


Entropy of pure states: not all wave functions are born equal

Denis K. Sunko^{a,*} 

Department of Physics, Faculty of Science, University of Zagreb, 10000 Zagreb, Croatia

Received 16 September 2021, Accepted 17 November 2021

Abstract – Many-body Hilbert space has the algebraic structure of a finitely generated free module. All N -body wave functions in d dimensions can be generated by a finite number of $N!^{d-1}$ of generators called *shapes*, with symmetric-function coefficients. Physically the shapes are vacuum states, while the symmetric coefficients are bosonic excitations of these vacua. It is shown here that logical entropy can be used to distinguish fermion shapes by information content, although they are pure states whose usual quantum entropies are zero. The construction is based on the known algebraic structure of fermion shapes. It is presented for the case of N fermions in three dimensions. The background of this result is presented as an introductory review.

Keywords: Many-body Hilbert space, Free module, Logical entropy, Quantum entropy, Bargmann space

Introduction

The realization that the space of physical states is a Hilbert space is arguably the milestone separating Bohr's "old" quantum mechanics from its modern form. It came about when von Neumann connected Heisenberg's and Schrödinger's formulations as two different coordinatizations of the same abstract vector space [1]. Heisenberg's alternant form [2] for many-fermion wave functions became the formal basis of all future developments under the name of Slater determinants [3]. Mathematically, these are *primitive realizations* of the Pauli principle: using them in calculations is a formal safeguard against violating the principle.

It appears now that these foundational developments did not go as far as mathematically possible [4]. Many-body Hilbert space has an additional finer structure, which is difficult to visualize in the standard formulations. Specializing to fermions from now on, the difficulty can be illustrated by the difference between *trivial* and *non-trivial* zeros of the wave function.

For definiteness, consider the Hermite functions $\psi_n(x) \sim H_n(x)e^{-x^2/2}$, which are a complete orthonormal basis of solutions of the one-dimensional harmonic oscillator in the Hilbert space $L^2(\mathbb{R})$. Because they are in real space, they can be orthogonal only by having n distinct zeros on the line for each n . These zeros are trivial: they appear because we insist on a real-space coordinatization, perhaps in the belief that real space is somehow "easier."

When the one-body wave functions are combined in a Slater determinant, non-trivial zeros appear as a consequence of the Pauli principle. These constitute the *nodal surface* of the many-body wave function [5]. However, they are notationally hidden in the real-space formulation, where the nodal surface of a superposition of Slater determinants appears as an interference effect of complicated oscillations imposed by the trivial zeros.

In 1961, Bargmann [6] recast the oscillator problem by an integral transform to a space $F(\mathbb{C})$, the *Bargmann space* of entire functions $f : \mathbb{C} \rightarrow \mathbb{C}$ such that,

$$\int_{\mathbb{C}} |f(t)|^2 d\lambda(t) < \infty, \quad (1)$$

with scalar product,

$$(f, g) = \int_{\mathbb{C}} \bar{f}g d\lambda(t), \quad d\lambda(t) = \frac{1}{\pi} e^{-|t|^2} d \operatorname{Re} t d \operatorname{Im} t. \quad (2)$$

*Corresponding author: dko@phy.hr

^aSupported by the Croatian Science Foundation Project No. IP-2018-01-7828.

The transform of the oscillator basis is just $\psi_n(x) \leftrightarrow t^n/\sqrt{n!}$. Bargmann's transform was historically the first of so-called wavelet transforms, with important contemporary applications to image analysis and signal processing. His motivation was however purely theoretical: to have a Hilbert space in which t and ∂_t would be conjugates. Indeed, they act as creation and destruction operators of excitations represented by the wave function t^n .

From the present point of view, the merit of Bargmann's formulation is that quantum numbers appear as powers, so that $t^n t^m = t^{n+m}$, and trivial zeros are manifestly trivial: they are the n -fold zeros of simple monomials t^n . By contrast, indices of real special functions behave obscurely under multiplication. In Bargmann space, the oscillations needed for one-body orthogonality are taken up by the phase, because t is a complex number. Many-body zeros appear as zeros of polynomials constructed from monomials in several variables. The nodal surface is the locus of zeros of a complex polynomial. Such objects have been studied for a long time, and are among the basic building blocks of algebraic geometry [7]. Thus Bargmann's transform allows many-body wave functions to be studied directly by classical methods developed in the nineteenth century for the study of *geometric invariants*: polynomials invariant with respect to various symmetries.

The finer structure, mentioned above, concerns invariants of the Pauli principle, called *shapes* [4]. There is a finite number of them, as described in Section 2. In Section 3, an explicit algorithm is given to generate them. The algorithm is just taking derivatives of polynomials, so it evidently reduces the information content of shape wave functions in each step. This observation raises the question how to quantify the information, because the shapes are pure states, so that all their standard quantum entropies, based on the density matrix, are zero. In Section 4, the problem is solved by applying the concept of logical entropy. The last section places this development in a wider context.

Many-body Hilbert space as a free module

Historical background

The basic idea of algebraic geometry appeared already in antiquity, in relation with the so-called Delian problem. Menaechmos (380–320 BC) was apparently the first to notice that the solution of the algebraic equation $x^3 = 2$ could be constructed geometrically as the intersection of two parabolas, $y = x^2$ and $2x = y^2$. The idea is that an unknown algebraic object can be presented as an intersection of known geometric objects.

