Chemical Physics 173 (1993) 377–383
North-Holland

Neural network solution of the Schrödinger equation for a two-dimensional harmonic oscillator

J. Androsiuk, L. Kułak and K. Sienicki

Institute of Technical Physics and Applied Mathematics, Technical University of Gdańsk, Majakowskiego 11/12, 80-952 Gdańsk, Poland

Received 3 December 1992

We present computer simulations of a neural network capable of learning to perform transformations generated by the Schrödinger equation required to find eigenenergies of a two-dimensional harmonic oscillator. We show that this task can be achieved by a not fully connected back-propagation neural network containing 49 input neurons, 3 hidden layer neurons and 1 output neuron. The investigated neural network turns out to be capable of predicting eigenenergies with an average error of less than one percent. We demonstrate that the CPU time required to teach a neural network of performing the transformation produced by the Schrödinger equation is about 2 min to reach 41000 learning iterations, thus making foreseeable a direct application of a neural network in this and other more complex physical and chemical problems. A discussion of the errors due to the generalization of acquired knowledge is presented and related to a limited number of examples in learning mode and the number of neurons in the hidden layer. Decreasing the number of neurons in the hidden layer increases the apparent ability of the neural network for generalization.

1. Introduction

The last decade has seen a remarkable upsurge of interest in neural networks and neural computing [1–3]. The interest varies from theoretical studies to pragmatic implementations of neural networks. At the present state, neural networks, neural computations and computing are very promising methods of investigating complex multidimensional scientific and pragmatic issues. The neural network is a highly parallel dynamical system with the topology of a direct graph that can carry out information processing by means of its state response to a continuous or initial input [4]. At present it appears that neural networks are particularly suited for studies of complex physical and chemical processes. From another perspective, physical and chemical processes can be analyzed using the methods of the complexity theory. From our standpoint the complexity theory is meaningful as it raises the question of whether or not phenomena that are known to be algorithmic are actually algorithmic in a useful way [5]. Although there is no indisputable answer, contemporary numerical methods of quantum chemistry may be categorized as NP problems. Owing to intrinsic features, neural networks can be used to tackle computational difficulties of physics and chemistry. Though the theoretical background of neural networks is in some extent established [1–4], numerical simulations of neural networks require some ingenuity in creating neural structures and learning rules which outperform other possible neural structures in their learning mode [6].

Although neural networks are capable of performing complex tasks, their implementation in physical and chemical research requires at this preliminary stage a number of basic studies to provide a foundation for investigating more perplexing questions. In this perspective Darsey et al. [7] presented results concerning solutions of the Schrödinger equation for some model potential energy functions employing neural network simulations. They investigated one of the most simple solutions of the Schrödinger equation for a two-dimensional harmonic oscillator. It was shown that their neural network can forecast eigenvalues of test cases with an average error of a few percent. Certainly, this result is a very interesting one. It provides the neural network the usefulness to learning the transformation produced by the Schrödinger equation. However, approximately 12 h on a VAX 2000 workstation in learning mode to simulate the transformation of the Schrödinger equation for this very simple task raises the question of the usefulness of the method.

In this paper we present similar studies of a neural network solution of the Schrödinger equation for a two-dimensional harmonic oscillator. We show that the time required to teach a neural network to perform the transformation produced by the Schrödinger equation can be *much shorter*, thus making foreseeable a direct application of neural networks in this and more complex problems.

2. Neural network

Searching literature one encounters a broad spectrum of various types of neural networks that were designed for particular applications [1-3]. In this work our interest is in teaching a neural network to perform the transformation produced by the Schrödinger equation, and for this purpose the back-propagation network (BPN) appears to be the best one [1-3]. To simulate the BPN we used a threshold sigmoidal function of the form

$$f(h)=(1+e^{-2\beta h})^{-1}\,, \tag{1}$$

which satisfies the condition $f(h)+f(-h)\equiv 1$. In eq. (1) the parameter β has the meaning of an inverse of temperature, $T=\beta^{-1}$, by the formal analogy of this equation to the Fermi function. The BPN is depicted in fig. 1 and the equations describing the state of the network are given by

