Received: 6 September 2022 | Accepted: 18 November 2022

DOI: 10.1002/pamm.202200005

Measure concentration and the Schrödinger equation

Harry Yserentant[1,*]

[1] Technische Universität Berlin, Institut für Mathematik, 10623 Berlin, Germany

This talk pursued the aim to represent the solutions of the electronic Schrödinger equation as traces of higher-dimensional functions. This allows to decouple the electron-electron interaction potential but comes at the price of a degenerate elliptic operator replacing the Laplace operator on the higher-dimensional space. The surprising observation is that this operator can without much loss again be substituted by the Laplace operator, the more successful the larger the system under consideration is. This is due to a nontrivial concentration of measure phenomenon that has much to do with the random projection theorem known from probability theory and can, for example, serve as a building block for the construction of iterative methods that map sums of products of orbitals and geminals onto functions of the same type.

1 Introductory remarks

The electronic Schrödinger equation establishes a connection between chemistry and physics. It describes systems of finitely many electrons that interact with a given, fixed set of nuclei and among each other. The neglect of the motion of the nuclei can be heuristically justified by the fact that the nuclei are much heavier than the electrons and move therefore on a different time scale. A complete mathematical justification is, of course, a considerably more difficult task. For a system of N electrons moving in the field of K nuclei of charge Z_ν clamped at the positions a_ν, the corresponding Hamilton operator reads

$$H \;=\; -\,\frac{1}{2}\sum_{i=1}^{N}\Delta_i \;-\;\sum_{i=1}^{N}\sum_{\nu=1}^{K}\frac{Z_\nu}{\|x_i-a_\nu\|}\;+\;\frac{1}{2}\sum_{i,j=1}^{N}\frac{1}{\|x_i-x_j\|}. \tag{1.1}$$

It acts on functions that depend on the variables $x_1, x_2, \ldots, x_N$ associated with the positions of the electrons in the three-dimensional space. To get rid of the problems with the electron-electron interaction terms, it is an obvious idea, which has already in the early days of quantum mechanics been successfully applied to calculate the ground state of the Helium atom and paved the way for the general acceptance of quantum mechanics in its present form, to introduce the differences of the electron positions as additional variables. To be precise, let

$$m = 3\times N, \quad n = 3\times N + 3\times\frac{(N-1)N}{2}, \tag{1.2}$$

and let the vectors in $\mathbb{R}^m$ and $\mathbb{R}^n$, respectively, be partitioned into subvectors in the position space $\mathbb{R}^3$. Let the subvectors of the vectors in $\mathbb{R}^m$ and the first N of the subvectors of the vectors in $\mathbb{R}^n$ be labeled by the indices $i = 1,\ldots,N$ and the remaining subvectors of the vectors in $\mathbb{R}^n$ by the index pairs (i,j), with components $i,j = 1,\ldots,N$ and $i<j$. We try to represent the solutions of the electronic Schrödinger equation then in the form $u(x) = U(Tx)$, where the matrix T maps the vectors x in $\mathbb{R}^m$ into the vectors $y = Tx$ in $\mathbb{R}^n$ with the subvectors

$$y_i = x_i, \quad y_{ij} = \frac{x_i - x_j}{\sqrt{2}} \tag{1.3}$$

and U is a function from $\mathbb{R}^n$ to $\mathbb{R}$ that can ideally be well approximated, say, by a sum of products of orbitals, functions depending only on the position $y_i = x_i$ of a single electron, and geminals, functions of the variables y_{ij} associated with the differences $x_i - x_j$ of the electron positions. The scaling factor $\sqrt{2}$ is not an absolute must, but will considerably simplify the later presentation. The potential decouples in this new set of variables completely and splits into a sum of terms that depend only on one of the components y_i or y_{ij}. The question is what happens with the first term in the Hamilton operator, which is associated with the kinetic energy of the electrons. The talk tried to answer this question. For background information and references to the literature, see [1] and [2], the papers on which the talk was essentially based. A detailed exposition of the results presented in this short review, including complete proofs, can be found in [3].