One can embed Menaechmos' intersection of parabolas into an infinite family of polynomials:

$$R(x, y) = P \cdot (x^2 - y) + Q \cdot (2x - y^2). \quad (3)$$

When P and Q are c -numbers, this construction is a vector space. If P and Q are arbitrary polynomials in (x, y) , the two parabolas are not independent, because setting $P = 2x - y^2$ and $Q = y - x^2$ gives $R = 0$, so that P and Q are not even unique for any given R . Most interesting are intermediate cases, when P and Q are “sensibly restricted,” say by requiring that they be symmetric in x and y : $P(x, y) = P(y, x)$, and similarly for Q . In that case, it is said that the two quadratic forms generate a *free module* over the ring of symmetric polynomials. Notably, the two parabolas are not themselves members of this ring. Algebraically, they could be independent members of any set, which is why the module is called free. (As the example shows, freedom is affected by the choice of coefficient ring, which defines linear independence.) If the module is defined by a property, instead of by generators, then the question arises of the number of generators necessary to obtain all elements which possess the desired property.

In 19th century, a significant body of research focused on symmetry as such a defining property. For example [8]: how many *independent* rotational invariants can be constructed out of N vectors \mathbf{v}_i in d dimensions? For $N \leq d$, it turns out that they are precisely the $N(N+1)/2$ scalar products $x_{ij} = \mathbf{v}_i \cdot \mathbf{v}_j$, $1 \leq i < j \leq N$. Therefore, taken as independent variables, the x_{ij} generate a polynomial ring of all possible rotational invariants.

The case $N > d$ presents a generic complication, of which three vectors in two dimensions are the simplest example. The angles between them must satisfy $\alpha + \beta = \gamma$, which implies relations connecting the scalar products, so the number of independent rotational invariants is smaller than $N(N+1)/2$. Such relations are called *syzygies*. They are an important practical obstacle to generating only independent invariants in concrete situations, and we shall see below that they also appear in the fermion many-body problem.

In 1891, Hilbert achieved a historical breakthrough in the study of invariants [9]. He showed that all complex-polynomial invariants of classical and finite symmetry groups were finitely generated, i.e. could be written in the form of a free module,

$$\sum_{i=1}^D h_i g_i, \quad (4)$$

with a finite number D of generators g_i , and h_i belonging to a polynomial ring. Because wave functions in Bargmann space are complex polynomials, and the Pauli principle is a symmetry, we can make direct use of this result in the construction of many-fermion wave functions.

Two-fermion wave functions

Take two fermions in three dimensions. Staying with the oscillator, the position coordinates in real space are the arguments of six Hermite functions, which are Bargmann-transformed to powers of six complex variables (t_i, u_i, v_i) with $i = 1, 2$. This arrangement as two points in \mathbb{C}^3 is physically natural. However, it should be kept in mind that real-space intuition does not necessarily transfer to wave-function space directly. For example, it must be checked what the angular momentum operators look like in Bargmann space. Luckily they look the same as in real space, so that solid harmonics in Bargmann space can be obtained from those in real space simply by replacing Cartesian coordinates (x, y, z) with (t, u, v) [10]. In particular, the triplet (t_i, u_i, v_i) transforms as a vector, which supports the above grouping of variables into points from a physical viewpoint, marking the first step in developing a geometric visualization of wave-function space.

Denoting $t = t_1 - t_2$ etc., one can write the most general antisymmetric wave function of two fermions in three dimensions by inspection:

$$\Psi(t, u, v) = P \cdot t + Q \cdot u + R \cdot v + S \cdot tuv, \quad (5)$$

where the question remains, how to restrict the polynomials P, Q, R, S . Evidently they must be symmetric under particle exchange $1 \leftrightarrow 2$. To keep the generators independent (the module free), they must not contain terms linear in t, u , or v . Physically, it is also sensible to restrict them to relative coordinates. Thus we arrive at,

$$\Psi(t, u, v) = P(t^2, u^2, v^2) \cdot t + Q(t^2, u^2, v^2) \cdot u + R(t^2, u^2, v^2) \cdot v + S(t^2, u^2, v^2) \cdot tuv. \quad (6)$$

We have constructed Ψ as a member of a free module, finitely generated by (t, u, v, tuv) . The latter quartet are generating invariants of the Pauli principle, just as the scalar products generate all invariants of the orthogonal group. The rotationally invariant formulation, including center-of-mass excitations, is described in Ref. [10].

Wait, one could ask, isn't $tuv = t \cdot u \cdot v$ a syzygy? Actually it would be, and the basis (t, u, v) would suffice, if we were to allow the coefficients to be *any* symmetric polynomial, such as uv , making $tuv = uv \cdot t$ linear in t . Our restriction to squares has excluded that, but it seems ad hoc. There is a general and quite natural motivation which has the same effect [4].

The symmetric-function coefficients are interpreted physically as bosonic excitations. This interpretation is natural in Bargmann space, because powers of the variables count excitation quanta.¹ By the same token, the generators are vacuum states. It is a physical requirement for bosonic excitations that they can be quantized in a manner analogous to the electromagnetic field. It means that excitations factorize across the space directions in such a way that they are symmetric under particle exchange in each direction *separately*: a plane wave is a boson, and all waves can be expanded in plane waves. This restriction to bosonic excitations naturally implies the one above, because the lowest-degree symmetric functions of relative coordinates in any one direction are t^2, u^2 , and v^2 . The polynomial coefficients are thus superpositions of physical bosonic excitations $t^{2k}u^{2l}v^{2m}$ only. The restriction to bosonic excitations determines the generating invariants. In particular, one can count them in general [4]. There are M^{d-1} generators for N identical particles (fermions or bosons) in d dimensions. These generators or vacuum states are called *shapes*.

