



SpectraFox: A free open-source data management and analysis tool for scanning probe microscopy and spectroscopy

Michael Ruby

Freie Universität Berlin, Arnimallee 14, 14195 Berlin, Germany

Received 1 February 2016; received in revised form 1 April 2016; accepted 5 April 2016

Abstract

In the last decades scanning probe microscopy and spectroscopy have become well-established tools in nanotechnology and surface science. This opened the market for many commercial manufacturers, each with different hardware and software standards. Besides the advantage of a wide variety of available hardware, the diversity may software-wise complicate the data exchange between scientists, and the data analysis for groups working with hardware developed by different manufacturers. Not only the file format differs between manufacturers, but also the data often requires further numerical treatment before publication. SpectraFox is an open-source and independent tool which manages, processes, and evaluates scanning probe spectroscopy and microscopy data. It aims at simplifying the documentation in parallel to measurement, and it provides solid evaluation tools for a large number of data.

© 2016 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Keywords: Scanning probe microscopy; Scanning probe spectroscopy; Data overview; Data management; Data analysis; Batch processing

Code metadata

Current code version	2016.04.01.0
Permanent link to code/repository used of this code version	https://github.com/ElsevierSoftwareX/SOFTX-D-16-00022
Legal Code License	LGPLv3.0
Code versioning system used	git
Software code languages, tools, and services used	VB.NET, C#
Compilation requirements, operating environments & dependencies	.NET Framework 4.5
If available Link to developer documentation/manual	http://wiki.spectrafox.com
Support email for questions	contact@spectrafox.com

Software metadata

Current software version	2016.04.01.0
Permanent link to executables of this version	https://github.com/ElsevierSoftwareX/SOFTX-D-16-00022
Legal Software License	LGPLv3.0
Computing platform/Operating System	Microsoft Windows Vista or higher (32- or 64-bit)
Installation requirements & dependencies	.NET Framework 4.5, data acquisition tools: SPECS/Nanonis, Omicron/Matrix, Createc, Nanotec/WSxM
If available, link to user manual—if formally published include a reference to the publication in the reference list	http://wiki.spectrafox.com
Support email for questions	contact@spectrafox.com

E-mail address: m@mikeruby.net.

1. Motivation and significance

Scanning probe microscopy and spectroscopy has become one of the key technologies in modern surface science. This is due to its capabilities in real space investigation of surface structures and their local electronic or magnetic environment. The technique has forked into many different variations, which specialized to different fields of research [1–6]. They all have in common that a sharp probe, for instance the apex of a metal wire, scans the sample surface in close vicinity. The measured quantity reveals the interaction between tip and sample, which may be, e.g., an electric current in scanning tunneling microscopy, or a force in atomic force microscopy.

Independent of the exact operation principle of the microscope, two types of data can be generated. In scanning probe *microscopy* (SPM) the output is a two-dimensional matrix of values representing the spatial variation of a specific sample parameter. The most prominent example is the surface topography. In scanning probe *spectroscopy* (SPS) the output is a one-dimensional trace of values, which is recorded at a single spatial location, where the reaction of the sample is measured on the variation of a junction parameter. For instance, in scanning tunneling spectroscopy the sample bias voltage is changed to measure the differential conductance of the junction.

Many independent research laboratories and manufacturers have improved SPM and SPS in the last decades. This includes the construction of different data acquisition systems, of which most of them use a proprietary file format to store their data. This complicates the exchange between scientists. It is especially problematic for research groups dealing with hardware from different manufacturers. The quality of the data analysis tools provided by the hardware suppliers also differs in their quality and features. So far only two free tools address these issues [7,8]. Only one of them is available under terms of an open-source license. The focus of these programs lies in handling microscopy images, rather than spectroscopy files.

SpectraFox is a free, independent, and open-source data management and analysis tool, which closes this gap, and puts a focus on the handling of spectroscopy files. It is suited to accompany a project during an ongoing measurement to keep the overview of the measured data, and to extract data quickly for documentation. After the measurement, SpectraFox provides tools for elaborated data analysis and evaluation. It includes many features, such as simple numerical data treatments applicable to a whole batch of files, or more complex features, such as non-linear least square fitting, and advanced visualization techniques. Finally it can export all results, e.g., to formats used in commercial software, such as OriginLab's Origin, Wolfram's Mathematica or Mathworks Matlab, which may be used to create a publication-ready output.