2 Trace functions and the Laplace operator

To keep the presentation simple, we are in the following mainly concerned with functions $U:\mathbb{R}^n\to\mathbb{R}$, n a potentially high if not very high dimension, that possess a then also unique representation

$$U(y) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \int \widehat{U}(\omega)\,\mathrm{e}^{\,\mathrm{i}\,\omega\cdot y}\,\mathrm{d}\omega \tag{2.1}$$

* Corresponding author: e-mail yserentant@math.tu-berlin.de

in terms of an integrable function $\widehat{U}$, their Fourier transform. Such functions are by the Riemann-Lebesgue theorem uniformly continuous and vanish at infinity. The space $W_0(\mathbb{R}^n)$ of these functions is under the norm

$$\|U\| = \left(\frac{1}{\sqrt{2\pi}}\right)^n \int |\widehat{U}(\omega)|\,\mathrm{d}\omega \tag{2.2}$$

a Banach space and even a Banach algebra. Let T be a still arbitrary $(n \times m)$-matrix of full rank $m < n$ and let

$$u : \mathbb{R}^m \to \mathbb{R} : x \to U(Tx) \tag{2.3}$$

be the trace of a function in $U \in W_0(\mathbb{R}^n)$. As the functions in $W_0(\mathbb{R}^n)$ are uniformly continuous, the same holds for their traces. As there is a constant c with $\|x\| \le c\,\|Tx\|$, the trace functions (2.3) vanish at infinity, too. The next lemma gives a criterion for the existence of partial derivatives of the trace functions, where we use the common multi-index notation.

Lemma 2.1 *Let U be a function in $W_0(\mathbb{R}^n)$ and let, for α be given and with T^t the transpose of T, the functions*

$$\omega \to (\mathrm{i}\,T^t\omega)^\beta \widehat{U}(\omega), \quad \beta \le \alpha, \tag{2.4}$$

be integrable. The trace function (2.3) *of this function U possesses then the partial derivative*

$$(\mathrm{D}^\alpha u)(x) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \int (\mathrm{i}\,T^t\omega)^\alpha \widehat{U}(\omega)\,\mathrm{e}^{\,\mathrm{i}\,\omega\cdot Tx}\,\mathrm{d}\omega, \tag{2.5}$$

which is like u itself uniformly continuous and vanishes at infinity.

For partial derivatives of order one, this can be shown with the help of the dominated convergence theorem applied to the corresponding difference quotients. For partial derivatives of higher order, the proposition follows by induction.

Let $W_0^2(T)$ be the space of the functions $U \in W_0(\mathbb{R}^n)$ with finite (semi)-norm

$$|U|_T = \left(\frac{1}{\sqrt{2\pi}}\right)^n \int \|T^t\omega\|^2\,|\widehat{U}(\omega)|\,\mathrm{d}\omega, \tag{2.6}$$

where $\|\cdot\|$ denotes again the Euclidean norm. The traces of these functions are by Lemma 2.1 twice continuously differentiable. Let $\mathcal{L} : W_0^2(T) \to W_0(\mathbb{R}^n)$ be the degenerate elliptic differential operator given by

$$(\mathcal{L}U)(y) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \int \|T^t\omega\|^2\,\widehat{U}(\omega)\,\mathrm{e}^{\,\mathrm{i}\,\omega\cdot y}\,\mathrm{d}\omega. \tag{2.7}$$

For the functions $U \in W_0^2(T)$ and their traces (2.3), by Lemma 2.1

$$-(\Delta u)(x) = (\mathcal{L}U)(Tx) \tag{2.8}$$

holds. That is, $-\Delta u$ is itself the trace of a higher-dimensional function $\mathcal{L}U$. The matrix T^t maps the $\mathbb{R}^n$ into the lower-dimensional, and in the case of the matrix given by (1.3) much lower-dimensional $\mathbb{R}^m$. The dimension $n - m$ of its kernel is in such cases much higher than the dimension m of its range. The more surprising is the fact that the Euclidean norm of $T^t\omega$ is, for the matrix T assigned to the Schrödinger equation, on most of the $\mathbb{R}^n$ almost equal to the Euclidean norm of ω itself. The fraction of the vectors ω on the unit sphere of the $\mathbb{R}^n$ for which the Euclidean norm of $T^t\omega$ differs from one by more than a given small amount tends exponentially to zero as the number of electrons goes to infinity. This is due to a nontrivial concentration of measure phenomenon, which has a lot to do with the random projection theorem (see Lemma 5.3.2 in [4], for example) from probability theory. It means that the operator (2.7) behaves more or less like the negative Laplace operator

