INSTITUTE OF PHYSICS PUBLISHING NANOTECHNOLOGY
Nanotechnology **17** (2006) S173–S177 doi:10.1088/0957-4484/17/7/S12

Novel amplitude and frequency demodulation algorithm for a virtual dynamic atomic force microscope

J Kokavecz[1,2], Z Tóth[3], Z L Horváth[2], P Heszler[3] and Á Mechler[3,4]

[1] Institute for Engineering and Materials Science, University of Szeged, PO Box: 406, H-6701 Szeged, Hungary
[2] Department of Optics and Quantum Electronics, University of Szeged, PO Box: 406, H-6701 Szeged, Hungary
[3] Research Group on Laser Physics of Hungarian Academy of Sciences, PO Box: 406, H-6701 Szeged, Hungary
[4] School of Chemistry, Monash University, Clayton, VIC 3800, Australia

E-mail: kjanos@physx.u-szeged.hu (J Kokavecz)

Received 18 August 2005, in final form 7 November 2005
Published 10 March 2006
Online at stacks.iop.org/Nano/17/S173

Abstract
Frequency-modulated atomic force microscopy (FM-AFM; also called non-contact atomic force microscopy) is the prevailing operation mode in (sub-)atomic resolution vacuum applications. A major obstacle that prohibits a wider application range is the low frame capture rate. The speed of FM-AFM is limited by the low bandwidth of the automatic gain control (AGC) and frequency demodulation loops. In this work we describe a novel algorithm that can be used to overcome these weaknesses. We analysed the settling times of the proposed loops and that of the complete system, and we found that an approximately 70-fold improvement can be achieved over the existing real and virtual atomic force microscopes. We show that proportional–integral–differential controllers perform better in the frequency demodulation loop than conventional proportional–integral controllers. We demonstrate that the signal to noise ratio of the proposed system is 5.7×10^{-5}, which agrees with that of the conventional systems; thus, the new algorithm would improve the performance of FM-AFMs without compromising the resolution.

1. Introduction

Frequency modulation (FM) mode (also called non-contact mode) is a promising operation mode of the atomic force microscope (AFM) in ultrahigh vacuum environment, reportedly achieving real atomic resolution imaging of various substrates on a routine basis [1–4]. Besides topography, FM mode also yields the map of the energy dissipated by the tip–sample interaction. FM mode has also been used for measuring the strength of covalent bonds, manipulating atoms and recording high resolution force–distance curves [5–7].

FM mode image calculations and force–distance curve simulations, mandatory for the correct interpretation of the experimental data, have been performed on various pretences. Most simulations are based on the *static* approximation [8, 9], where the tip oscillates at a steady state without frequency or dissipation transients, and automatic gain control (AGC) and frequency demodulation loops work ideally. This approximation is legitimate only at low scan speeds when the probe has enough time to reach an equilibrium frequency and dissipation, and the transients in the AGC and frequency demodulation loops have also been decayed.

Recently, virtual FM-AFM machines were introduced to overcome the limitations of the static approximation [10–12]. The virtual machine treats the tip and the control loops as dynamical systems and handles the transients realistically. Moreover, the structure and the system parameters of the virtual machine were determined from a precise experimental

evaluation of a real AFM [10]. Consequently, a virtual FM-AFM can be used to analyse the image formation and stability as well as the dynamics of the AGC and frequency demodulation loops. Virtual FM-AFM simulations provided the explanation for the spatial shift and contrast inversion between topographical and dissipation images and also for the extreme sensitivity of the dissipation to a tip change [12].

Recent image calculations have shown that the slow response of the AGC and frequency demodulation loops causes significant image distortion at 100 nm s^{-1} scan speed [10]. The response time of the control loops is limited by the low pass filters which are used within the loops. For example, the automatic gain control loop bandwidth is around 0.5 kHz while the frequency demodulator loop bandwidth is limited to ∼1 kHz. Assuming 270 kHz cantilever resonance frequency, one can calculate that during the amplitude and frequency measurements several hundred oscillations are averaged. However, even a single oscillation contains the requested information. Here, we present a complete virtual FM-AFM which uses fast automatic gain control and frequency demodulation loops. The algorithm is based on the Fourier method, which allows the determination of the amplitude and phase from a single oscillation. Accordingly, the settling time of the system can be shortened by more than an order of magnitude. The proposed algorithm is suitable not only for virtual image calculations, but also could be realized on a real AFM.

