{i=1}^{n} |\chi_i\rangle\langle\chi_i|\psi_k\rangle = \sum_{i=1}^{n} |\chi_i\rangle C_{i,k}, \tag{1}$$

also represents a possible quantum state; here $\hat{P}_i$ stands for the operator projecting onto state $|\chi_i\rangle$, while $C_{i,k}$ denotes the corresponding linear expansion coefficient. Within the basis of orthonormal state vectors, $\langle\chi_i|\chi_j\rangle = \delta_{i,j}$, normalization of the superposed $|\psi_k\rangle$vector requires that

$$\langle\psi_k|\psi_k\rangle = \sum_{i=1}^{n} \langle\psi_k|\chi_i\rangle\langle\chi_i|\psi_k\rangle = \sum_{i=1}^{n} |C_{i,k}|^2 = 1. \tag{2}$$

This normalization condition provides the interpretation of $C_{i,k}$ as an amplitude ($a_i^{(k)}$) of the probability ($p_i^{(k)}$) of finding quantum system in state $|\chi_i\rangle$provided that a full state of the system is described by the superposed vector $|\psi_k\rangle$,

$$p_i^{(k)} = |a_i^{(k)}|^2 = \langle\psi_k|\hat{P}_i|\psi_k\rangle = |C_{i,k}|^2, \qquad \sum_{i=1}^{n} p_i^{(k)} = 1. \tag{3}$$

In other words, $p_i^{(k)}$ can be interpreted as an expectation value of the projection operator $\hat{P}_i$ taken in the quantum state $|\psi_k\rangle$. In turn, the $\hat{D}_k = |\psi_k\rangle\langle\psi_k|$ projector is called the density operator and gives rise to the density matrix $\mathbf{D}_k$ in the basis of $|\boldsymbol{\chi}\rangle$:

$$\mathbf{D}_k = \langle\boldsymbol{\chi}|\hat{D}_k|\boldsymbol{\chi}\rangle = \langle\boldsymbol{\chi}|\psi_k\rangle\langle\psi_k|\boldsymbol{\chi}\rangle = \mathbf{C}_k\mathbf{C}_k^{\dagger}, \tag{4}$$

where $\mathbf{C}_k$ denotes the column vector of the corresponding linear combination coefficients; normalization condition (2) immediately implies that $\mathbf{D}_k$ is an idempotent matrix,

$$\mathbf{D}_k^2 = \mathbf{C}_k\mathbf{C}_k^{\dagger}\mathbf{C}_k\mathbf{C}_k^{\dagger} = \mathbf{C}_k\left[\sum_{i=1}^{n} |C_{i,k}|^2\right]\mathbf{C}_k^{\dagger} = \mathbf{C}_k^{\dagger}\mathbf{C}_k = \mathbf{D}_k. \tag{5}$$

Diagonal elements of $\mathbf{D}_k$ have the same probabilistic interpretation as the elements of the p-vector from (3), while the off-diagonal elements can be interpreted as amplitudes of the probabilities of finding (due to measurement) the quantum system given by $|\psi_k\rangle$ to be described simultaneously (in a statistical sense) by a pair of different vector states, $|\chi_i\rangle$ and $|\chi_j\rangle$,

$$\mathbf{D}_k \equiv \mathbf{A}_k(\boldsymbol{\chi},\boldsymbol{\chi}) = \left\{A_k(i,j) = C_{i,k}C_{k,j}^{\dagger}\right\}. \tag{6}$$

What follows, the elements of the corresponding joint probability matrix $\mathbf{P}_k(\boldsymbol{\chi},\boldsymbol{\chi})$ read:

$$P_k(i,j) = |A_k(i,j)|^2 = |C_{i,k}|^2|C_{j,k}|^2 = p_i^{(k)}p_j^{(k)}, \tag{7}$$

with normalization condition

$$\sum_{i=1}^{n}\sum_{j=1}^{n} P_k(i,j) = \sum_{i=1}^{n} p_i^{(k)} = \sum_{j=1}^{n} p_j^{(k)} = 1. \tag{8}$$