$$-(\Delta U)(y) = \left(\frac{1}{\sqrt{2\pi}}\right)^n \int \|\omega\|^2\,\widehat{U}(\omega)\,\mathrm{e}^{\,\mathrm{i}\,\omega\cdot y}\,\mathrm{d}\omega \tag{2.9}$$

applied to functions U defined on the higher-dimensional space.

3 The underlying measure concentration effect

Let $n > m$ and let A be a real $(m \times n)$-matrix of rank m. The kernel of such a matrix has the dimension $n - m$ and hence can be a large subspace of the $\mathbb{R}^n$. Nevertheless, the set of all x for which

$$\|Ax\| \ge \delta\,\|A\|\|x\| \tag{3.1}$$

holds fills, in the high-dimensional case, often almost the complete $\mathbb{R}^n$ once δ falls below a certain bound; the norms are as in the previous section the Euclidean norm on the $\mathbb{R}^m$ and the $\mathbb{R}^n$, respectively, and the matrix norm is the assigned spectral norm. To describe this effect in more detail, we introduce on the $\mathbb{R}^n$ the probability measure

$$\lambda(M) = \frac{1}{n\nu_n} \int_{M \cap S^{n-1}} \mathrm{d}\eta, \tag{3.2}$$

where ν_n is the volume of the n-dimensional unit ball and $n\nu_n$ is the area of the unit sphere S^{n-1}. We apply it to the sectors consisting of the $x \in \mathbb{R}^n$ for which $\|Ax\| < \delta \|A\| \|x\|$ holds to measure their opening angle. For orthogonal projections, matrices with one as the only singular value, the measure of these sectors possesses a closed integral representation.

Theorem 3.1 *Let the $(m \times n)$-matrix P, $m < n$, be an orthogonal projection. Then*

$$\lambda\big(\{x \,|\, \|Px\| < \delta \|x\|\}\big) = F(\delta), \quad 0 \le \delta < 1, \tag{3.3}$$

holds, where the function $F(\delta) = F(m, n; \delta)$ is defined by the integral expression

$$F(\delta) = \frac{2\,\Gamma(n/2)}{\Gamma(m/2)\Gamma((n-m)/2)} \int_0^\delta (1-t^2)^\alpha t^{m-1}\,\mathrm{d}t \tag{3.4}$$

and the exponent $\alpha \ge -1/2$ is given by

$$\alpha = \frac{n-m-2}{2}. \tag{3.5}$$

It takes nonnegative values for dimensions $n \ge m+2$.

If the difference of the dimensions n and m is even, the function (3.4) is for even m an even and for odd m an odd polynomial of degree $n-2$ in δ. A closed representation of these polynomials is given in [2]. For practical purposes, it is, however, more advantageous not to rely on such representations and to evaluate $F(\delta)$ numerically by means of a quadrature rule. As F takes the value $F(1) = 1$, $F(\delta) = F(\delta)/F(1)$ holds, so that there is no need to evaluate the Gamma function.

The function (3.4) always represents a lower bound for the area ratios under consideration.