2. Description of the virtual NC-AFM machine

In FM mode, the cantilever oscillation is regulated to provide a constant amplitude at its resonance frequency by applying positive feedback through a phase shifter and an automatic gain control amplifier. The phase shifter is needed to eliminate the resonance frequency shift caused by the feedback loop. It has been shown that the ideal AGC loop should have $\pi/2$ phase shift at all frequencies [13]. Forces acting on the tip attached to the lower part of the cantilever cause modulation of the oscillator output (e.g. frequency), which is determined by a frequency demodulator. To measure the topography, a third control loop called the distance regulation loop is applied, which keeps the frequency shift constant by actuating the base of the cantilever. The block diagram of the system is shown in figure 1.

2.1. Cantilever motion and the AGC loop

In our model, the dynamics of the tip is governed by equation (1):

$$m^* x''(t) = -kd(t) - \left(\frac{m^*\omega_0}{Q} - G(t)\right) d'(t) - F_{ts}(x, x'). \tag{1}$$

Here, x denotes the position of the tip apex relative to the sample surface while d represents the cantilever deflection. The following cantilever parameters are included in the model: spring constant (k), undamped angular resonance frequency (ω_0) and quality factor (Q). The phase shifter and the variable gain amplifier are modelled by the $G(t)d'(t)$ term. Here, $G(t)$ stands for the loop gain and is also used as dissipation signal.

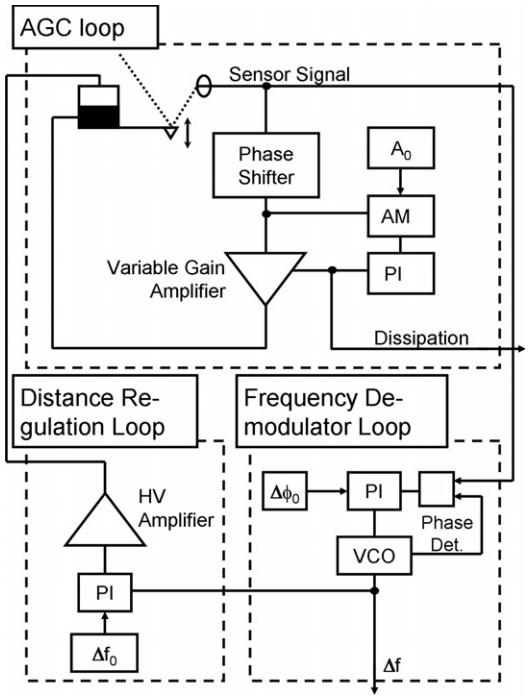

Figure 1. Block diagram of the FM-AFM. The vibration amplitude of the cantilever is maintained constant by the AGC loop. The frequency shift (Δf) is continuously monitored by the frequency demodulation loop. The distance regulation loop holds the frequency shift at a predefined constant value by actuating the base position of the cantilever. The overall response speed of the system is limited since the AGC and frequency demodulation loops average over several hundred oscillations.

The Q-control technique uses a similar feedback loop having a constant gain to modify the effective quality factor of the probe in amplitude modulated mode [13, 14]. In FM mode, the gain is varied in order to maintain constant vibration amplitude. Assuming a sinusoidal deflection signal, which is legitimate for typical operating conditions [15], the phase difference between $d(t)$ and $d'(t)$ is exactly $\pi/2$; therefore, $d'(t)$ can be used as the output of the phase shifter. Although this representation of the phase shifter is simple and accurate, several papers represent the phase shifter by a time delayed term [5, 16, 17]. However, the time delayed representation results in varying phase shift as the frequency changes. Accordingly, the eigenfrequency of the system is slightly modulated by the AGC loop, leading to artifacts in the topography [13]. A further disadvantage of the time delayed representation is that the accurate solution of the time delayed model requires significantly shorter integration steps, which prolongs the calculation time [18].