Since its launch, SpectraFox has proven its capabilities in the daily work of several research groups, and contributed to multiple publications [9–13]. This article describes the main features of SpectraFox. It extends the detailed instruction manual available online. We demonstrate its capabilities in a real world example, where we show the evaluation of a set

of physical data recorded with a low-temperature scanning tunneling microscope under ultra-high vacuum conditions.

2. Software framework

2.1. Software architecture

SpectraFox is based on Microsoft's .NET Framework 4.5. Its user interface is based on multiple windows for each functionality. Multi-threading is heavily used, which takes advantage of modern multi-core processors. For complex numerical problems the computation capabilities can be further improved by using a dedicated graphic processor via the OpenCL or CUDA interfaces.

The software is released under an open-source license and features a plugin-interface. Both allow an easy extension of the software's capabilities by the user, such as custom file handling routines, data treatment procedures, or models for the non-linear least square fitting.

SpectraFox's architecture is sketched in Fig. 1. It manages data files by dedicated project folders. The typical usage scenario is to point the software to the folder containing the output of a manufacturer's measurement software. SpectraFox scans the directory for compatible files, and creates an index which is stored in a cache-file. This index contains measurement related information included in the file-headers, and allows to search for data with specific properties. As scanning all files may be a heavy task for projects containing many thousand files, the usage of a cache-file speeds up subsequent loading times. For larger projects, this architecture retains a low memory footprint.

The output generated by SpectraFox's data treatment routines is stored separately from each source file. This always keeps the original data untouched. It protects from data corruption, thus forbidding any manipulation of scientific data.

2.2. Software functionalities

SpectraFox is a tool that collects and displays measurement related meta-information extracted from the headers of the data files. For instance, the user comments, the record location of spectra or images, the record time, or the junction properties, such as feedback conditions or voltages. This collection of information is displayed in an overview for all data-files, together with preview images of the data itself. The list is sortable, and can be filtered or searched for files with specific properties. Furthermore, SpectraFox does not treat spectroscopy data separated from microscopy images. Instead it connects spectra recorded at different locations of the sample with related images.

A primary goal of SpectraFox is to simplify the processing of a large number of data-files. Therefore all basic operations applied to spectroscopy files are applicable not only to one file, but can be executed, once set up, to many files simultaneously. These basic tasks for spectroscopy data comprise

1. offsetting data columns, or adding multiple columns,
2. multiplying data columns with a scalar factor or values of another column,

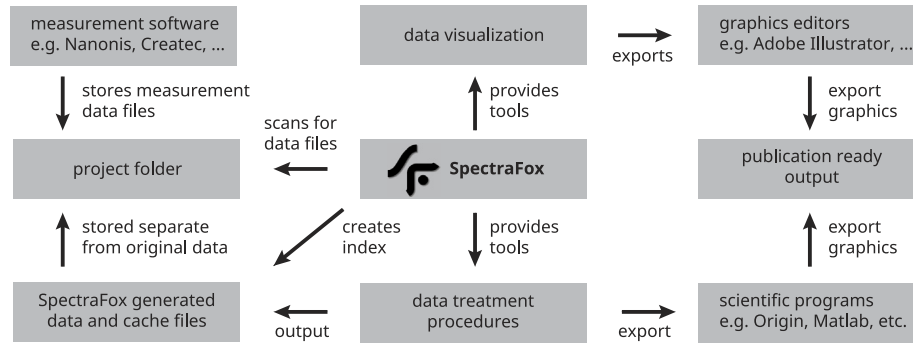


Fig. 1. Architecture of SpectraFox. The output of the hardware manufacturer’s measurement software is stored in a project folder. SpectraFox scans the folder for compatible files and creates an index of their properties. It further provides data treatment procedures and analysis tools. Their output is stored together with the cached property index in the project folder. The data files can also be exported to a file format accepted by common scientific programs. Tools for data visualization can export their images for further processing to other graphics editors.

3. cropping data to arbitrary ranges,
4. averaging of multiple data columns within one file, or between several files,
5. smoothing data columns using adjacent average or Savitzky–Golay filters,
6. reversing the order of data in a column,
7. calculating the numeric derivative of a column,
8. normalizing data columns to a constant value in a given range to compare different files,
9. “regauging” differential conductance data measured with a lock-in amplifier to restore the physical units, and
10. performing a deconvolution of the frequency shift and current signals recorded in frequency-modulated atomic force microscopy.