Theorem 3.2 *Let A be a nonvanishing matrix of dimension $m \times n$, $m < n$. Then one has*

$$\lambda\big(\{x \,|\, \|Ax\| < \delta \|A\| \|x\|\}\big) \;\ge\; F(m, n; \delta). \tag{3.6}$$

Upper bounds for the area ratios depend in general on the singular values of the matrix, in the extreme case on its condition number, the ratio of its maximum and its minimum singular value. This is fortunately not the case for the matrices $A = T^t$ assigned to the Schrödinger equation. The Euclidean norm of the vector $Tx \in \mathbb{R}^n$ is given by

$$\|Tx\|^2 = \sum_{i=1}^N \|x_i\|^2 + \frac{1}{4} \sum_{i=1}^N \sum_{j=1}^N \|x_i - x_j\|^2 \tag{3.7}$$

or, after rearrangement, with the rank three map $T_0 x = x_1 + x_2 + \cdots + x_N$ by

$$\|Tx\|^2 = \frac{N+2}{2} \|x\|^2 - \frac{1}{2} \|T_0 x\|^2. \tag{3.8}$$

The $(m \times m)$-matrix $T^t T$ has therefore only two distinct eigenvalues, the eigenvalue 1 of multiplicity $3 \times 1 = 3$ and the eigenvalue $(N+2)/2$ of multiplicity $3 \times (N-1)$, that is, $m-3$. The singular values of the matrix T^t are therefore

$$\sigma_i = 1 \text{ for } i \le 3, \quad \sigma_i = \sqrt{\frac{N+2}{2}} \text{ for } i \ge 4. \tag{3.9}$$

The spectral norm of the matrix T^t is $\sqrt{(N+2)/2}$.

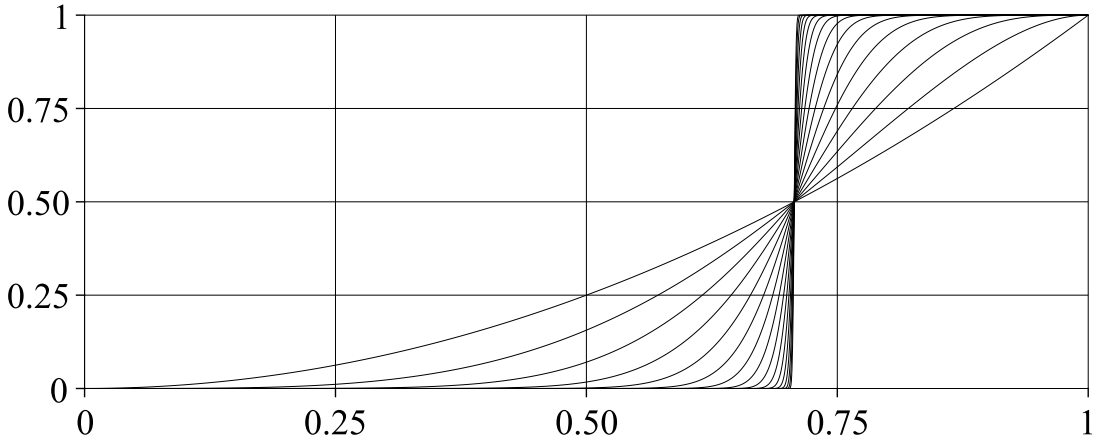

Fig. 1: The functions $F(\delta) = F(m, n; \delta)$ from equation (3.4) for the dimensions $m = 2^k$, $k = 1, \ldots, 16$, and $n = 2m$

Theorem 3.3 *Let A be a matrix of dimension $m \times n$, $m < n$, with singular values $\sigma_k = \sigma_m$ for $k > m_0$. The corresponding area ratios then satisfy the estimate*

$$\lambda\big(\{x \,|\, \|Ax\| < \delta \, \|A\| \|x\|\}\big) \leq F(m - m_0, n; \delta). \tag{3.10}$$

The next theorem describes the limit behavior of the function (3.4) when the dimensions increase and tend to infinity. The subsequent estimates are expressed in terms of the function

$$\phi(\vartheta) = \vartheta \, \exp\left(\frac{1 - \vartheta^2}{2}\right). \tag{3.11}$$

It increases on the interval $0 \leq \vartheta \leq 1$ strictly, attains at the point $\vartheta = 1$ its maximum value one, and decreases from there again strictly to its limit value zero.