2.2. Amplitude and frequency demodulation algorithm

To determine the amplitude and phase from a single oscillation the Fourier method can be used. The integration should be performed for exactly one (or integer number of) oscillation(s). However, the oscillation frequency may change in frequency modulated mode, so the integration time should be varied with the actual frequency as depicted in figure 2. To determine the amplitude ($A(t)$) and phase difference ($\Delta\phi(t)$), equations (2*c*)

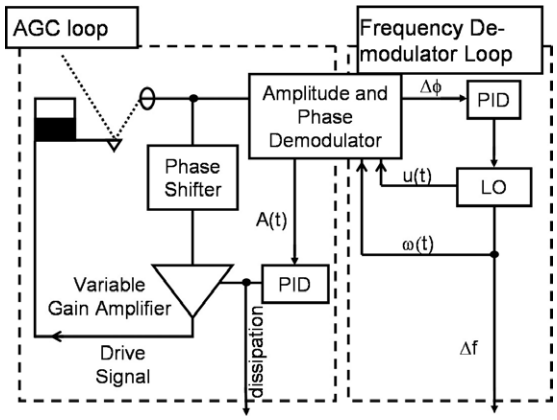

Figure 2. Block diagram of the proposed AGC and frequency demodulation loops. Amplitude and phase measured from one oscillation using the Fourier method. The calculated amplitude ($A(t)$) is sent to a PID controller, which controls the gain of the AGC loop. The phase difference ($\Delta\phi$) between the cantilever deflection and the local oscillator (LO) signals is also maintained constant by a PID controller which regulates the LO frequency ($\omega(t)$). To perform the amplitude and phase difference measurement for exactly one period, the frequency and phase ($u(t)$) of the LO serve as input of the Fourier method.

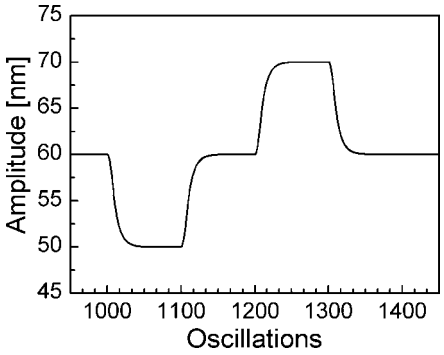

Figure 3. AGC loop performance. The desired amplitude was set to 60 nm and decreased to 50 nm between 1000 and 1100 oscillations while it was increased to 70 nm between 1200 and 1300 oscillations. The settling time of the AGC loop is ∼40 oscillations. In this simulation, we used a free cantilever whose quality factor was 5×10^4.

and (2*d*) are used, respectively:

$$A_{\sin}(t) = \frac{\omega(t)}{\pi} \int_0^{\frac{2\pi}{\omega(t)}} d(t-\tau)\sin(u(t-\tau))\,\mathrm{d}\tau \qquad (2a)$$

$$A_{\cos}(t) = \frac{\omega(t)}{\pi} \int_0^{\frac{2\pi}{\omega(t)}} d(t-\tau)\cos(u(t-\tau))\,\mathrm{d}\tau \qquad (2b)$$

$$A(t) = [A_{\sin}^2(t) + A_{\cos}^2(t)]^{1/2} \qquad (2c)$$

$$\Delta\phi(t) = \arctan\left(\frac{A_{\cos}(t)}{A_{\sin}(t)}\right). \qquad (2d)$$

The calculated amplitude is fed into a proportional–integral–differential (PID) controller, which is implemented as follows (equation (3)):

$$w^{i+1} = w^i - P(v^i - v_0) - I\sum_{j=0}^{M}(v^{i-j} - v_0) - D(v^i - v^{i-1}). \quad (3)$$

Here, w^i, w^{i+1} and v^i, v^{i+1} stand for the output and input values in the ith and $(i+1)$th time steps, while v_0 would be the desired value. The constants P, I and D are the proportional, integral and differential gains, respectively. Finally, M denotes the length of the array which is used for summation.

Frequency demodulation is based on the phase locked loop concept. The phase difference between the deflection and the local oscillator's signal ($\Delta\phi(t)$) is calculated using equation (2*d*) and is sent to the PID controller (equation (3)) which regulates the frequency of the local oscillator ($\omega(t)$). The local oscillator frequency is used to calculate the integration time to be used in the next time step. The phase of the local oscillator $u(t)$ is the integral of the local oscillator's frequency (equation (4)):

$$u(t) = u_0 + \int_0^t \omega(t)\,\mathrm{d}t. \qquad (4)$$

The virtual FM-AFM solves the equation of motion of the probe (equation (1)) using a fourth order Runge–Kutta algorithm. The cantilever deflection value was stored in a first-in–last-out array. Next, the amplitude ($A(t)$) and phase difference ($\Delta\phi(t)$) were calculated using equations (2*c*) and (2*d*). Typically, 5000 data points were used, which represented one period. The calculated amplitude and phase differences were passed to the PID controllers (equation (3)), and the gain ($G(t)$) and local oscillator frequency ($\omega(t)$) were updated. The phase of the local oscillator ($u(t)$) was determined using equation (4), and the first-in–last-out array was resized in order to perform the integration in the next time step for exactly one period. Finally, the new base position of the cantilever was calculated from the local oscillator frequency using the PID algorithm (equation (3)).