Parenthetically, the above factorization requirement applies to bosonic *excitations*, like density waves, not to conserved bosons like helium atoms. The latter have their own shapes in full analogy with the fermion ones [4], but have been much less studied so far.

Geometrically, the $2!^3 - 1 = 4$ generators (t, u, v, tuv) can be grouped into two objects: a vector (t, u, v) and a pseudoscalar tuv . The latter transforms like a signed volume under rotations and reflections in \mathbb{C}^3 , and is in fact the Bargmann transform of the wave function $\psi_1(x)\psi_1(y)\psi_1(z) \sim xyz$ in \mathbb{R}^3 . All excitations of two fermions in three dimensions can be classified according to whether they contain the generator tuv or not. In this way, a kinematic reason is found for the existence of bands in spectra of finite systems. The spectrum of two fermions has two bands because there are exactly two distinct rotationally invariant ways to comply with the Pauli principle [10].

Algorithmic construction of shapes

The algorithm

A general constructive algorithm to obtain all shape wave functions has been published in Ref. [4]. It is valid for fermions and bosons in all dimensions. However, it is neither transparent nor efficient. For fermions in three (and all odd) dimensions, the following algorithm has been proven to generate *only* shapes [11], however the conjecture that it generates *all* shapes is still unproven. All statements referring to the entropy of shapes in the latter part of this article refer to shapes generated by this algorithm, thus, strictly speaking, to the so-defined subset of shapes. The experimental evidence, including some counting results in very large spaces, nevertheless indicates that the conjecture is most likely true.

¹ In real space, only real nodes count excitations, and a symmetric function multiplying an antisymmetric one does not necessarily introduce new ones.

The algorithm finds generators of three-dimensional many-fermion Hilbert space over the space of bosonic excitations as symmetrized derivatives of the triple product (*source shape*),

$$\mathcal{D}_N = \Delta_N(t)\Delta_N(u)\Delta_N(v), \quad (7)$$

where $\Delta_N(t)$ is the Vandermonde form in N variables t_1, \dots, t_N , similarly for u and v :

$$\Delta_N(t) = \begin{vmatrix} t_1^{N-1} & \dots & t_N^{N-1} \\ \vdots & & \vdots \\ t_1 & \dots & t_N \\ 1 & \dots & 1 \end{vmatrix} = \prod_{1 \leq i < j \leq N} t_i - t_j. \quad (8)$$

The Vandermonde form is the Bargmann-space version of the ground-state Slater determinant in one dimension. The algorithm is: multiply three ground-state Slater determinants, one in each direction, and find all distinct symmetrized derivatives. For large N , this is easier said than done [11], but that issue is outside the scope of this paper.

The source shape is the unique shape of highest possible degree [12]. The symmetrized derivatives are iterates of the operator,

$$\nabla^{(a,b,c)} = \sum_{i=1}^N \frac{\partial^a}{\partial t_i^a} \frac{\partial^b}{\partial u_i^b} \frac{\partial^c}{\partial v_i^c} =: (T^a U^b V^c), \quad (9)$$

where the parenthesis notation is easier to use in concrete expressions. A single parenthesis is called a *word*, iterating words creates a *sentence*, naturally commutative. Notably, the majority of shapes cannot be obtained with a single word, as elaborated below.

The case $N = 2$ is easily seen to conform to this scheme. We have,

$$\mathcal{D}_2 = \begin{vmatrix} t_1 & t_2 \\ 1 & 1 \end{vmatrix} \begin{vmatrix} u_1 & u_2 \\ 1 & 1 \end{vmatrix} \begin{vmatrix} v_1 & v_2 \\ 1 & 1 \end{vmatrix} = (t_1 - t_2)(u_1 - u_2)(v_1 - v_2). \quad (10)$$

The non-zero derivatives are,

$$\nabla^{(1,1,0)} \mathcal{D}_2 = (TU) \mathcal{D}_2 = 2(v_1 - v_2), \quad (11)$$

and cyclically, accounting for the four shapes guessed above.

One-letter words always give zero, $(T^k) \mathcal{D}_N = 0$, because they act on a single Vandermonde form, in which each variable appears in one column. For any determinant, priming and summing the entries column-wise is the same as row-wise, e.g.,

$$\begin{vmatrix} a'_{11} & a_{12} \\ a'_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & a'_{12} \\ a_{21} & a'_{22} \end{vmatrix} = \begin{vmatrix} a'_{11} & a'_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} \\ a'_{21} & a'_{22} \end{vmatrix}.$$

For a Vandermonde form, the right-hand terms have either equal rows or a zero row, so the result follows. In particular, there are no fermion shapes of degree one less than \mathcal{D}_N .