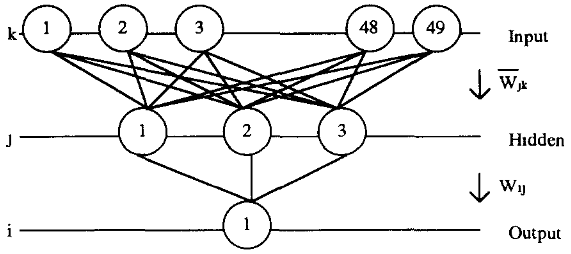

Fig. 1. Schematic architecture of a feed-forward neural network with one hidden layer of neurons. In the simulations presented in this work we used 49 input neurons, 3 hidden layer neurons and 1 output neuron.

$$S_i=f(h_i),\quad h_i=\sum_j w_{ij}s_j-\vartheta_i\,, \tag{2}$$

$$s_j=f(\bar h_j),\quad \bar h_j=\sum_k \bar w_{jk}\sigma_k-\bar\vartheta_j\,. \tag{3}$$

In the learning mode we used the generalized delta rule [1] which is a typical learning procedure utilizing "data-results" pairs. After the neural network is fed with the data, it generates the response and compares it with the requested exact values. During the learning process, the weights of neural connections are changed to minimalize the difference between values expected σ_i^k and generated by the network ζ_i^μ. The learning process is repeated until the error defined by the function

$$\Delta(w_{ij},\vartheta_i,\bar w_{jk},\bar\vartheta_j)=\tfrac{1}{2}\sum_\mu\sum_i\,[\zeta_i^\mu-f(h_i^\mu)]^2 \tag{4}$$

becomes as small as possible. In eq. (4) $\mu=1, ..., p$ and p is the number of cases used in training the network. Thus we are interested in finding a global minimum of the multidimensional function $\Delta(\ )$ given by eq. (4). While simulating the neural network we used the gradient method to evaluate the above function. For this purpose the gradient of $\Delta(\ )$, with respect to every parameter, was calculated and used in implementing the computer code.

3. Theoretical preliminaries

In this work we wish to show that a neural network can learn the transformations produced by the Schrödinger equation. Therefore we felt that for the sake of awareness we should briefly describe the theoretical steps which are necessary to solve the Schrödinger equation for a two-dimensional harmonic oscillator. We are interested in the solution of the stationary Schrödinger equation, where the Hamilton operator is given by a two-dimensional potential energy

$$V=\tfrac{1}{2}m(\omega_x^2x^2+\omega_y^2y^2)\,, \tag{5}$$

where ω_x and ω_y are the oscillator frequencies.

Separating variables, $\Psi(x,y)=\varphi_1(x)\varphi_2(y)$, in the Schrödinger equation and introducing new variables $x_1=x\sqrt{m\omega_x/\hbar}$ and $x_2=y\sqrt{m\omega_y/\hbar}$ one easily obtains

$$\frac{d^2\varphi_1}{dx_1^2}+(\varepsilon_1-x_1^2)\varphi_1=0\,, \tag{6}$$

$$\frac{d^2\varphi_2}{dx_2^2} + (\varepsilon_2 - x_2^2)\varphi_2 = 0 , \qquad (7)$$

where $\varepsilon_1 = 2E_1/\hbar\omega_x$ and $\varepsilon_2 = 2E_2/\hbar\omega_y$ and $E_1 + E_2 = E$. The analytical solution of eqs. (6) and (7) can be obtained in the well-known form

$$\varphi_1(x) = (m\omega_x/\hbar)^{1/4}(2^{n_1}n_1!\sqrt{\pi})^{-1/2} \times\exp(-m\omega_x x^2/2\hbar)\, H_{n_1}(x\sqrt{m\omega_x/\hbar}) , \qquad (8)$$

$$\varphi_2(y) = (m\omega_y/\hbar)^{1/4}(2^{n_2}n_2!\sqrt{\pi})^{-1/2} \times\exp(-m\omega_y y^2/2\hbar)\, H_{n_2}(y\sqrt{m\omega_y/\hbar}) , \qquad (9)$$

where $H(\ )$ is the Hermite polynomial. From eqs. (6) and (7) one can get $\varepsilon_1 = 2n_1 + 1$ and $\varepsilon_2 = 2n_2 + 1$ and the eigenenergies are given by

$$E = \hbar\omega_x(n_1 + \tfrac{1}{2}) + \hbar\omega_y(n_2 + \tfrac{1}{2}) . \qquad (10)$$