Some of these features may seem trivial, but SpectraFox is the only software package that bundles them for a direct application to a batch of data files.

More complex tasks require individual settings. They are thus performed for each file separately. Here, an important feature is the non-linear least square fit tool for spectroscopy data. SpectraFox contains many fit functions for common physical problems, such as Gaussian or Lorentzian distributions, or Fano and Frotalineshapes [14–16], relevant in interference processes. More specialized fit routines provide models developed for superconductor physics and inelastic tunneling spectroscopy, in particular dealing with inelastic spin excitations. The first requires the convolution of two sets of data, and the latter performs a diagonalization of the zero-field spin Hamiltonian [17]. All fit models can be combined and coupled. This allows to fit several spectral features within one spectrum, even if they overlap each other. For specialized problems we encourage the user to extend SpectraFox by plugin libraries. A highlight is the possibility to fit two datasets simultaneously, and with intercoupled parameter sets, which is – to the authors best knowledge – a unique feature of SpectraFox among other data analysis tools. This feature may be helpful in many scenarios. For instance, in measurements at different but physically equal locations a coupled fit leads to an optimized result.

Furthermore, a core feature of SpectraFox is the visualization of spectroscopy data. Besides plotting single files, SpectraFox can plot multiple spectra on top of each other, or as

a stacked waterfall plot. Multiple spectra can be visualized as well with respect to their spatial location in two-dimensional color plots. Here, the spectral intensity is represented by a color intensity, which is plotted on top of a related topographic image. This is especially suited for data recorded in a grid of spatial points, but may be used as well for arbitrarily placed spectra. A common practice in SPS is to record multiple spectra along a straight line, to investigate the evolution of spectral features along the lateral position. SpectraFox visualizes this evolution in a color plot, where the spectral intensity is plotted with respect to the position, and the quantity varied within the spectrum.

SpectraFox, initially not intended to provide functionalities for dealing with microscopy images since others have specialized to this [7,8], contains nevertheless some helpful features for this purpose. For instance a tool to extract line profiles across such images. A unique feature of SpectraFox among other free SPM analysis tools is the possibility to combine multiple images recorded at different lateral locations to one single overview image.

3. Implementation and physical concepts

In Section 2.2 several basic features provided by SpectraFox are enumerated. Most of them are self-explaining, so that we refer to the user manual for details about their handling. Here, we want to focus on giving details of two tools that deal with specialized problems in scanning tunneling spectroscopy (STS) and atomic force microscopy (AFM).

In STS-spectra one is interested in the differential conductance (dI/dV) with respect to the voltage difference of the tip-sample junction. This signal is proportional to the local density of states of the sample, under the assumption of a constant density of states of the tip [18]. In this case it is a measure of the local electronic structure. Since the quantity is the derivative of the current with respect to the sample voltage, it is prone to all kinds of noise. Instead of recording the signal directly one often uses a lock-in technique to enhance the signal-to-noise ratio. As a drawback the absolute physical units of the signal are lost. SpectraFox provides the possibility to restore these units—a procedure which we call “regauging”. The simplest way requires to know the settings of the lock-in amplifier

that were applied during the measurement. The lock-in signal can then easily be regauged back to the absolute physical units. However, if the values are not available, SpectraFox offers an alternative way to regauge the lock-in signal. It requires the simultaneously recorded current across the junction. The procedure first smooths the signal of the current, and then calculates its numeric derivative with respect to the bias voltage. The output has the correct physical dimensions and contains all features of the lock-in signal, however with a high noise level. In general, the lock-in signal can contain an offset y_0 , and may be linearly stretched by a factor m with respect to the original signal: $y(x) = y_0 + mx$. By performing a fast minimum-bracketing least square fit of the high quality lock-in signal to the noisy derivative of the current, we obtain the parameters y_0 and m , and can restore the absolute physical dimensions.

