

**An evolutionary strategy for
complex structure determination
by Low Energy Electron Diffraction**

**Paul Quinn (B.Sc.)
School of Physical Sciences
Dublin City University**

A thesis submitted to



for the degree of
Masters of Science

**Research Supervisor
Dr. A.A. Cafolla**

September 2001

Contents

List of Tables	v
List of Figures	vi
Abstract	viii
Declaration	ix
Acknowledgments	x
Chapter 1 Introduction	1
1.1 Surface Science	1
1.1.1 Surface Crystallography	2
1.1.2 Aim of the Thesis	2
Chapter 2 Experimental and Theoretical Details	4
2.1 Introduction	4
2.2 Basic Diffraction	5
2.3 Experimental Arrangement	8
2.4 LEED as a qualitative probe	9
2.5 LEED as a quantitative probe	11

<i>CONTENTS</i>	iii
2.6 Tensor LEED	15
2.7 Comparison of Theory to Experiment	15
2.8 Surface Composition	19
2.8.1 Auger Electron Spectroscopy	19
2.8.2 X-ray Photoelectron Spectroscopy	21
Chapter 3 LEED study of the Ni(110) and Ni(111) surfaces	24
3.1 Introduction	24
3.2 Experimental Details	24
3.3 Computational Issues	25
3.4 Results and Discussion	26
3.5 Conclusion	35
Chapter 4 An Evolutionary Strategy for LEED	37
4.1 Introduction	37
4.2 Review of Current Global Search Methods	39
4.2.1 Simulated Annealing	39
4.2.2 Genetic Algorithms	40
4.3 Review of Global Search Methods in LEED	42
4.4 Differential Evolution	45
4.4.1 Overview	45
4.4.2 Population Initialisation	46
4.4.3 Mutation	47
4.4.4 Crossover	47
4.4.5 Selection	48
4.5 Ir(110)-(2x1) Missing Row - Row Pairing Model	48

<i>CONTENTS</i>	iv
4.6 Results	50
4.7 Discussion	55
4.8 Conclusion	57
Chapter 5 Summary and Conclusion	59
5.1 Summary	59

List of Tables

3.1	Best-fit interplanar spacings and Pendry R-factors, R_p obtained for the Ni(110) and Ni(111) surfaces. d_{12}, d_{23} and d_{34} represent the distances from the first to second, second to third and third to fourth Ni layers respectively.	32
3.2	Comparison of the results of this study of the Ni(110) surface with previous studies of this surface. Δd represents the difference between the interlayer spacing found from experiment and the bulk Ni(110) value.	33
4.1	Example of the crossover procedure used in Genetic Algorithms The chromosome fragments printed in italics and bold letters in the Nth generation recombine to form the italic and bold printed chromosomes in the Nth+1 generation Note that mutation has occurred in the formation of the Nth+1 italic chromosome	42
4.2	Comparison of search strategies as applied to the Ir(110)(2x1) surface over a $1\text{\AA} \times 1\text{\AA} \times 1\text{\AA}$ parameter space	55

List of Figures

2.1	Diffraction from a one dimensional chain of atoms	5
2.2	Intensity profile resulting from diffraction from a 1-D atom chain .	6
2.3	Digram of the change in wavevector Δk in the diffraction process	7
2.4	Experimental setup for a LEED	9
2.5	Diffraction from a bulk crystal. Note how scattering from a plane produces both a reflected (backscattered) and transmitted (forward scattered) beam set.	13
2.6	Energetics of the Auger process	20
2.7	An example of the photoemission process, where an incident X-ray photon transfers its energy to a 1s core level electron	22
3.1	Graph of the best fit theory to expt IV curves for the (1,0), (0,1) and (1,1) diffraction beams for the Ni(110) surface	27
3.2	Graph of the best fit theory to expt IV curves for the (0,2),(2,0) and (3,0) diffraction beams for the Ni(110) surface	28
3.3	Graph of the best fit theory to expt IV curves for the (2,1),(1,2) and (1,3) diffraction beams for the Ni(110) surface	29
3.4	Graph of the best fit theory to expt IV curves for the (1,1),(2,0) and (0,2) diffraction beams for the Ni(111) surface	30