Although the obtained solution is a textbook example we have shown it to illustrate the analytical steps required to find the solution of the Schrödinger equation. In the following analysis we are interesting in finding the ground state eigenenergy of a two-dimensional harmonic oscillator. Therefore the quantum numbers $n_1 = n_2 \equiv 0$ and the eigenenergy is given by

$$E = \tfrac{1}{2}\hbar(\omega_x + \omega_y) . \qquad (11)$$

For simplicity we have also assumed that $\hbar \equiv 1$. Now we shall show that a neural network can learn the above transformation generated by the Schrödinger equation.

4. Results and discussion

In order to find the best neural network we analysed two types of networks; the fully connected BPN and the not fully connected BPN. The obtained results show that the not fully connected BPN outperforms the fully connected BPN. For the same number of iterations the error in predicting the solution of the Schrödinger equation was approximately two times higher for the fully connected BPN. Therefore we shall only show and discuss the results obtained for the not fully connected BPN. The network described in section 2 was trained using the following parameters: x, $y \in \langle -3, 3\rangle$ with the increment $+1$, $\omega_x \in \langle 0.4, 0.9\rangle$, $\omega_y \in \langle 1.0, 1.9\rangle$ with the increments 0.1 and $\omega_x = \omega_y = 1$. We performed simulations on a PC 486/50 MHz computer and the CPU time required to reach 41000 iterations was less than two minutes. After the network was trained with the above parameters we generated eigenenergies for different values ω_x and ω_y as depicted in table 1 for several numbers of learning iterations. One may notice that

Table 1
Results of predicting eigenenergy values (E) and the relative error (%) for several values of oscillator frequencies (ω_x, ω_y) and four different numbers of neural network learning iterations. The network was a not fully connected BPN containing 49 input neurons, 3 hidden layer neurons and 1 output neuron

ω_x	ω_y	16230		29462		35797		40600	
		E	%	E	%	E	%	E	%
0.60	1.00	0.807	0.88	0.807	0.88	0.807	0.88	0.807	0.88
0.70	1.00	0.855	0.59	0.856	0.71	0.856	0.71	0.856	0.71
0.80	1.00	0.910	1.11	0.910	1.11	0.910	1.11	0.905	0.56
0.90	1.00	0.970	2.10	0.966	1.68	0.966	1.68	0.966	1.68
0.40	1.00	0.734	4.86	0.733	4.71	0.731	4.43	0.717	2.43
0.40	1.40	0.922	2.44	0.921	2.33	0.919	2.11	0.914	1.56
0.50	1.45	0.981	0.62	0.981	0.62	0.980	0.51	0.980	0.51
0.55	1.40	0.974	0.10	0.974	0.10	0.973	0.21	0.974	0.10
0.60	1.90	1.253	0.24	1.245	0.40	1.243	0.56	1.247	0.24
0.95	1.90	1.407	1.26	1.412	0.91	1.411	0.98	1.419	0.42
0.50	1.90	1.217	1.42	1.205	0.42	1.202	0.17	1.202	0.17
0.60	1.90	1.253	0.24	1.245	0.40	1.243	0.56	1.247	0.24
0.70	1.90	1.294	0.46	1.290	0.77	1.289	0.85	1.298	0.15
0.80	1.90	1.338	0.88	1.339	0.81	1.338	0.88	1.338	0.88
0.90	1.90	1.384	1.14	1.388	0.86	1.387	0.93	1.392	0.57