3. Results and discussion

First, we tested the AGC loop performance. In this test the desired amplitude was modulated in steps as can be seen in figure 3. The desired amplitude was 60 nm, which was decreased to 50 nm between 1000 and 1100 oscillations and increased to 70 nm between 1200 and 1300 oscillations. The settling time of the amplitude signal was ∼40 oscillations, which is very small if we take into account that the quality factor was 5×10^4. Since this test was performed for a free cantilever, that is, the surface forces were neglected, the settling time without amplitude regulation had been $\sim 5 \times 10^4$ oscillations. We did not observe any significant differences between the cases when the amplitude increased or decreased, a result which indicates that the loop gain changes its sign, that is, it jumps from positive feedback to negative.

Next, we tested the response of the frequency demodulation loop (figure 4). Here, we applied a sinusoidal signal to the input of the loop and changed the frequency. The relative frequency was changed between 0.999 and 1.001, a range of normal FM operation. The settling time of the frequency demodulation loop was ∼20 oscillations, which is slightly better than that of the AGC loop.

The settling time of the frequency demodulation strongly depends on the type and proper setting of the controller.

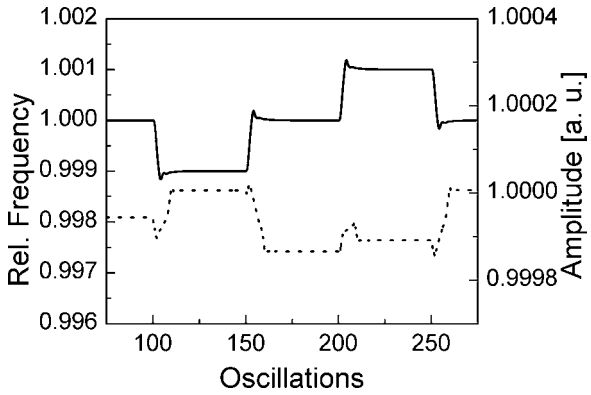

Figure 4. Performance of the frequency demodulation loop and noise of the amplitude signal. We applied a $\sin(2\pi f)$ signal to the input of the frequency demodulation loop and varied f between 0.999 and 1.001. The frequency signal (——) and the amplitude signal (· · · · · ·) are plotted. Deviation of the amplitude signal from unity originates from the resizing of the integration array and the integration time slightly differs from the periodic time of the actual frequency. In the present calculation, the integration array contained 5000 elements and the relative RMS noise of the amplitude signal was 5.7×10^{-5}.

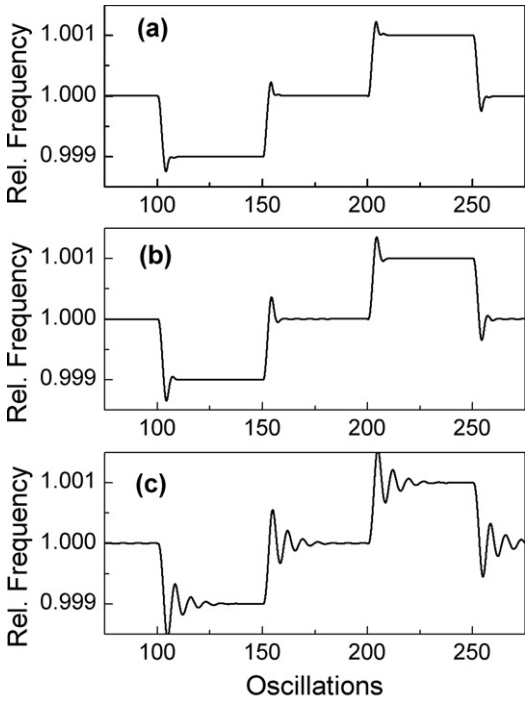

Figure 5. Importance of the differential gain. We applied the same signal to the frequency demodulation loop as in figure 4 and varied the differential gain among values 0.25 (figure 5(a)), 0.17 (figure 5(b)) and 0.1 (figure 5(c)). The proportional and integral gains were not altered. The proportional gain was 10^{-5} while the integral gain was set to zero.