Recalling Bayes's theorem [38], which refers to the symmetry of probabilities of simultaneous events, we may write

$$P_k(i,j) = P_k(i|j)p_j^{(k)} = P_k(j|i)p_i^{(k)} = P_k(j,i), \tag{9}$$

where $P_k(j|i)$ is the corresponding conditional probability of the jth vector state given the ith one (from the statistical point of view, the superposition of states leads to uncertainty of the exact quantum-mechanical state identification); comparing the above formula with (7) we straightforwardly get

$$P_k(i|j) = p_i^{(k)}, \qquad \text{and} \qquad P_k(j|i) = p_i^{(k)}, \tag{10}$$

which implies that in the superposed state $|\psi_k\rangle$the statistics of $|\chi_j\rangle$ and $|\chi_i\rangle$ states from a large number of measurements on replicas of the system are independent (no correlation). Therefore, the conditional probability (transition) matrix $\mathbf{P}_k(\boldsymbol{\chi}|\boldsymbol{\chi})$ is a first-order matrix and comprises the exact copies of the probability vector $\mathbf{p}^{(k)} = \left\{p_i^{(k)}\right\}$ in each row, i.e.

$$\mathbf{P}_k(\boldsymbol{\chi}|\boldsymbol{\chi}) = \left[\mathbf{p}^{(k)},\ldots c,\mathbf{p}^{(k)}\right]^{\mathrm{T}}, \qquad \mathrm{diag}\mathbf{P}_k(\boldsymbol{\chi}|\boldsymbol{\chi}) = \mathbf{p}^{(k)}. \tag{11}$$

If the quantum system in question is described exactly by a single superposed vector, $|\psi_k\rangle$, then $\mathbf{p}^{(k)}$ represents a stationary probability distribution in a Markov sense [39] and, in accordance with the Perron-Frobenius theorem [40],

$$\lim_{x\to\infty}[\mathbf{P}_k(\boldsymbol{\chi}|\boldsymbol{\chi})]^x = \mathbf{1}\mathbf{P}_k(\boldsymbol{\chi}|\boldsymbol{\chi}), \tag{12}$$

it represents a normalized left eigenvector of the conditional probability matrix associated with the unit eigenvalue; in quantum mechanics, this state is usually referred to as the pure state. It is worth noting that for any normalized distribution $\mathbf{p}^0 = (p_1^0,\ldots c,p_n^0)$ the transition matrix $\mathbf{P}_k(\boldsymbol{\chi}|\boldsymbol{\chi})$ always recovers a stationary distribution, i.e.

$$\mathbf{p}^{(k)} = \mathbf{P}_k(\boldsymbol{\chi}|\boldsymbol{\chi})^{\mathrm{T}}\mathbf{p}^0. \tag{13}$$

Unlike the pure state, a mixed quantum state generally does not reflect proper quantum-state features (so it is not possible to assign a state vector to the system). Instead, it constitutes a statistical ensemble of pure states $\{|\psi_k\rangle\}$ given by the probability vector $\boldsymbol{p} = \{\boldsymbol{p}_k\}$ and the associated density operator in the following form:

$$\hat{D} = \sum_{k=1}^{n} \boldsymbol{p}_k\hat{D}_k = \sum_{k=1}^{n} \boldsymbol{p}_k|\psi_k\rangle\langle\psi_k|, \qquad \sum_{k=1}^{n} \boldsymbol{p}_k = 1. \tag{14}$$

In the basis of pure state vectors $\{|\chi_i\rangle\}$ it reads

$$\hat{D} = \sum_{i=1}^{n}\sum_{j=1}^{n} \hat{P}_i\hat{D}\hat{P}_j = |\boldsymbol{\chi}\rangle\mathbf{C}\boldsymbol{p}\mathbf{C}^{\dagger}\langle\boldsymbol{\chi}| = |\boldsymbol{\chi}\rangle\mathbf{D}\langle\boldsymbol{\chi}|, \tag{15}$$

and hence

$$\mathbf{D} = \mathbf{C}\boldsymbol{p}\mathbf{C}^{\dagger} = \left\{D_{i,j} = \sum_{k=1}^{n} \boldsymbol{p}_k C_{k,i}^{\dagger}C_{j,k}\right\}. \tag{16}$$

In contrast to its pure-state counterpart, $\hat{D}$ is generally not idempotent,

$$\hat{D}^2 \neq \hat{D}, \qquad \sum_{j=1}^{n} |D_{i,j}|^2 \neq D_{i,i} = p_i, \tag{17}$$

which means that the element of the symmetric matrix $\mathbf{D}$ no longer represents a probability amplitude of simultaneous events and the probability distribution in $\mathbf{p} = \mathrm{diag}\mathbf{D}$ is not stationary in a Markov's sense. However, a special case of statistical ensemble with the associated idempotent density operator (up to a constant) is a uniform probabilistic mixture of m pure states $\{|\psi_k\rangle\}$. For the sake of simplicity let us assume that

$$\boldsymbol{p} = \left\{\underbrace{m^{-1},\ldots c,m^{-1}}_{m}, \quad \underbrace{0,\ldots c,0}_{n-m}\right\}. \tag{18}$$