3.5	Graph of the best fit theory to expt IV curves for the (1,1),(2,0) and (0,2) diffraction beams for the Ni(111) surface	31
3.6	Graph of the raw data for the (1,0)	34
3.7	Graph of the raw data for the (0,1)	34
4.1	A general overview of the Simulated Annealing algorithm	41
4.2	A general overview of the Differential Evolution algorithm	46
4.3	Schematic Diagram of the mutation and crossover processes in the Differential Evolution algorithm. (a) A mutant vector i is formed by choosing three random target vectors and adding the weighted difference of two of them to a third vector (b) A trial vector is generated by swapping the parameters in the target vector with parameters in the mutant vector. Swapping is determined by the crossover probability CR. A random number is generated for each parameter in the vector. If this number is \leq CR, then the parameters are swapped	49
4.4	A top view of the Ir(2x1) surface	51
4.5	Histogram resulting from DE method (1) with a uniform initial population	52
4.6	Histogram resulting from DE method (2) with a uniform initial population	53
4.7	Histogram resulting from DE method (2) with a gaussian deviate initial parameter population	54

Abstract

Low Energy Electron Diffraction (LEED) has been used extensively in the study of surface structures. Structures are typically determined in a trial and error process whereby theoretical calculations for an initially proposed structure are refined until a good fit to experimentally measured data is found. In this thesis LEED has been used to study the structure of the clean Ni(110) and Ni(111) surfaces. For the clean Ni(110) surface a 7% relaxation in the outermost Ni layer spacing is found compared to 0% for the Ni(111) surface. The variation in surface relaxation is explained in terms of the reduced valence electron density and atomic packing density at the surface. The trial and error approach to determining structures is not only time consuming but can also be highly inaccurate. An automated approach to structure determination using an Evolutionary Strategy called Differential Evolution is developed in this work. The algorithm is found to be extremely fast and reliable and provides a 25-37% increase in speed compared to the best competing method.

Declaration

I hereby certify that this material, which I now submit for assessment on the programme of study leading to the award of Masters of Science is entirely my own work and has not been taken from the work of others save and to the extent that such work has been cited and acknowledged within the text of my own work.

Signed: *Paul Chen*

ID Number: 96970855

Date: *13/9/01*

Acknowledgments

First and foremost I would like to thank my supervisor, Tony Cafolla. He has an infectious enthusiasm for all things scientific which to this day inspires me to get off my ass and learn new things. Thanks. Next I would like to thank Eilish for patiently listening to my rants and putting up with me. I thank her for her continual support and encouragement. You're a good friend. Philip Ryan, Darren Walsh, John Moloney and Mark Gibson all get a mention for forcing me to drink large quantities of ale for their amusement. The lecturers at DCU, in particular Greg Hughes and Brian Lawless, have been a great help along the way. I would like to thank my grandmother, Christine, my parents, Noeleen and Paul, and my brothers, Robert and Graham for their support and encouragement. Last but not least I'd like to dedicate this thesis to Sinead and Julie. Thanks for the love and laughter.

Chapter 1

Introduction

1.1 Surface Science

One of the most fundamental properties of any material is its structure. In the study of bulk crystals the determination of their bulk structure is essential to an understanding of the electronic and chemical properties of the material. At the surface of any material the forces and stresses on the atoms are different from those in the bulk. This generally results in a different structure at the surface than that in bulk which in turn affects the chemical and electronic properties at the surface. Surface Science is the study of how a material interacts with its environment through its surface and the study of the structure and electronic properties of the material at its surface. Surface Science is a very broad discipline and some of the many areas being investigated include heterogeneous catalysis, microelectronics, adhesion, lubrication, corrosion, coatings, solid-solid and solid-liquid interfaces.

1.1.1 Surface Crystallography

Several experimental techniques exist to study surface structure. Among them are Ion scattering, X-ray Standing Waves (XSW) and Surface X-ray diffraction but by far the most common technique is Low Energy Electron Diffraction (LEED). LEEDs popularity is due to a number of facts.

- It is simple to use, and by UHV standards a cheap technique. Other techniques tend to require specialised equipment such as synchrotron radiation sources, ion sources, or high voltage electron guns.
- LEED theory is well developed
- It is an extremely flexible technique that can be applied to the study of both metal and semiconductor surfaces.
- It can be used to examine growth, periodicity and order on a surface at a superficial level or in conjunction with computational models it can be used to determine bond lengths and atomic arrangement.