We experienced that fast settling time was impossible to achieve using a commonly applied PI controller, but it was quite easy with a PID controller, which indicates that the differential gain has a strong influence on the settling time. To show the importance of the differential gain, let us consider that there is a constant frequency (ω_{in}) at the input of the frequency demodulation loop while the frequency of the local oscillator is $\omega(t)$. It is easy to see that the rate of the phase difference is equal to the difference between the two frequencies (equation (5)):

$$(\Delta\phi)'(t) = \omega(t) - \omega_{\text{in}}. \tag{5}$$

When substituting the phase difference as input and the local oscillator frequency as output into the PID algorithm (equation (3)) we get the following expression:

$$\omega^{i+1} - \omega^i = -P(\Delta\phi^i - \Delta\phi_0) - I\sum_{j=0}^{M}(\Delta\phi^{i-j} - \Delta\phi_0) - D(\Delta\phi^i - \Delta\phi^{i-1}). \tag{6}$$

Here, $\Delta\phi_0$ is the desired phase difference which is set by the user. Hence, we can assume that the desired phase difference is zero without loss of generality. Moreover, if we assume infinitely short time steps then we can write time derivatives instead of the differences. Substituting equation (5) into (6) we get

$$(\Delta\phi)''(t) = -P(\Delta\phi)(t) - I\int(\Delta\phi)(t)\mathrm{d}t - D(\Delta\phi)'(t). \tag{7}$$

Equation (7) is identical with the equation of motion of a damped oscillator, provided that the integral term is neglected ($I = 0$). In this case, the differential gain plays the same role as the damping. Hence, the best system performance can be achieved in the aperiodic case. The importance of the differential gain is demonstrated in figures 5(a)–(c). Here, we can observe that a decrease of the differential gain results in more oscillations before reaching the equilibrium state.

Unfortunately, the aperiodic case cannot be achieved since $\Delta\phi$ calculated by the algorithm is averaged for one cycle, therefore equation (5) does not stand. This is even more pronounced in real AFMs, where several hundred oscillations are averaged, therefore the phase signal is not the time derivative of the frequency difference, but rather an integral of it; therefore, the integral gain becomes crucial in the phase locked loop control instead of the differential gain.

Noise of the amplitude measurement can be seen in figure 4. Amplitude noise originates from the frequency changes, since as the frequency changes the length of integration time has to be changed, too. However, the integration time cannot be modified at will since the deflection values are available only at those time points where they were calculated. To minimize this error, we chose 5000 points per period, which allowed a satisfactorily accurate choice of integration time. The relative root-mean-square noise of the amplitude was 5.7×10^{-5}. Alternatively, the noise can be reduced by application of a window function [19].

Finally, we tested the settling time of the whole system. In this test the base position of the cantilever was modulated by a 0.5 nm square signal. To compare our virtual AFM to existing real and virtual machines we used the same parameters as in [10]. Consequently, the resonance frequency of the probe was 270 kHz while the quality factor and cantilever spring constant were 4.5×10^4 and 30 N m^{-1}, respectively. Results are depicted in figure 6.

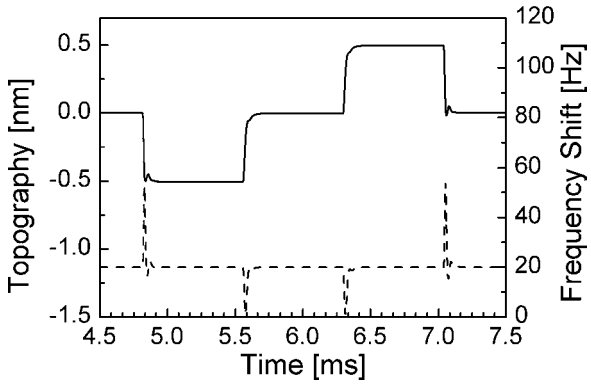

Figure 6. Response time of the complete system. Here we used the same data as in [10]. The response time of our system is about 0.1 ms while existing real and virtual AFMs' response time is 7 ms.

The settling time of our virtual AFM is about 0.1 ms, which is 70 times faster than that of the other machine (7 ms). The significantly shorter response time is a result of the fast amplitude and phase measurement as well as the appropriately tuned controller.

