Using simulated annealing algorithms to solve the Schrödinger equation in muonic atoms

I.S. Kardaras[1] and O.T. Kosmas[2]

[1] Theoretical Physics Section, University of Ioannina, GR-45110 Ioannina, Greece
[2] Chair of Applied Dynamics,University of Erlangen-Nuremberg, Germany

E-mail: iokard@uoi.gr, odysseas.kosmas@ltd.uni-erlangen.de

Abstract. In many physical problems, the computation of exact wave functions for muons (particles about two hundred times heavier than electrons), bound in the extended Coulomb field created by the atomic nucleus, is required. Even though the problem is trivial under the assumption of point-like nuclear systems, the consideration of the nuclear finite-size necessitates the use of advantageous numerical techniques. In the case of non-relativistic bound muons, the solution of the Schrödinger equation is reliable, but for a relativistic description the solution of the Dirac equations for the bound muon is needed. In the present contribution, as a first step, we attempt to derive a method for solving the Schrödinger equation on the basis of simulated annealing algorithms. To this end, one may optimize appropriate parametric expressions for the wave function of a muon orbiting around complex nuclei by employing the simulated annealing method recently constructed to minimize multi parametric expressions in several physical applications.

1. Motivation

As is well known, when negative muons, μ^-, produced in a meson factory (e.g. at the Fermilab, USA, or at the J-PARC, Japan), slow down in matter, it is possible for them to be captured in atomic orbits (in this way muonic atoms are produced). Afterwards, fast electromagnetic cascades bring the muon of the muonic atom down to the innermost (1s or 2p) quantum orbits [1]. A bound in such an orbit muon may disappear either by decay known as muon decay in orbit or by capture by the nucleus. The main reaction-channel of the muon capture is the ordinary muon capture represented by the reaction [1]

$$\mu_b^- + (A, Z) \rightarrow (A, Z-1)^* + \nu_\mu \,, \tag{1}$$

where (A, Z) denotes the initial atomic nucleus with mass number A and proton number Z (A and Z are considered integer numbers). Even though processes like the reaction (1) have been the subject of extensive experimental and theoretical investigations started on the early 50's, recently the interest has been revived due to the important role they play in physical applications and astrophysics. In these studies mostly a mean value of the muon wave function, $\Phi_\mu(r)$, with r being the spherical coordinate, has been utilized. However, for a reliable description of the reaction (1) and of any reaction having the same initial state with it, i.e. a muon orbiting around an atomic nucleus (A, Z), the derivation of an exact muon wave function is necessary.

It is the purpose of our present work to construct an advantageous method that would provide an accurate muon wave functions by solving the Schrödinger equation (or the Dirac equations)

which obeys a muon bound in the extended Coulomb field of the nucleus of such a muonic atom. Our method is exploiting the advantages of the simulated annealing algorithms [2].

2. Brief description of the Formalism

In several approaches the description of the muon capture process (1) is based on non-relativistic muons [1, 3]. Then, one needs to solve the Schrödinger equation which gives the radial part of the muon wave function $\Phi_\mu(r)$, or the known as the reduced radial wave function $U(r)$ defined as

$$U(r) = r\Phi_\mu(r)\,. \tag{2}$$

In the latter case, the Schrödinger equation is written as [4]

$$-\frac{\hbar^2}{2\mu}\frac{d^2}{dr^2}U(r) + V(r)U(r) = EU(r)\,. \tag{3}$$

where E is the total energy of the muon (energy eigenvalue). In the latter equation μ denotes the reduced mass of the muon-nucleus system given by

$$\frac{1}{\mu} = \frac{1}{m_\mu} + \frac{1}{Zm_p + (A-Z)m_n}, \tag{4}$$

where m_μ is the muon rest mass and m_p and m_n stand for the mass of the proton and neutron, respectively. In (3) $V(r)$ represents the extended Coulomb potential originating from the finite-size nuclear charge-density distribution $\rho(r)$ which for various spherically symmetric nuclei is available from electron scattering data [4]. For a point-like nucleus, Eq. (3) can be solved trivially. However, by taking into consideration the finite size of the nucleus, $V(r)$ is obtained from the expression

$$V(r) = -e^2\int\frac{\rho(r')}{|r-r'|}d^3r'\,. \tag{5}$$

(e denotes the electric charge of the proton) Then, in order to solve the Schrödinger equation (3) one needs to apply special numerical integration techniques [4, 5, 6].