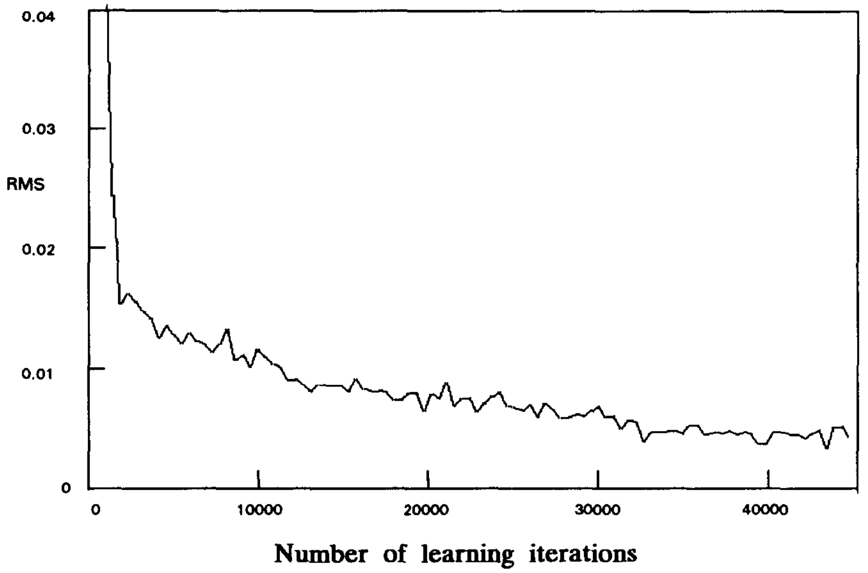

Fig. 2. The root-mean-square normalized error (rms) as a function of number of iterations of a feed-forward back-propagation neural network. The network contains 49 input neurons, 3 hidden layer neurons and one output neuron.

the error in predicting the eigenenergy decreases with the number of learning iterations. One should bear in mind that the CPU time for predicting the eigenenergy is instantaneous in comparison to the CPU in the learning mode. Fig. 2 illustrates the changes of the root-mean-square normalized error (rms) as a function of number of the learning iterations. This error is decreasing rapidly. However, above 10^3 iterations the network is reaching its best performance and the RMS is slowly converging with an increased number of iterations. These simulations show that the designed BPN is very efficient in predicting the eigenenergy of a two-dimensional harmonic oscillator. Moreover, our network outperforms the network described in ref. [7]. Firstly, comparing CPU times, our network is approximately 360 times faster. We expect that the CPU time of the simulated network would be even shorter should we have used a VAX 2000 workstation, as used in ref. [7]. Secondly, the BPN used by us is also more simple than the respective network of Darsey et al. [7]. It contains 49 input neurons, 3 hidden layer neurons and 1 output neuron in comparison to Darsey's et al. [7] network which contains 40 hidden layer neurons. Thirdly, contrary to Darsey et al. [7], our network was not fully connected. We also simulated several BPNs with different numbers of hidden neurons and observed that the network with 3 hidden layer neurons is the finest in its performance for this particular application.

For the sake of comparison we present in table 2 the results of the predicted eigenenergy values obtained by us and Darsey et al. [7]. It is noticed that the RMS obtained by our network is superior. It is

Table 2
Comparison of the results of neural network simulations of eigenenergies (E) of two-dimensional harmonic oscillator frequencies (ω_x, ω_y). The results for 4501 iterations were taken from ref. [7]

ω_x	ω_y	4501		40666	
		E	%	E	%
0.60	1.00	0.839	4.84	0.807	0.88
0.70	1.00	0.872	2.62	0.856	0.71
0.80	1.00	0.895	0.63	0.905	0.56
0.90	1.00	0.907	4.51	0.966	1.68
0.40	1.40	0.879	2.30	0.914	1.56
0.50	1.45	0.967	0.79	0.980	0.51
0.55	1.40	0.973	0.19	0.974	0.10
0.60	1.90	1.256	0.51	1.247	0.24
0.95	1.90	1.418	0.46	1.419	0.42
0.50	1.90	1.185	1.22	1.202	0.17
0.70	1.90	1.305	0.39	1.298	0.15
0.80	1.90	1.353	0.19	1.338	0.88
0.90	1.90	1.390	0.73	1.392	0.57