The case of three fermions

For three fermions, there are $6^2 = 36$ shapes, shown as a lattice connected by derivatives in Figure 1. The generating function for fermion shapes [4] reads in this case,

$$q^9 + 3q^7 + 7q^6 + 6q^5 + 6q^4 + 10q^3 + 3q^2, \quad (12)$$

the last term $3q^2$ referring to the textbook triplet ground state of degree 2, i.e.,

$$\begin{vmatrix} t_1 & t_2 & t_3 \\ u_1 & u_2 & u_3 \\ 1 & 1 & 1 \end{vmatrix} = t_{12}u_{23} - u_{12}t_{23}, \quad (13)$$

and cyclically, where t_{12} is shorthand for $t_1 - t_2$ and so on. While the textbook expression is on the left, the right-hand form arises naturally when taking derivatives of \mathcal{D}_3 . In the figure, the source shape \mathcal{D}_3 of degree $3N(N-1)/2 = 9$ is the white disk at the top. The remaining shapes are shown in levels according to their degree. Shapes accessible by a single word are shown as red disks, with the corresponding derivatives (words) depicted by red arrows. For example, the three

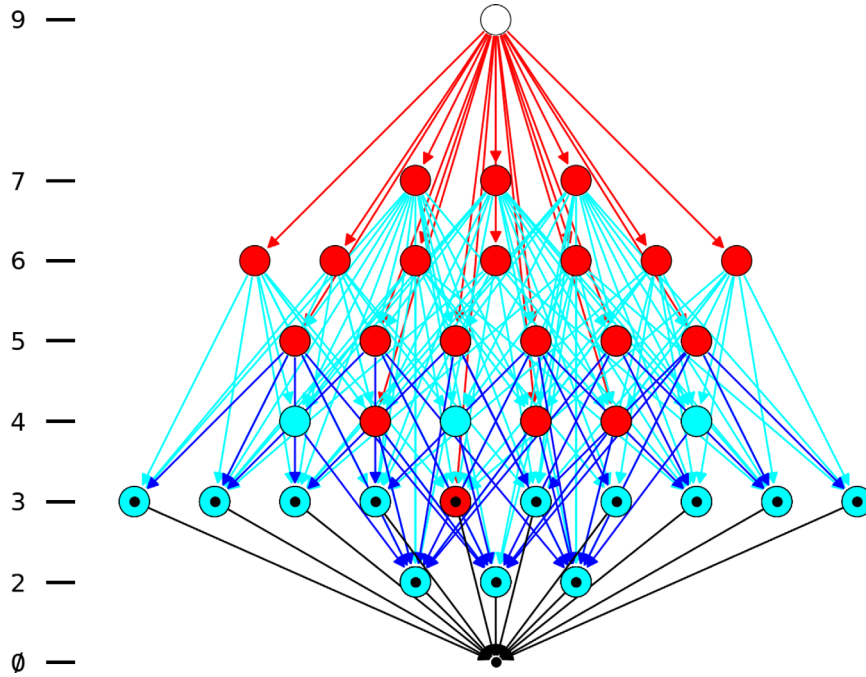


Figure 1. Lattice of shapes for $N = 3$, graded by degree. The top white dot is the state \mathcal{D}_3 , cf. equation (7). Each arrow represents a symmetrized derivative (word) $(T^a U^b V^c)$, cf. equation (9). The word labels are omitted from the arrows for graphical reasons. The bottom black dot is the zero polynomial (not the constant 1). For further details, see the text.

shapes of degree 7 are respectively $(TU)\mathcal{D}_3$, $(TV)\mathcal{D}_3$, and $(UV)\mathcal{D}_3$. As a further example, the red arrow pointing to the single red disk of degree 3 is the word $(T^2 U^2 V^2)$, which means that this red disk is the state,

$$(T^2 U^2 V^2)\mathcal{D}_3 \sim t_{12}u_{12}v_{12} - t_{13}u_{13}v_{13} + t_{23}u_{23}v_{23}. \tag{14}$$

This example shows why no other single word can give a shape of degree 3. It would involve a single derivative higher than 2, e.g. $(T^3 U^2 V)$, but 2 is the highest power appearing in the Vandermonde forms for $N = 3$. The shapes which can be reached with two words, but not one, are shown in cyan, which exhausts all the shapes for $N = 3$. The second words are depicted by cyan arrows, some of which give alternative paths to the red disks. Some of the red and cyan disks can also be reached with three words, where the third is depicted by a blue arrow. Finally, black dots mark shapes on which *any* symmetrized derivative gives zero, which case is depicted by black arrows connecting them to a black dot, symbolizing the zero polynomial, which naturally becomes the bottom of the lattice.

The sharp-eyed reader will have noticed that the first two rows of red disks all have only one arrow pointing at them, while several point at each of the six disks in the third row, of degree $3N(N - 1)/2 - 4 = 5$. These are syzygies among the sentences. They appear as soon as the total degree of the sentence exceeds N , i.e. the total degree of shapes is less than $3N(N - 1)/2 - N$. In other words, if N were equal to 4 (or greater), all eighteen sentences of total degree 4,

$$\begin{aligned} &(T^3 U), (T^2 U^2), (TU)^2, (TU^3), (T^3 V), (T^2 UV), (TV)(TU), (TU^2 V), (UV)(TU), \\ &(U^3 V), (T^2 V^2), (TV)^2, (TUV^2), (UV)(TV), (U^2 V^2), (UV)^2, (TV^3), (UV^3), \end{aligned} \tag{15}$$

would give rise to precisely all the 18 distinct shapes of degree $3N(N - 1)/2 - 4$. For $N = 3$, there are not enough variables to distinguish them all. Instead, the six words above which contain a third derivative give zero, while the remaining 12 are “crushed together” to give all 6 distinct non-zero shapes of degree 5 for $N = 3$, each repeated twice, e.g. $(T^2 U^2)\mathcal{D}_3 \sim (TU)^2\mathcal{D}_3$. Indeed, magnifying the figure, one can count a total of 15 arrowheads pointing to the six disks in the third row. The extra three come from the three two-word sentences, like $(TV)(TU)$, which generate distinct paths upon commutation.