By multiplying (3) by $U(r)$ and integrating we obtain

$$-\frac{\hbar^2}{2m}\int U(r)\frac{d^2}{dr^2}U(r)dr + \int V(r)U(r)^2dr = E\int U(r)^2dr. \tag{6}$$

Integrating by parts the term with the second derivative (using boundary conditions the term $U(r)dU(r)/dr$ vanishes) and solving for the energy E the latter equation, we find

$$E = N_0\int_0^\infty[\frac{\hbar^2}{2\mu}(\frac{dU(r)}{dr})^2 + V(r)U^2(r)]dr, \tag{7}$$

where N_0 is the normalization factor

$$N_0 = \left[\int_0^\infty U^2(r)dr\right]^{-1}. \tag{8}$$

In solving the second order differential equation (3), Eq. (7) acts as a constraint.

We mention that $U(r)$ obeys the boundary condition $U(r=0)=0$ and has asymptotic behavior $U(r) \sim e^{-kr}$, with $k > 0$ (usually $0.1 \leq k \leq 0.6$). Equation (3) has been solved previously by assuming that $U(r)$ can be written as a superposition of appropriately parametrized sigmoid functions as explained below [4].

3. Solving the Schrödinger equation with artificial neural networks

At first we derive a grid from $r = 0$ up to a point $r = b$ (b is dependent on the nuclear system in question, $30 \leq b \leq 100fm$) where the wave function practically vanishes, and denote it by r_j, for $j = 1, 2, \ldots, n$. Then (3) must hold at every point r_j of the grid. This is equivalently expressed as

$$\sum_{j=1}^{n} \left[\left(-\frac{\hbar^2}{2\mu} \frac{d^2}{dr^2} + V(r_j) - E \right) U(r_j) \right]^2 = 0 \tag{9}$$

The artificial neural networks approach, chosen to solve the latter equation, consists in parameterizing $U(r)$ and then minimizing the left-hand side of (9) divided by the normalization $\int U(r)^2 dr$, so as to avoid the trivial solution $U(r) = 0$ everywhere [7]. To this aim we use the parametrization

$$U(r) = re^{-kr} N(r, \mathbf{u}, \mathbf{v}, \mathbf{w}), \qquad k > 0, \tag{10}$$

where $N(r, \mathbf{u}, \mathbf{v}, \mathbf{w})$ is a feed-forward artificial neural network with one hidden layer and one input unit (r). The biases are denoted by $\mathbf{u} = (u_1, u_2, \ldots, u_m)$, where m is the number of hidden units. The weights to the hidden layers are denoted by $\mathbf{w} = (w_1, w_2, \ldots, w_m)$ and the weights to the output by $\mathbf{v} = (v_1, v_2, \ldots, v_m)$. The hidden layer units have sigmoid activations of the form $f(x) = (1 + e^{-x})^{-1}$. Thus, the artificial neural network is defined as

$$N(r, \mathbf{u}, \mathbf{v}, \mathbf{w}) = \sum_{i=1}^{m} u_i f(w_i r + v_i). \tag{11}$$

To obtain the precise expression for the reduced radial wave function $U(r)$, we insert this form in equation (10). By combining the latter equation with (9) we get

$$\mathcal{C} = \sum_{j=1}^{n} \left[\left(-\frac{\hbar^2}{2\mu} \frac{d^2}{dr^2} + V(r_j) - E) \right) \left(r_j e^{-kr_j} \sum_{i=1}^{m} u_i f(w_{ji} r_j + v_i) \right) \right]^2 = 0. \tag{12}$$

In the present paper, inspired by the simulated annealing method of Ref. [2], we will derive an algorithm to optimize the latter expression which would play the role of a total cost function $\mathcal{C}$.