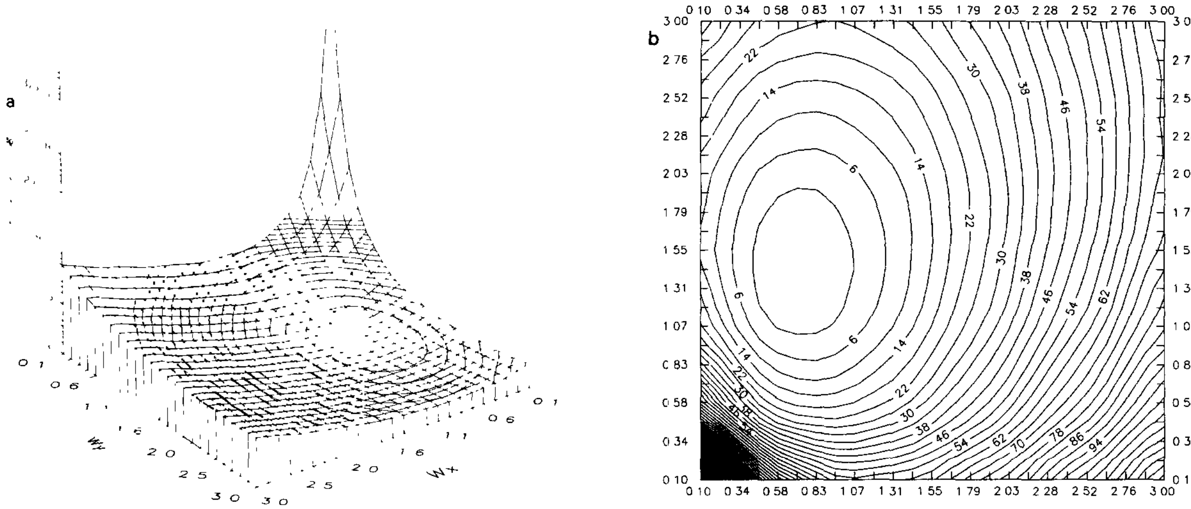

Fig. 3. The relative error (%) in predicting the eigenenergy of a two-dimensional harmonic oscillator for different values of oscillator frequencies ω_x and ω_y. The results were obtained for a fully connected BPN containing 49 input neurons, 40 hidden layer neurons and 1 output neuron after 2×10^5 learning iterations. Note that due to difficulties with the drawing program the maximum relative errors are in opposite corners. In fact they should be in the same corner of the $\{\omega_x, \omega_y\}$ surface.

very illustrative to portray the overall performance of the neural network using three-dimensional plots. In figs. 3 and 4 the error in predicting the eigenenergy of a two-dimensional harmonic oscillator for different values of oscillator frequencies ω_x and ω_y is depicted. Fig. 3 shows the respective results obtained by us for 2×10^5 learning iterations, while simulating the fully connected BPN described by Darsey et al. [7] (49 input neurons, 40 hidden neurons and 1 output neuron). Fig. 4 depicts similar results for 2×10^5 learning iterations, but obtained for a not fully connected BPN containing 49 input neurons, 3 hidden