Theorem 3.4 *Let ξ be the square root of the dimension ratio m/n. For $\delta < \xi$ and $\xi < \delta$, respectively, then one has*

$$0 \, \leq \, F(m, n; \delta) \leq \, \phi\left(\frac{\delta}{\xi}\right)^m, \quad 0 \, \leq \, 1 - F(m, n; \delta) \leq \, \phi\left(\frac{\delta}{\xi}\right)^m. \tag{3.12}$$

If the dimension ratio $\delta_0^2 = m/n$ is kept fixed or only tends to δ_0^2, the functions (3.4) thus tend to a step function with jump discontinuity at δ_0. Figure 1 reflects this behavior. We summarize our findings therefore once more as follows and relate them to the prospective jump positions.

Theorem 3.5 *Let A be a nonvanishing matrix of dimension $m \times n$, $m < n$, and let ξ be the square root of the dimension ratio m/n. For $\vartheta > 1$, then one has*

$$\lambda\big(\{x \,|\, \|Ax\| \geq \vartheta\xi \|A\| \|x\|\}\big) \leq \phi(\vartheta)^m. \tag{3.13}$$

The theorem states in particular that the norm of Ax exceeds the value $\xi\|A\|\|x\|$ by more than a moderate factor $\vartheta > 1$ only on a very small, de facto negligibly sector, an observation that is of great importance for the analysis of iterative methods. Under the much more restrictive assumptions from Theorem 3.3, Theorem 3.5 possesses a counterpart for values $\vartheta < 1$.

Theorem 3.6 *Let $n > m$ and let A be a nonvanishing $(m \times n)$-matrix with singular values $\sigma_k = \sigma_m$ for $k > m_0$. If $m' = m - m_0$ and ξ' is the square root of m'/n, then*

$$\lambda\big(\{x \,|\, \|Ax\| < \vartheta\xi' \|A\| \|x\|\}\big) \leq \phi(\vartheta)^{m'} \tag{3.14}$$

holds for all ϑ in the interval $0 < \vartheta < 1$.

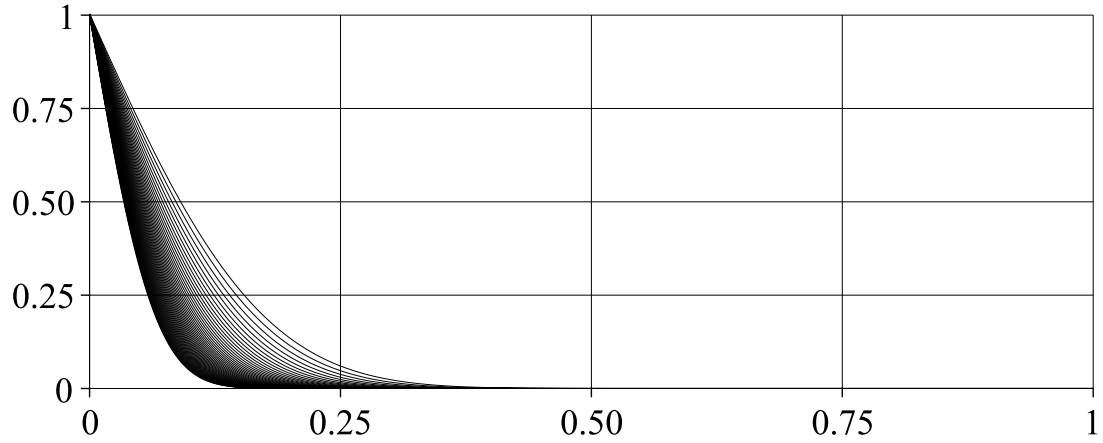

Fig. 2: Upper bounds for the probability that the condition (4.1) is violated as function of ε for $N = 8, \ldots, 64$ electrons.