This “crushing together” is similar to what happens between scalar products when the space dimension is too low. It reflects a physical fact: when the number of particles is fixed, several long-enough deexcitation sequences can lead to the same state, which is an algebraic version of the fermion sign problem [13]. Further considerations along these lines can be found in Ref. [11].

The observation relevant for this work is that derivatives of polynomials obviously reduce their information content [12]. Therefore, the information content of all pure states is not the same, and some way should be found to quantify it.

Logical entropy applied to shapes

Identification of Pauli distinctions

Logical entropy was extensively described in a recent review [14]. Unlike the usual information entropy, which counts the number of bits necessary to distinguish members of some set, logical entropy counts the distinctions themselves. Like all entropies, it has a microscopic and a coarse-grained (probabilistic) realization. Only the former is of concern here.

In the logical-entropy approach, for any given partition of a set, each pair of elements belonging to different subsets in the partition is called a *distinction*. The logical entropy is just the total number of such pairs (distinctions). For example, if the set $\mathcal{U} = \{a, b, c, d\}$ is partitioned into two subsets $\mathcal{U}_1 = \{a, b\}$ and $\mathcal{U}_2 = \{c, d\}$, the total number of distinctions made by this partition is 4: (a, c) , (a, d) , (b, c) , (b, d) . Hence the logical entropy of the partition is $h(\{\mathcal{U}_1, \mathcal{U}_2\}) = 4$. Notably, the pairs are counted only once.

For a given set, the largest logical entropy is obtained by partitioning it into one-element subsets (*singletons*). It is equal to the number of distinct-element pairs, i.e. $W(W - 1)/2$ if W is the cardinality of the initial set. In the above example, the singleton partition would be $\mathcal{W} = \{\{a\}, \{b\}, \{c\}, \{d\}\}$. The logical entropy of this partition is the cardinality of its direct-product square without the diagonal $\mathcal{I} = \{(a, a), (b, b), (c, c), (d, d)\}$, divided by two, because (a, b) is the same distinction as (b, a) :

$$h(\mathcal{W}) = \text{card}((\mathcal{W} \times \mathcal{W}) \setminus \mathcal{I})/2 = 6. \quad (16)$$

The logical entropy of the singleton partition corresponds to the *microscopic* entropy in standard statistical physics, with which it shares the property that it is fixed by the underlying set (\equiv microscopic phase space), so it is maximal and cannot change. The logical entropy of coarser partitions, like $\{\mathcal{U}_1, \mathcal{U}_2\}$ above, similarly corresponds to the coarse-grained (relevant) entropies in physics, which are defined by probability distributions over the microscopic states. These increase towards the singleton (microscopic) value as the partitions become finer (the distributions more uniform).

The defining expression for the Vandermonde form (8) is a product of differences of all pairs of variables: if there are N variables, there are $N(N - 1)/2$ such differences. In the language of logical entropy, the number of terms in the Vandermonde product counts the number of distinct pairs of variables. This insight naturally provides the idea how to ascribe logical entropy to the highest-degree quantum shape \mathcal{D}_N : count the total number of factors in it, so $h(\mathcal{D}_N) = 3N(N - 1)/2$ in three dimensions. Taking derivatives naturally reduces the number of factors, so it is immediately obvious that the information content measured by the logical entropy will be reduced. The key idea is that distinctions of variables (t_i, t_j) , i.e. pairs of distinct elements of the set of variables, are encoded as algebraic differences $(t_i - t_j)$ of the same variables in the wave function.

While the operationalization of this idea requires some technical finesse, it should be kept in mind that there is nothing contrived about the phenomenon which it seeks to quantify. It has been rigorously proven that the polynomial \mathcal{D}_N is a shape [12], and that all its symmetrized derivatives are shapes [11]. It is an algebraic truism that taking derivatives of a polynomial reduces its information content, all the way down to the zero polynomial. Thus the proposition that all wave functions do not have the same information content is on firm mathematical ground, even though the particular measure of that content, proposed here, need not be the only one possible.

Encoding distinctions as differences has the obvious advantage that one can test for their presence algebraically. This observation gives a tentative definition: a many-body shape S *distinguishes* two one-body wave functions t_i and t_j if $a \neq b \Rightarrow S(t_i = a, t_j = b) \equiv -S(t_i = b, t_j = a) \neq 0$. The definition tests for the presence of a factor $t_i - t_j$ in the expression for S . It is limited to shapes because no more is needed, as shown below. In the general case, one must find a way to extend it to distinctions made by derivatives of \mathcal{D}_N , which are not fully factorizable. Indeed, the Pauli principle does not require antisymmetry in any given space coordinate, unlike the case for waves, discussed above.

There are three subtleties to consider now. First, our definition allows us in principle to distinguish directions in space as well, for example: if exchanging $t_i \leftrightarrow u_i$ changes the sign of the shape, evidently it has some orbital symmetry. We must limit ourselves only to those distinctions which test the Pauli principle.

Second, we are really looking for distinctions in excess of the minimum required by the Pauli principle, which is that a wave function changes sign when *triplets* (t_i, u_i, v_i) and (t_j, u_j, v_j) are exchanged. Thus all shapes make $N(N - 1)/2$ distinctions (the number of pairs of triplets for N particles), which could even be subtracted from whatever logical entropy they have, to find the entropy relative to that baseline. Notably, multiplying the shapes with any symmetric function cannot change the number of distinctions according to our definition. Thus, limiting the definition to shapes is not really a limitation, but a recognition that only shapes make distinctions.