A powerful capability of AFM is to measure the non-linear interaction force between tip and sample. A common technique to obtain this is frequency-modulated AFM. Here, the cantilever is oscillated at its resonance frequency. The force between tip and sample leads to a shift of this value with respect to the tip-sample distance $\Omega(z)$, which is tracked by a feedback circuit. The popular qPlus tuning fork design [19] allows to simultaneously record the current across the junction. Due to the high oscillation frequency of the cantilever the output is the time-averaged tunneling current $I_a(z)$. SpectraFox can perform a deconvolution of both signals to extract the interaction force $F(z)$ out of the first signal, and the instantaneous tunneling current $I(z)$ out of the latter one [20,21]:

$$F(z) = \frac{2k}{\omega_r} \int_z^\infty \left[1 + \frac{a^{1/2}}{8\sqrt{\pi(t-z)}} \right] \Omega(t) - \frac{a^{3/2}}{\sqrt{2(t-z)}} \frac{d\Omega(t)}{dt} dt, \quad (1)$$

$$I(z) = I_a(z+a) - \int_z^\infty \sqrt{\frac{2a}{t-z}} \times \left[\frac{dI_a(t)}{dt} - \sqrt{\frac{2}{\pi}} \frac{dI_a(t+a)}{dt} \right] dt. \quad (2)$$

Here, a is the oscillation amplitude, ω_r is the resonance frequency, and k is the spring constant of the cantilever.

For the quantitative evaluation of other physical processes SpectraFox provides a non-linear least square fit engine. It supports the Nelder–Mead and Levenberg–Marquardt optimization algorithms [22–24]. The first is more robust against an inaccurate guess of the initial parameter set, the latter is faster in approaching an optimized minimum.

For the elementary fit functions we refer to the user manual. Here we want to explain two complex fit models that target particular physical problems. The first model convolves two functions. For instance, this is important in scanning tunneling spectroscopy, if the model of a constant density of states of the tip is not applicable. In this case the current integral of the tunnel junction has to be calculated, to separate spectral features originating from the density of states of the tip from those of the sample. This is equivalent to a convolution of the tip and sample

density of states, $\rho_{t,s}$, respectively.

$$I(V, z) = \int \rho_s(V + \varepsilon) \rho_t(\varepsilon) \Delta f(V, \varepsilon) |M(V, z)|^2 d\varepsilon. \quad (3)$$

Here, V is the sample bias voltage, z is the tip-sample distance, $\Delta f(V, \varepsilon)$ is the difference in the occupation of the density of states, and $M(V, z)$ is the tunnel matrix element [25]. To account for additional line broadening, the output is furthermore convolved with a Gaussian function. Each convolution procedure needs a large amount of processing power, especially, if the fit models for the density of states contain slowly decaying functions. SpectraFox uses multi-threading and a cache algorithm to reduce calculation times. This integral approach to calculate the full current also reproduces artificial spectral features. A prominent example is the appearance of a negative differential conductance. This happens if two sharp resonances in tip and sample are convolved with each other. Although the voltage difference is increased, the drop in the density of states leads to a decreasing current. This procedure has proven its capabilities in tunneling spectroscopy experiments on superconducting samples, which are carried out with a superconducting tip. In this case the tip and sample density of states are expressed by the BCS theory [26] for conventional superconductors. Additionally, magnetic impurities may introduce resonances with Lorentzian lineshapes within the superconducting gap. For further details we refer to Refs. [10–13], where data of this kind has been evaluated using SpectraFox.

The second fit model targets inelastic tunneling spectroscopy of spin excitations in adsorbates [27,28,17]. In this case, the tunneling current is obtained by diagonalizing the spin Hamiltonian for a zero-field splitting of spin states by a ligand field:

$$H_{zfs} = DS_z^2 + E(S_x^2 - S_y^2). \quad (4)$$

Here, S is the spin operator, and E and D give the strength of the in-plane and out-of-plane magnetic anisotropy, respectively. As an input, the fit model requires the spin state and the temperature of the system. The latter determines the occupation of the anisotropy split states. It then calculates the transition matrix element between the different spin orientations. This gives the tunneling probability, which increases as soon as the bias voltage reaches the threshold energy to flip the spin of the adsorbate. In essence, it leads to an increased slope of the current across the junction, or to a step in the differential conductance. The signal can then be fitted to the experimental data.

4. Illustrative example

The capabilities of SpectraFox are illustrated by evaluating a real world example. We show a subset of data published in Ref. [13], which has been recorded with a low-temperature scanning tunneling microscope under ultra-high vacuum conditions. The measurement software Nanonis is bundled together with the digital–analog converter hardware by SPECS.