1.1.2 Aim of the Thesis

In this thesis the basic background to LEED will be explained and the accuracy of the theory will be demonstrated through the application of LEED to study of the clean Ni(110) and the Ni(111) surfaces. LEED theory is extremely well developed [1,2] but the study of complex structural models is complicated by the possibility that several solutions can exist with only one solution being the correct one. This problem will be addressed with the implementation of an Evolutionary Strategy [3] called Differential Evolution [4]. Results will be presented on the algorithms performance and compared to alternate algorithms.

References

- [1] M.A. VanHove, W.H. Weinberg,C.M. Chan *Low-Energy-Electron-Diffraction* Springer Verlag,(1986)
- [2] M.A. Van Hove, S.Y. Tong *Surface Crystallography by LEED* Springer-Verlag (1979)
- [3] D.E. Goldberg *Genetic Algorithms is search,optimization and machine learning* Addison-Wesley (1989)
- [4] R.Storn, K. Price, *Jn. Global Optimization* **11** (1997) 341

Chapter 2

Experimental and Theoretical Details

2.1 Introduction

The structure at the surface of a material can be significantly different to the bulk. The atoms at the surface experience appreciably different forces from those in the bulk due the reduction in symmetry and loss of neighbouring atoms. The structural changes that result from this reduction in coordination can be contractions or expansions of the surface layers, referred to as relaxations or they can be a complete restructuring of the surface plane referred to as reconstructions.

Low Energy Electron diffraction is one of the oldest diffraction techniques applied to the study of surfaces dating back to the experiments of Davisson and Germer in 1927. LEED is used in two basic ways, firstly as a qualitative probe of surface order and symmetry, and secondly as a quantitative tool for determining exact surface structure. LEED has become the technique to beat for surface

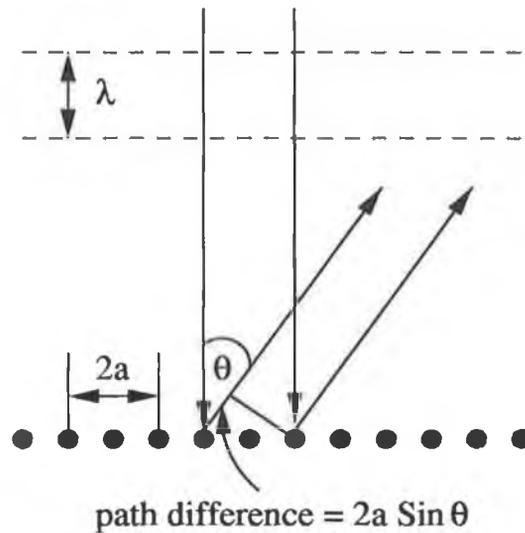


Figure 2.1: Diffraction from a one dimensional chain of atoms

structure determination, with currently over 500 surface structures solved with this technique alone. One of the biggest successes of LEED theory was the elucidation of the structure of the infamous Si(111)-(7x7) surface [2]. The DAS (Dimer-Adatom-Stacking Fault) model proposed by (Van Hove, Tong et al) was later vindicated with the arrival of the STM, which although not measuring bond lengths did reveal the basic structure predicted by LEED to be correct. This chapter will start with an introduction to simple diffraction in one and two dimensions and how this applies to LEED. The method by which exact structural information can be obtained from the diffraction process will then be outlined.

2.2 Basic Diffraction

Consider a one-dimensional chain of atoms separated by a distance a , with an electron beam incident on them as shown in figure 2.1. Consider the backscatter-

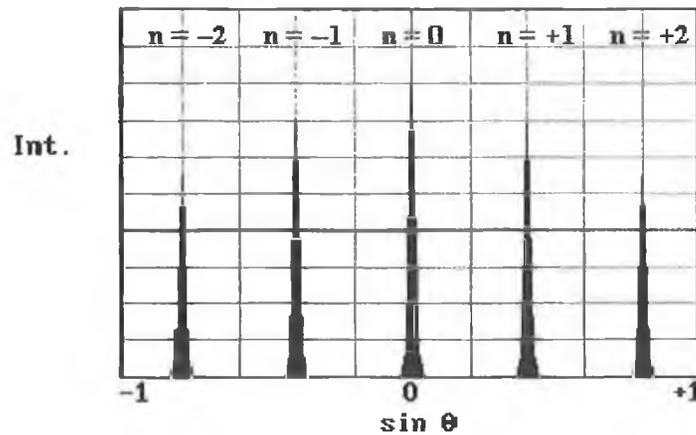


Figure 2.2: Intensity profile resulting from diffraction from a 1-D atom chain

ing wavefront from the second nearest neighbour¹ atoms at a well defined angle θ to the surface normal then there is a path difference (pd) in the distance the radiation has to travel from the scattering centres to the detector. The size of this path difference is $2a \sin \theta$ and must be equal to an integral order of wavelengths for constructive interference to occur i.e.,

$$pd = 2a \sin \theta = h\lambda \quad (2.1)$$

where λ is the wavelength of the incident electron beam and h is an integer. The intensity profile produced by such a diffraction condition is shown in figure 2.2.