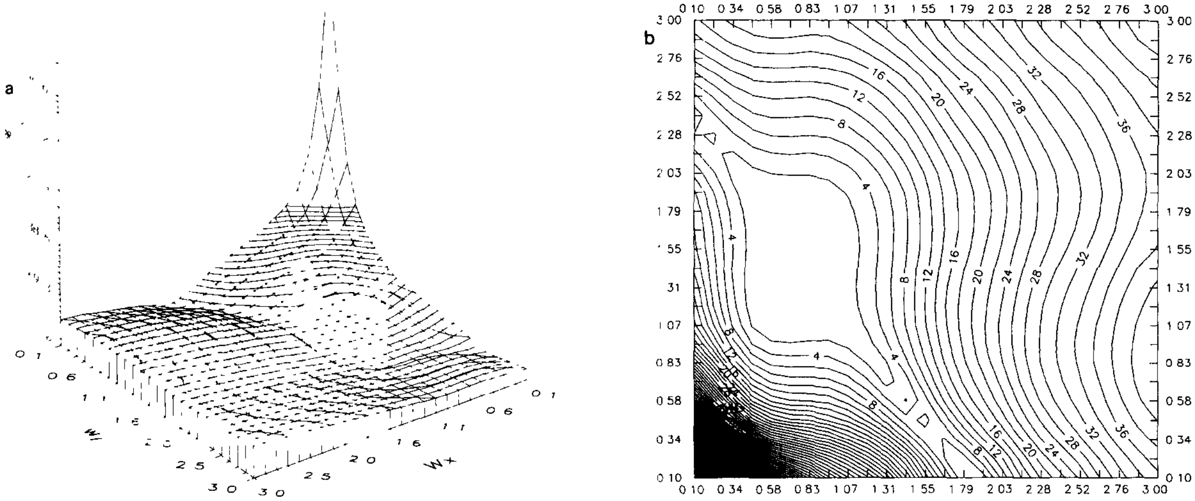

Fig. 4. The relative error (%) in predicting the eigenenergy of a two-dimensional harmonic oscillator for different values of oscillator frequencies ω_x and ω_y. The results were obtained for a fully connected BPN containing 49 input neurons, 3 hidden layer neurons and 1 output neuron after 2×10^5 learning iterations. Note that due to difficulties with the drawing program the maximum relative errors are in opposite corners. In fact they should be in the same corner of the $\{\omega_x, \omega_v\}$ surface.

layer neurons and 1 output neuron. Comparing these figures one notices several interesting features. The eigenenergy predications for both networks have an attractor basin. However, the network with three hidden layer neurons shows a narrow attraction channel across the $\{\omega_x, \omega_y\}$ surface, contrary to a network with 40 hidden neurons where a shallow attractor appears. One should also notice that the not fully connected BPN with three hidden neurons outperforms the network with 40 hidden neurons. In both cases the smallest relative error in predicting eigenenergy values is in proximity to values of oscillator frequencies that were used in learning mode ($\omega_x \in \langle 0.4, 0.9 \rangle$ and $\omega_y \in \langle 1.0, 1.9 \rangle$). From figs. 3 and 4 one can observe that both networks have a poor performance in predicting eigenenergies outside oscillator frequencies that were used in training. The networks poor performance is particularly evident for small values of oscillator frequencies. This behaviour is directly related to the Hamilton operator of two-dimensional potential energy as given by eq. (5) and the slope of this surface for small values of ω_x and ω_y.

Fig. 5 illustrates the error in predicting the eigenenergy of a not fully connected BPN for 40666 learning iterations. The network contained 49 input neurons, 3 hidden layer neurons and 1 output neuron as described above for the results presented in tables 1 and 2. Comparing its performance with the network used to obtain the results as depicted in fig. 3, one can see that this network gives the best results. A very good performance of this network yields a wide attractor. While increasing the oscillator frequencies, the error in predicting eigenenergies is slowly increasing. However, requesting the network to predict eigenenergies for small oscillator frequencies leads to a rapid increase of the error that is related to the slope of two-dimensional potential energy for small values of ω_x and ω_y.

The presented results also raise the question whether the neural network can generalize the acquired knowledge in a meaningful way. It was shown that we succeeded in this task, but not always. When using a new and previously unknown input that notably deviated from the learning phase, the network was not capable to generalize the acquired knowledge. The degree of generalization was also related to the number of neurons in the hidden layer. We observed that the ability of generalization was decreasing as the number of hidden neurons was increasing. Subsequently our network with three hidden neurons outperformed the network investigated by Darsey et al.'s [7] in predicting the eigenenergy of a two-dimensional harmonic oscillator.