4 Back to the Laplace operator

The last two theorems apply, because of (3.9), to the matrices $A = T^t$ assigned to the Schrödinger equation. They form the basis of our argumentation. Let $0 < \varepsilon < 1$. For a randomly chosen vector ω in the frequency space, the probability that

$$(1-\varepsilon)\xi' \, \|T^t\| \|\omega\| \le \|T^t\omega\| < (1+\varepsilon)\xi \, \|T^t\| \|\omega\| \tag{4.1}$$

holds is by these two theorems, because of $\phi(1 \pm \varepsilon) < \exp(-c\,\varepsilon^2)$ with $c = -\ln(\phi(2))$, at least

$$1 - 2\exp(-c\,\varepsilon^2 m'), \tag{4.2}$$

where the dimensions (1.2) and the exponent $m' = 3 \times (N-1)$ depend on the number N of particles, the quantities ξ and ξ' are the square roots of the dimension ratio m/n and of m'/n, and the constants

$$\xi' \, \|T^t\| = \sqrt{1 - \frac{2}{N(N+1)}}, \quad \xi \, \|T^t\| = \sqrt{1 + \frac{1}{N+1}} \tag{4.3}$$

enclose the value one and tend to one as N goes to infinity. The fraction of the vectors ω on the unit sphere of the $\mathbb{R}^n$ for which the Euclidean norm of $T^t\omega$ differs from one by more than a given small amount thus tends as claimed exponentially to zero as the number of electrons goes to infinity. A quantitatively significantly better lower bound than (4.2) for the probability that (4.1) holds can be directly derived from Theorem 3.2 and Theorem 3.3. In terms of the function (3.4), it reads

$$F(m, n; (1+\varepsilon)\xi) - F(m', n; (1-\varepsilon)\xi') \tag{4.4}$$

and deviates for increasing particle number less and less from the probability

$$F(m, n; (1+\varepsilon)\xi) - F(m, n; (1-\varepsilon)\xi) \tag{4.5}$$

that an orthogonal projection from the $\mathbb{R}^n$ to the $\mathbb{R}^m$ maps a randomly chosen unit vector to a vector of length between $(1-\varepsilon)\xi$ and $(1+\varepsilon)\xi$. For some small to medium size systems, Fig. 2 shows the resulting upper bounds for the probability that the condition (4.1) is violated. Since the products of the matrix T^t with vectors $e \in \mathbb{R}^n$ in the three-dimensional subspaces assigned to the particle positions and their differences have the norm $\|T^t e\| = \|e\|$, they satisfy the condition

$$\xi' \, \|T^t\| \|e\| < \|T^t e\| < \xi \, \|T^t\| \|e\| \tag{4.6}$$

and thus, independent of ε, the condition (4.1). This establishes a link to hyperbolic cross spaces and thereby indirectly also to tensor product approximations, and not least to the mixed regularity of electronic wave functions [5]. The inequality tells us that the Fourier transforms of functions in the corresponding hyperbolic cross spaces are inherently concentrated in the subregions of the frequency space on which the norm of $T^t\omega$ does not much differ from that of ω.

What does all this mean? For high electron numbers, at the latest when statistical physics comes into play, the norm of $T^t\omega$ is almost equal to that of ω for all ω outside of a tiny, negligible sector. This allows to replace the operator (2.7) in such cases by the negative Laplace operator (2.9) and the original Hamilton operator in the corresponding sense by a decoupled Hamilton operator acting upon the higher-dimensional functions. But also for moderate particle numbers, the negative Laplace operator (2.9) remains a good approximation to the operator (2.7), surely good enough to serve as a building block for the construction of rapidly convergent iterative methods to determine the lowermost eigenvalues of the Hamilton operator (1.1).

Acknowledgements Open access funding enabled and organized by Projekt DEAL.

References

[1] H. Yserentant, Numer. Math. **146**, 219–238 (2020).
[2] H. Yserentant, SIAM J. Matrix Anal. Appl. **43**, 464–478 (2022).
[3] H. Yserentant, The Laplace operator, measure concentration, Gauss functions, and quantum mechanics, https://arxiv.org/abs/2208.03957.
[4] R. Vershynin, High-Dimensional Probability (Cambridge University Press, Cambridge, UK, 2018).
[5] H. Yserentant, ESAIM: M2AN **45**, 803–824 (2011).