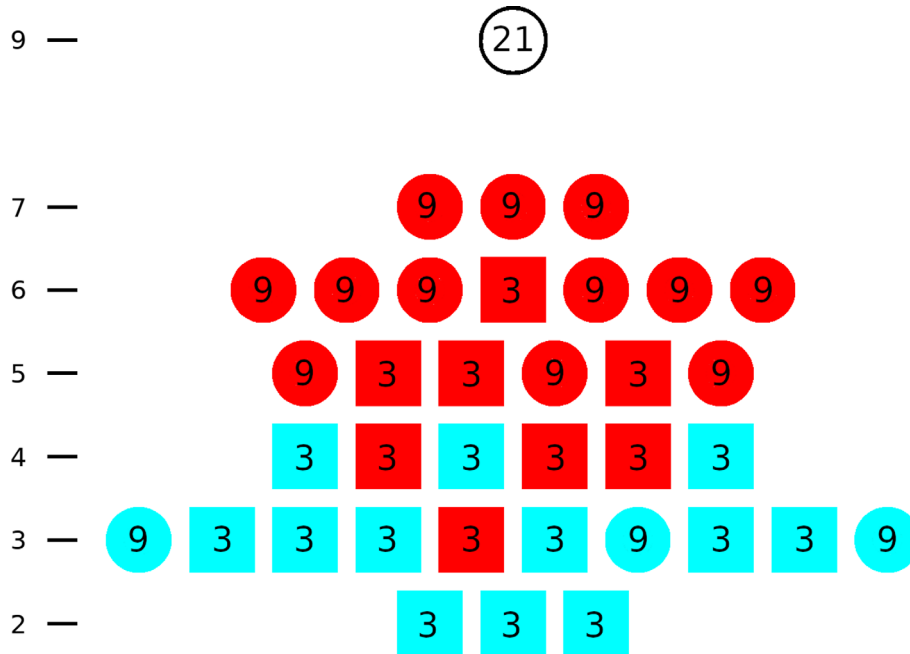


Figure 2. Values of the logical entropies of the states in Figure 1, in the same lattice arrangement and color-coding. White circle: $h = 21$. Colored circles: $h = 9$. Colored squares: $h = 3$.

The shape \mathcal{D}_N implements the Pauli principle redundantly, because it is the product of three one-dimensional fermionic ground states, so it is antisymmetric in each direction by itself. In addition to the triplet distinctions, it makes $3N(N - 1)/2$ distinctions of single variables. When it is subject to the operator (TU) (say), the shape $(TU)\mathcal{D}_N$ will make both the $N(N - 1)/2$ *singlet distinctions* coming from the Vandermonde form $\Delta_N(v)$, which was not touched by this operator, and the $N(N - 1)/2$ *triplet distinctions* like all others. The appearance of different levels of distinctions – singlet and triplet – at the same time is the third subtlety. Once it is noticed, there is no reason not to consider *doublet* distinctions as well. They do not change the sign, because two transpositions give an even permutation. Thus we are led to the following refined definition.

A shape S is said to *make a (Pauli) distinction* if the algebraic identity,

$$S \equiv \text{sign}(\pi)\pi(S), \tag{17}$$

holds TRUE when $\pi(S)$ is a *permutation testing the Pauli principle*, which acts on S by exchanging its variables in one of the following ways, with prescribed $\text{sign}(\pi)$:

1. a triplet exchange, $(t_i, u_i, v_i) \leftrightarrow (t_j, u_j, v_j)$ for some pair of particle indices (i, j) , and $\text{sign}(\pi) = -1$;
2. a doublet exchange, $(a_i, b_i) \leftrightarrow (a_j, b_j)$ for some pair of particle indices (i, j) , where (a, b) is any of (t, u) , (t, v) , (u, v) , and $\text{sign}(\pi) = 1$;
3. a singlet exchange, $a_i \leftrightarrow a_j$ for some pair of particle indices (i, j) , where a is any of t, u, v , and $\text{sign}(\pi) = -1$.

The (Pauli) logical entropy of S , denoted $h(S)$, is defined to be the total number of distinctions made by S according to the above definition. For example, \mathcal{D}_3 makes 21 distinctions: 3 triplet, 9 doublet, and 9 singlet. The next shape $(TU)\mathcal{D}_3$ makes the standard 3 triplet distinctions, but it contains the Vandermonde form $\Delta_3(v)$ still untouched, for another 3 doublet and 3 singlet distinctions, or 9 in all. The logical entropies of all states in Figure 1 are shown in Figure 2. Note, for example, the single $h = 3$ state among the seven of degree 6. It is the state $(TUV)\mathcal{D}_3$, in which taking derivatives has spoiled the singlet (and doublet) distinctions in all three variables. The remaining six are variants of the form $(T^2U)\mathcal{D}_3$, in which one of the three Vandermonde determinants remains untouched. In a similar vein, one Vandermonde determinant remains untouched in the three $h = 9$ states among the ten of degree 3. They are $(TU)^3\mathcal{D}_3 \sim \Delta_N(v)$ and cyclically, so that the three 1D ground states are distinguished by logical entropy among the ten first-excited 3D shapes.

A simple consequence of the above definition is that a wave function which is a superposition of shapes with symmetric-function (boson-excitation) coefficients will have the same logical entropy as the shape component with the lowest logical entropy. This observation extends the definition, originally enunciated for pure shapes, to all wave functions.