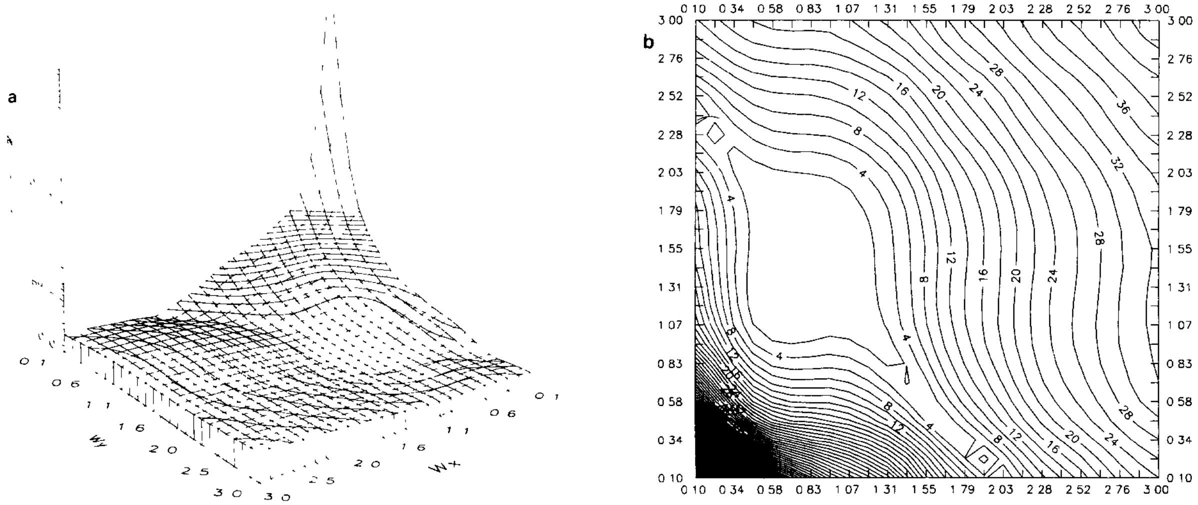

Fig. 5. The relative error (%) in predicting the eigenenergy of a two-dimensional harmonic oscillator for different values of oscillator frequencies ω_x and ω_y. The results were obtained for a not fully connected BPN containing 49 input neurons, 3 hidden layer neurons and 1 output neuron after 40666 learning iterations. Note that due to difficulties with the drawing program the maximum relative errors are in opposite corners. In fact they should be in the same corner of the $\{\omega_x, \omega_y\}$ surface.

5. Summary

In this work we presented studies of a neural network capable of performing the transformations generated by the Schrödinger equation required to find eigenenergies of a two-dimensional harmonic oscillator. We have shown that this task can be achieved by a not fully connected back-propagation neural network for short CPU times, thus making possible the investigation of more complex problems. Our work also demonstrates that some ingenuity is necessary to create the most appropriate neural network to solve a particular problem, as in our case the two-dimensional harmonic oscillator. However, the main conclusion resulting from this and Darsay et al.'s [7] work is that neural networks can be used to investigate more perplexing questions related to basic issues of physics and chemistry.

Acknowledgement

We wish to thank Professor C. Bojarski of the Institute of Technical Physics and Applied Mathematics, Technical University of Gdańsk for his interest, encouragement and support of this research. The authors would also like to thank Tomasz Jankowski for his NN (naturally neural) support and enthusiasm.

References

[1] J.M. Zurada, Introduction to artificial neural systems (West Publishing Company, St. Paul, MN, 1992).

[2] J. Dayhoff, Neural network architectures – an introduction (Van Nostrand Reinhold, New York, 1990);
R. Eckmiller and C. v.d. Mahlsburg, eds., Neural computers, NATO ASI Senes, Vol. F41 (Springer, Berlin, 1988).

[3] A. Maren, C. Harsten and R. Rap, Handbook of neural computing applications (Academic Press, New York, 1990);
H. Haken, ed., Neural and synergetic computers (Springer, Berlin, 1988).

[4] T. Kohonen, Self-organisation and asociative memory, 3rd Ed. (Springer, Berlin, 1989) pp. 250.

[5] M.B. Pour-El and I. Richards, Computability in analysis of physıcs (Sprınger, New York, 1989);
M.R. Garey and D.S. Johnson, Computers and intractability: a guide to the theory of NP-completeness (Freeman, San Francisco, 1979);
S. Wolfram, Phys. Rev. Letters 54 (1985) 735.

[6] M. Caudıll, AI Expert, January 1991, pp. 56;
M. Caudill and C. Butler, Naturally intelligent systems (MIT Press, Cambridge, 1991).

[7] J.A. Darsey, D.W. Noid and B.R. Upadhyaya, Chem. Phys. Letters 177 (1991) 189; 181 (1991) 386 (E).