The experimental data is measured on the (111) surface of the superconductor Pb, where monolayer islands of organic

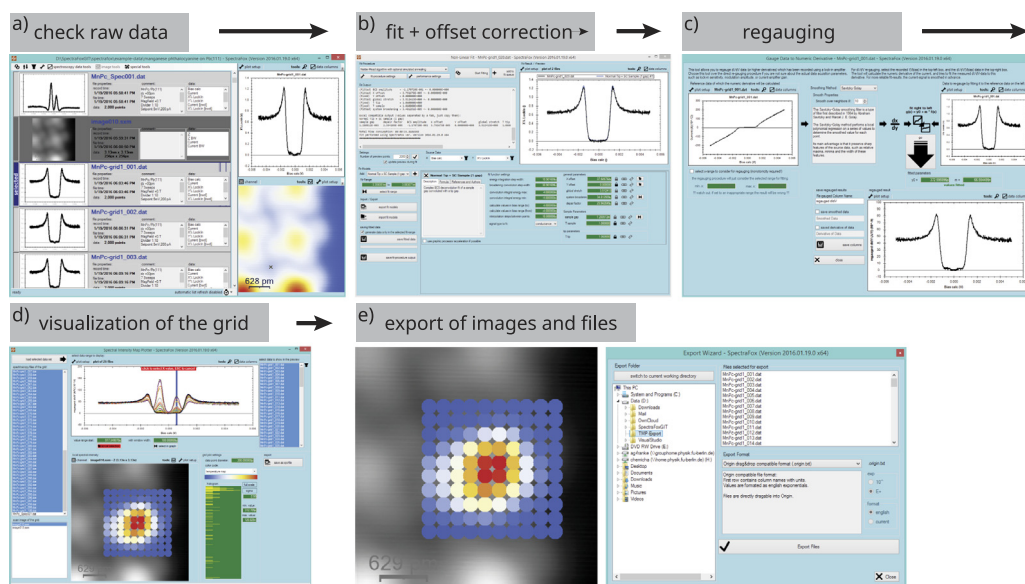


Fig. 2. Exemplary evaluation steps of a set of measurement data recorded with a scanning tunneling microscope. By using SpectraFox the sketched procedure needs less than 13 min to correct the offset of 100 files, to regauge the physical dimensions of their dI/dV -signal, to visualize their local spectral intensity on top of the topographic image, and to transfer the data to external scientific software. A video of all steps is available online at <http://dx.doi.org/10.1016/j.softx.2016.04.001>.

manganese phthalocyanine molecules are adsorbed. Pb is a conventional superconductor described by BCS theory. The magnetic moment of the metal center disturbs the superconducting pairing mechanisms, which creates subgap bound states [29]. These can be resolved in differential conductance spectroscopy. We have measured a dataset at a temperature of 1.1 K consisting of microscopy images and spectroscopy data recorded in a grid across a single molecule. The files can be found online in the example-data folder of the github repository.

To get an overview of the data, its record locations, and properties we start the evaluation by scanning the project folder for compatible files. We used a lock-in technique to acquire the spectroscopy data. The dI/dV -signal is thus given as an arbitrary voltage versus the bias voltage of the sample. Over the clean substrate the spectrum reflects the superconductor density of states. On top of a molecule, the spectrum contains bound-state resonances within the superconducting gap.

The individual steps of the evaluation are sketched in Fig. 2. We process 100 files in less than 13 min, which is documented in a video that is available online at <http://dx.doi.org/10.1016/j.softx.2016.04.001> together with this article. We first have to correct a tiny offset of $38 \mu\text{eV}$ in the bias voltage of all files. The value has been extracted from a fit of the spectrum of the bare substrate to the model of BCS theory [Fig. 2(b)]. After removing the offset, the second step is to recover the correct physical quantities of the lock-in signal. This is achieved by the regauging procedure described in Section 3 [Fig. 2(c)]. The third step is the visualization of the data of the grid. Therefore we overlay the spectral intensity with a topographic image, to get the exact localization of the subgap bound-states [Fig. 2(d)]. In this case we see clearly their increasing intensity on top of the molecule. Finally, all spectra with the regauged data are exported, so that they can be directly imported to other scientific programs for final processing [Fig. 2(e)].

5. Conclusions

SpectraFox is a promising tool to simplify daily tasks in scanning probe microscopy and spectroscopy. It provides data management functionalities for projects with large numbers of measurement files. Once set up, SpectraFox is capable to apply often used numeric procedures simultaneously to large numbers of files, which speeds up the evaluation process. SpectraFox also includes several specialized tools, for instance for non-linear least square fitting or for data visualization. Thereby it is important to note that all data created by SpectraFox is stored separately from the original files. This guarantees the source data never to be modified or rewritten.