Identification of the Pauli distinction sets

In the previous section, logical entropy is computed by a counting algorithm. The original definition [14] is by the cardinality of a set of distinctions (*ditset*). In order to show that the implementation is consistent with the definition, one must identify the ditset whose elements are counted by the algorithm. This issue is addressed now, separately for the different levels first.

The wave function is used as a test on a set of possible distinctions, and a distinction is counted if the test returns TRUE. By assumption, the wave function is a fermion wave function, which means that it encodes particle indistinguishability according to the Pauli principle: any permutation of particle indices will map the wave function onto itself up to the sign of the permutation.

For singlets, the counted set is a union of three ditsets, each generated as the distinctions of a partition into singletons $\{t_i\}$ etc. for each direction t , u , and v separately. This identification follows directly from the product form of the source shape \mathcal{D}_N , where each singlet distinction is encoded as a difference of the corresponding variables.

For triplets, the counted set is the ditset of a partition into *vector* singletons, precisely the triplets (t_i, u_i, v_i) . In other words, the underlying partition for the triplet distinctions is also into singletons, but consists of different mathematical objects. Thus we have identified the ditsets of the following partitions so far: three partitions into singletons of N complex points, and one partition into singletons of N complex triplets, so far.

The common geometric thread among these two ditsets is that a triplet defines a (signed) *volume* in \mathbb{C}^3 , while a point defines a *line segment* in \mathbb{C} . By the same token, doublets subtend *surfaces* in \mathbb{C}^2 , three times over, namely for each plane (t, u) , (t, v) , and (u, v) separately.

Thus the level-wise counting implementations of logical entropy are fully microscopic and consistent with the definition [14]. Each is always the maximally resolved one, consisting of singletons. These singletons refer to sets of different objects, rather than different partitions of the same set of objects, as one might have assumed. Importantly, it is impossible to count a ditset only partially, because of the indistinguishability of particles encoded in the wave function. If it returns TRUE on one distinction in an assumed ditset, it must return TRUE on all, in effect either admitting or rejecting the ditset as a whole. In this way, consistency of the definition of logical entropy with the filtering algorithm of the previous section is proven separately for singletons, doublets, and triplets.

Finally, to close the issue, it is obviously possible to define the total logical entropy as the sum of logical entropies of segments, surfaces, and volumes, because these are all different objects, so there cannot be double counting by construction. Hence the logical entropy of shapes counts the cardinality of the union of ditsets corresponding to the microscopic (maximal) partitions of the sets of variables whose exchange transforms the wave function into itself by the filtering of equation (17).

Discussion and conclusions

The definition of the logical entropy introduced above is not the only one possible. The singlet distinctions imply the other two kinds, as we have seen for \mathcal{D}_3 . Thinking of this implication as a *refinement* [14], one may well ask, why not always use the level of refinement which gives the maximal entropy? One cannot argue with a definition, of course, and such use, effectively replacing our *sum* over the sets of singlet, doublet, and triplet entropies with a *maximum*, would also be consistent. However, it turns out already on this small example that this choice flattens the histogram in Figure 2 significantly, assigning some shapes the same entropy which have different ones in the figure. Clearly, barring some special reason, one would prefer the more sensitive definition, in favor of the present choice.

The present formulation ascribes a *shape entropy* to pure quantum states, assigning them an intrinsic information content *a priori*. Indeed, Figure 1 can be interpreted as decoding a message. All the information about shapes is in the source shape \mathcal{D}_N , because taking derivatives is a deterministic algorithm. Nevertheless, there is practical value in carrying out the algorithm, just as there is in decompressing a file. For example, one discovers possible deexcitation sequences connecting the disks in the figure. Importantly, all states discovered in this way are easily proven to be pure shapes [11], so no selection among them is necessary, which would add information along the way. Conversely, to obtain the fully factorized form \mathcal{D}_N by iterated integration in multiple variables, starting from the well-known ground-state Slater determinants, requires significant information input in the form of well-chosen integration constants.

By contrast, all current formulations of quantum entropy, found in the literature [15, 16], are *a posteriori*. They assign zero entropy to pure states, while entropy appears only after a measurement, modelled by a projection, as first laid out by von Neumann in terms of the density matrix ρ [17]. This *a posteriori* character pertains both to von Neumann's original quantum entropy $-\text{Tr}(\rho \ln \rho)$ and to the logical quantum entropy [18, 19] defined as $\text{Tr}(\rho - \rho^2)$. Both are zero as long as the density matrix encodes a pure state, i.e. can be diagonalized with all eigenvalues equal to zero or one. The projection associated with a measurement makes the density matrix no longer idempotent, $\rho^2 \neq \rho$, causing both entropy formulas to become non-zero, so they clearly encode the information gain from the measurement.

The *a priori* information content in the shapes, measured by the shape entropy, values intrinsic geometric properties in wave-function space. It finds some wave functions to be more valuable than others, based on their symmetry properties. In this scheme, the most informative wave function is the source shape, which implements the Pauli principle most redundantly. It is proportional to itself not only under the exchange of triplets of variables, as all fermion wave functions must be, but also under the exchange of doublets and singlets. The obvious question is whether any physical meaning can be ascribed to such a valuation. The remainder of this discussion is devoted to a particular example of this kind. The physical issue is robustness, which makes physical states useful for quantum computation.