4. Optimization using simulated annealing techniques

We consider that the initial radial part of the μ^- wave function $U(r)$ is divided into several line segments, the number of which depends on the nuclear system in question. In this way, any radial wave function may be represented by a set of way-points $U_i = U(r_i)$. The total cost of the initial radial wave function $U(r)$ is readily calculated [see Eq. (12)]. Then, at any iteration step, every way-point of the wave function $U(r)$ is moved by an elementary length perpendicular to the line which connects the initial and the end points of the segment. This elementary length is specified by the Coulomb potential and the nuclear system (region in which the nuclear charge density $\rho(r)$ is, practically, non-zero). Every displacement has a positive or negative contribution to the total cost and it is accepted even if it has a positive contribution to the total cost with a probability, however, which depends on a parameter T (the temperature-parameter of the algorithm) that plays the same role as the temperature of a physical system. In simulated annealing methods the probability distribution is described by [2]

$$p(E) \equiv Prob(E) \approx exp(-E/T) \tag{13}$$

(the so-called Boltzmann probability distribution) where the temperature T and energy E are measured in MeV). The physical meaning of the latter equation is that a system in thermal

equilibrium at temperature T has its energy states probabilistically distributed according to their energy E. Even at low temperature, there is a chance for the system to get out of a local energy minimum in favor of finding a better, more global energy minimum. In other words, the system sometimes goes uphill as well as downhill, but the lower the temperature, the less likelihood for any significant uphill transition. Initially, the temperature-parameter T of the algorithm is high, but as the algorithm proceeds, the temperature is approaching to zero. At this point, only movements with negative contributions are accepted. This method, known as simulated annealing, is used to prevent the algorithm from being trapped in local minima through the above mechanism it proceeds further to find the global minimum (or minima) [2].

4.1. Description of the code for the minimization process

As mentioned before, simulated annealing mimics the annealing process to solve an optimization problem. The temperature parameter T [see Eq. (13)] controls the search which typically starts off high and is slowly "cooled" or lowered in every iteration. At each iteration a new point is generated and its distance from the current point is proportional to the temperature. If the new point has a better function value it replaces the current point and iteration counter is incremented. It is possible, however, to accept and move forward with a worse point, but the probability of doing so is directly dependent on the temperature T (this step, sometimes, helps identify a new search region in hope of finding a better minimum and protects the algorithm from being trapped in local minima).

Our objective function $\mathcal{C}$ is parameterized by the arguments $\mathbf{u}, \mathbf{v}, \mathbf{w}$ that act as constants during the optimization. These parameters can change to create a family of objective functions (they are not varied as part of the minimization). We should note that, there is no systematic way to decide if the calculated radial wave function $U(r)$ is the optimal one. However, in our cases the energy E is very critical and, thus, the optimal wave function $U(r)$ is close to that corresponding to the minimum energy. We are currently working on the derivation and checks of a MATLAB code for the minimization process and we expect to obtain results for concrete examples soon.

5. Summary and Conclusions

Exact bound-muon wave functions can be computed within the context of simulated annealing algorithms by solving the Schrödinger equation and can be compared with those obtained by previous numerical methods. Afterwards, improved wave functions for relativistic muons, coming out of similar solutions of the two-component Dirac equation could be obtained in a straightforward way by extending our present method.

References

[1] Measday D F 2001 Phys. Rep. **354** 243.
[2] Kosmas O T and Vlachos D 2012 Computers & Operations Research **39** 576.
[3] Suzuki T, Measday D F, and Roalsvig J P 1987 Phys. Rev. **C 35** 2212.
[4] Kosmas T S and Lagaris I E 2002 J. Phys. **G 28** 2907.
[5] Zinner N T, Laganke K, Vogel P Phys. Rev. **C 74** 024326, and References therein.
[6] Kardaras I S 2010 MSc. Thesis, Univ. of Ioannina.
[7] Tsoulos I G 2008 Appl. Math. and Comput. **203** 598.