SpectraFox is published under an open-source license, which ensures the software to stay independent of any manufacturer of measurement hardware. This also helps to maintain a homogeneous evaluation experience within and between groups which use different measurement equipment. SpectraFox is compatible with several data file formats. By now, it supports the handling of files generated by the systems of SPECS/Nanonis, Omicron/Matrix, Createc, and Nanotec Electronica/WSxM. The interested reader may help to contribute to support more manufacturers in the future by submitting example data to the project's github repository. We furthermore encourage the reader to contribute to the project, and extend the software to his own wishes.

Acknowledgments

We thank K.J. Franke, J.I. Pascual, G. Schulze, N. Hatter, C. Lotze, G. Ahmadi, and B.W. Heinrich for fruitful discussions, suggestions, and feedback to the software, its user manual and this publication. We also acknowledge financial support by the Deutsche Forschungsgemeinschaft through the collaborative research center SFB 658.

References

- [1] Pohl DW, Denk W, Lanz M. *Appl Phys Lett* 1984;44:651.
- [2] Binnig G, Rohrer H. *IBM J Res Dev* 1986;30:355.
- [3] Binnig G, Quate CF, Gerber Ch. *Phys Rev Lett* 1986;56:930.
- [4] Martin Y, Wickramasinghe HK. *Appl Phys Lett* 1987;50:1455.
- [5] Chang AM, Hallen HD, Harriott L, Hess HF, Kao HL, Kwo J, et al. *Appl Phys Lett* 1992;61:1974.
- [6] Wiesendanger R. *Rev Mod Phys* 2009;81:1495.
- [7] Gwyddion, Free SPM data analysis software (<http://gwyddion.net>).
- [8] Horcas I, Fernandez R, Gomez-Rodriguez JM, Colchero J, Gomez-Herrero J, Baro AM. *Rev Sci Instrum* 2007;78:013705.
- [9] Schulz F, Drost R, Hämäläinen SK, Demonchaux T, Seitsonen AP, Liljeroth P. *Phys Rev B* 2014;89:235429.
- [10] Ruby M, Heinrich BW, Pascual JI, Franke KJ. *Phys Rev Lett* 2015;114:157001.
- [11] Ruby M, Pientka F, Peng Y, von Oppen F, Heinrich BW, Franke KJ. *Phys Rev Lett* 2015;115:087001.
- [12] Ruby M, Pientka F, Peng Y, von Oppen F, Heinrich BW, Franke KJ. *Phys Rev Lett* 2015;115:197204.
- [13] Hatter N, Heinrich BW, Ruby M, Pascual JI, Franke KJ. *Nature Comm* 2015;6:8988.
- [14] Fano U. *Phys Rev* 1961;124:1866.
- [15] Frota HO. *Phys Rev B* 1992;45:1096.
- [16] Frank S, Jacob D. *Phys Rev B* 2015;92:235127.
- [17] Gatteschi D, Sessoli R, Villain J. *Molecular nanomagnets*. Oxford: Oxford University Press; 2006.
- [18] Chen CJ. *Introduction to scanning tunneling microscopy*. 2nd ed. Oxford: Oxford University Press; 2008.
- [19] Giessibl FJ. *Rev Mod Phys* 2003;75:949.
- [20] Sader JE, Jarvis SP. *Appl Phys Lett* 2004;84:1801.
- [21] Sader JE, Sugimoto Y. *Appl Phys Lett* 2010;97:043502.
- [22] Nelder JA, Mead R. *Comput J* 1965;7:308.
- [23] Levenberg K. *Q Appl Math* 1944;2:164.
- [24] Marquardt D. *SIAM J Appl Math* 1963;11:431.
- [25] Chen CJ. *Phys Rev B* 1990;42:8841.
- [26] Bardeen J, Cooper LN, Schrieffer JR. *Phys Rev* 1957;108:1175.
- [27] Heinrich AJ, Gupta JA, Lutz CP, Eigler DM. *Science* 2004;306:466.
- [28] Hirjibehedin CF, Lin C-Y, Otte AF, Ternes M, Lutz CP, Jones BA, et al. *Science* 2007;317:1199.
- [29] Shiba H. *Progr Theoret Phys* 1968;40:435.