A robust state is one that does not easily deexcite into lower states. One way to find it is to look for states of low energy which have a different configuration than states of even lower energy, so that both the thermodynamic forces and the Hamiltonian matrix elements along the deexcitation path are small. As already discussed elsewhere [10], the shape wave functions are band-heads in the spectra of finite systems. They fit both these properties, because band-heads are the lowest excited states which are of a different configuration than all lower states, which belong to other bands in the spectrum. For a quantum dot with three polarized electrons, one would therefore look for exceptionally robust states among the ten wave functions of degree 3 in Figure 1.

In order to construct a physically realizable state from them, one must consider rotational invariance next. Shapes either form a rotational multiplet by themselves, or else components of individual shapes are found embedded in rotational multiplets, which include bosonic excitations [10]. Combining the above considerations, a robust rotationally invariant state is expected to be a pure-shape multiplet. Extending the two-fermion analysis [10] to $N = 3$, one finds that the ten third-degree shapes in Figure 1 span two pure shape multiplets, a vector triplet (angular momentum $L = 1$) and septiplet ($L = 3$). Given that the ground state is a vector triplet itself, one would expect the septiplet to be exceptionally robust, because its electromagnetic deexcitation involves a forbidden $\Delta L = 2$ transition. Interestingly, it contains both the $h = 9$ and the $h = 3$ Cartesian-basis states shown in Figure 2. In this way, the high-entropy components are protected from decay by rotational invariance. The internal (projection) states of the septiplet are in principle amenable to further manipulation.

It is worth noting that three polarized electrons in a $3D$ well potential (quantum dot) have a total of $\binom{10}{3} = 120$ states involving the lowest three oscillator shells. The majority of these contain bosonic excitations (including center-of-mass ones), in other words are higher-excited band states, which are expected to deexcite quickly to their band-heads. Evidently, the above identification of seven special states among the 120, comprising the shape septiplet, is not trivial.

To conclude, a particular implementation of logical entropy has been shown to differentiate between pure states. These states are invariants of the Pauli principle, the shapes Ψ_i , which generate the full N -fermion Hilbert space as a free module, according to the formula,

$$\sum_{i=1}^{M^2} \Phi_i \Psi_i, \quad (18)$$

where the Φ_i are polynomials in the one-body wave functions, symmetric in each space direction separately (bosonic excitations). The shapes are geometric objects in wave-function space, and the shape entropy, as defined in this work, can be interpreted as a simple valuation of their basic transformation properties under particle exchange. Hopefully this approach can be leveraged to identify robust few-electron states suitable for quantum computation.

Conflicts of interest

The author declares no conflict of interest.

Acknowledgments

I would like to thank Franjo Sokolić for inviting me to the conference where this work was presented, and David Ellerman for encouraging me to write up the presentation as this paper.

References

1. von Neumann J (1927), Mathematische Begründung der Quantenmechanik. Nachr Ges Wissenschaften Göttingen Math-Phys Klasse 1927, 1–57.
2. Heisenberg W (1926), Mehrkörperproblem und Resonanz in der Quantenmechanik. Zeitschr für Physik 38(6), 411–426.
3. Slater JC (1929), The theory of complex spectra. Phys Rev 34, 1293–1322.
4. Sunko DK (2016), Natural generalization of the ground-state Slater determinant to more than one dimension. Phys Rev A 93, 062109.

5. Ceperley DM (1991), Fermion nodes. *J Stat Phys* 63(5), 1237–1267.
6. Bargmann V (1961), On a Hilbert space of analytic functions and an associated integral transform part I. *Commun Pure Appl Math* 14, 187–214.
7. Milne JS (2015), Algebraic Geometry (v6.01). Available at <https://www.jmilne.org/math/>.
8. Weyl H (1946), The classical groups: their invariants and representations, 2nd edn., Princeton University Press, Princeton.
9. Sturmfels B (2008), Algorithms in invariant theory, 2nd edn., Springer-Verlag, Wien.
10. Rožman K, Sunko DK (2020), Generic example of algebraic bosonisation. *Eur Phys J Plus* 135, 30.
11. Sunko DK (2020), Many-fermion wave functions: structure and examples, in: J. Bonča, S. Kruchinin (Eds.), *Advanced nanomaterials for detection of CBRN in NATO Science for Peace and Security Series A: Chemistry and Biology*, Springer, pp. 85–99.
12. Sunko DK (2017), Fundamental building blocks of strongly correlated wave functions. *J Supercond Nov Magn* 30(1), 35–41.
13. Hirsch JE (1985), Two-dimensional Hubbard model: numerical simulation study. *Phys Rev B* 31, 4403–4419.
14. Ellerman D (2017), Logical information theory: new logical foundations for information theory. *Log J IGPL* 25(5), 806–835.
15. Manfredi G, Feix MR (2000), Entropy and Wigner functions. *Phys Rev E* 62, 4665–4674.
16. Bosyk GM, Zozor S, Holik F, Portesi M, Lamberti PW (2016), A family of generalized quantum entropies: definition and properties. *Quantum Inf Process* 15(8), 3393–3420.
17. von Neumann J (1932), *Mathematical foundations of quantum mechanics*, Princeton University Press, Princeton.
18. Tamir B, Cohen E (2014), Logical entropy for quantum states [arXiv:1412.0616](https://arxiv.org/abs/1412.0616) [quant-ph].
19. Ellerman D (2018), Logical entropy: introduction to classical and quantum logical information theory. *Entropy* 20(9), 679.

Cite this article as: Sunko D.K. 2022. Entropy of pure states: not all wave functions are born equal. 4open, 5, 3.