The Fourier method is not the only procedure to determine the amplitude from a single oscillation. Amplitude can be measured just by two peak detectors which record the extremal values of the signal [20, 21]. However, this method is strongly affected by the noise, since only two data points are used. In the case of the Fourier method many more data points are used, leading to significantly lower noise. Similarly, the phase difference between two sinusoidal signals can be measured from a single oscillation, but this method has much worse signal to noise ratio than the Fourier method [22].

We emphasize that the proposed algorithm can be used not only for virtual FM-AFM calculations, but also for implementation in real systems, which in turn could open new fields for fast FM-AFM imaging.

4. Conclusion

In this work, we proposed a novel algorithm for the AGC and frequency demodulation loops of FM-AFM. By analysing the settling times of the loops and that of the complete system, we demonstrated that an approximately 70-fold improvement can be achieved in comparison to the settling time of existing real and virtual FM-AFMs. We demonstrated that proportional–integral–differential controllers perform better in the frequency demodulation loop than conventional proportional–integral controllers. A detailed analysis shows that the noise inherent in the algorithm would fall into the same range as the noise of the existing virtual and real machines. This fact promises improved performance without compromising the resolution. The concepts discussed here are valid for, and can be implemented in, a real FM-AFM, leading to new avenues for FM-AFM imaging.

Acknowledgments

Financial support from TS049872, T38152 and T046394 grants of the Hungarian Scientific Research Fund (OTKA), Australian Research Council Discovery grants DP0662816 and DP0663290, and the grant NKFP 3A/071/2004 of the Hungarian Ministry of Culture and Education are acknowledged with thanks. ÁM gratefully acknowledges his Monash Fellowship.

References

[1] Morita S, Wiesendanger R and Meyer E 2003 *Noncontact Atomic Force Microscopy* (Berlin: Springer)
[2] Giessibl F 2003 *Rev. Mod. Phys.* **75** 949
[3] Albrecht T R, Grütter P, Horne D and Rugar D 1991 *J. Appl. Phys.* **69** 668
[4] Binnig G, Gerber Ch and Quate C F 1986 *Phys. Rev. Lett.* **56** 930
[5] Gotsmann B, Anczykowski B, Seidel C and Fuchs H 1999 *Appl. Surf. Sci.* **140** 314
[6] Oyabu N, Custance O, Yi I, Sugawara Y and Morita S 2003 *Phys. Rev. Lett.* **90** 176102
[7] Lantz M A, Hug H J, Hoffmann R, van Schendel P J A, Kappenberger P, Martin S, Baratoff A and Güntherodt H J 2001 *Science* **291** 2580
[8] Pang C L, Raza H, Haycock S A and Thornton G 2002 *Phys. Rev.* B **65** 201401
[9] Foster A S, Barth C, Shluger A L and Reichling M 2001 *Phys. Rev. Lett.* **86** 2373
[10] Polesel-Maris J and Gauthier S 2005 *J. Appl. Phys.* **97** 044902
[11] Couturier G, Aimé J P, Salardenne J and Boisgard R 2001 *Eur. Phys. J.: Appl. Phys.* **15** 141
[12] Gauthier M, Pérez R, Arai T, Tomitori R and Tsukada M 2002 *Phys. Rev. Lett.* **89** 146104
[13] Mertz J, Marti O and Mlynek J 1993 *Appl. Phys. Lett.* **62** 2344
[14] Humphris A D L, Tamayo J and Miles M J 2000 *Langmuir* **16** 7891
[15] Nony L, Boisgard R and Aimé J P 1999 *J. Chem. Phys.* **111** 1615
[16] Hölscher H, Gotsmann B, Allers W, Schwartz U D, Fuchs H and Wiesendanger R 2001 *Phys. Rev.* B **64** 075402
[17] Gauthier M, Sasaki N and Tsukada M 2001 *Phys. Rev.* B **64** 085409
[18] Kokavecz J, Horváth Z L and Mechler Á 2004 *Appl. Phys. Lett.* **85** 3232
[19] Orfanidis S J 1995 *Introduction to Signal Processing* (Englewood Cliffs, NJ: Prentice-Hall)
[20] Ando T, Kodera N, Takai E, Maruyama D, Saito K and Toda A 2001 *Proc. Natl Acad. Sci.* **98** 12468
[21] Ando T, Kodera N, Naito Y, Kinoshita T, Furuta K and Toyoshima Y Y 2003 *Chem. Phys. Chem.* **4** 1196
[22] Stark M and Guckenberger R 1999 *Rev. Sci. Instrum.* **70** 3614