We note from this pattern that the pattern is symmetric about $\theta=0$ and that $\sin \theta$ is inversely proportional to a . Given this inverse proportionality it is useful to consider the diffraction process again by examining the wavevector of the incident beam. The incident electron beam has a wave vector k such that,

$$|k| = \frac{2\pi}{\lambda} \quad (2.2)$$

¹The choice of second nearest neighbour is simply to keep the arguments clearer later on

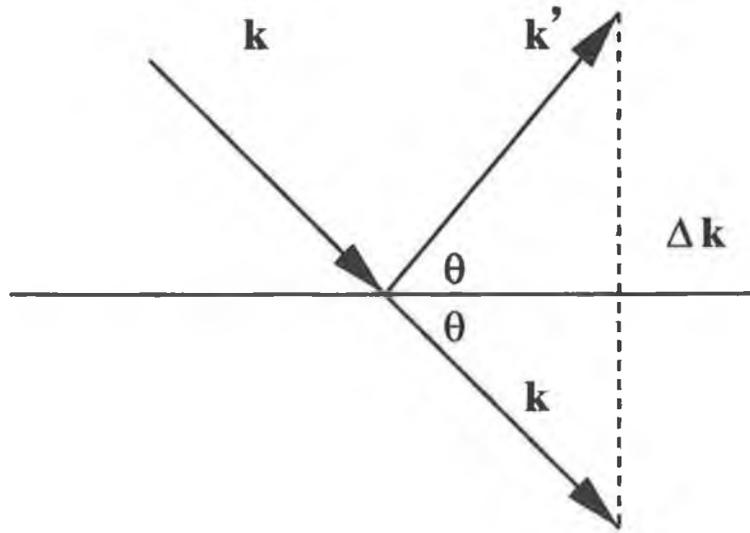


Figure 2.3: Digram of the change in wavevector Δk in the diffraction process

substituting for λ in equation 2.1 gives,

$$2|k| \sin \theta = \frac{2\pi}{a} h \quad (2.3)$$

But what does $2|k| \sin \theta$ represent? In figure 2.3 the backscattering of an electron beam is described. The length $2|k| \sin \theta$ represents a change in the direction of the wavevector. So diffraction occurs when the wavevector undergoes a change Δk such that ,

$$\Delta k = \frac{2\pi}{a} h \quad (2.4)$$

In two dimensions this becomes

$$\Delta k = \frac{2\pi}{a} h + \frac{2\pi}{b} k \quad (2.5)$$

where h and k are integers and b is the second dimension of the 2D lattice. A reciprocal lattice can be constructed with vectors a^* , b^* which relate to the real-

space vectors a, b by

$$a^* = \frac{2\pi}{a} \quad (2.6)$$

$$b^* = \frac{2\pi}{b} \quad (2.7)$$

The translational vector of the reciprocal lattice, G_{hk} , is

$$G_{hk} = ha^* + kb^* \quad (2.8)$$

which implies that

$$\Delta k = G_{hk} \quad (2.9)$$

or that every point on the reciprocal lattice corresponds to a diffraction maxima.

2.3 Experimental Arrangement

The basic experimental setup for performing a LEED experiment is shown in figure 2.4. An electron beam with a well defined energy is produced by an electron gun and is incident on a sample. The electrons are then backscattered from the sample surface onto a system of grids surrounding the LEED gun. The backscattered electrons will be of two types, elasticity scattered and inelastically scattered. The elastically scattered electrons are those which possess purely diffraction information and have not been influenced by phonons, plasmons etc. The inelastically scattered electrons which have lost energy due to electron-electron, phonon, and plasmon interactions, constitute the majority of the backscattered flux. These electron are removed by the grids as follows. After reaching the first grid G1 which is earthed the electrons are accelerated toward the fluorescent screen S, which has a 5kV potential on it. The grids G2 and G3 are held at a

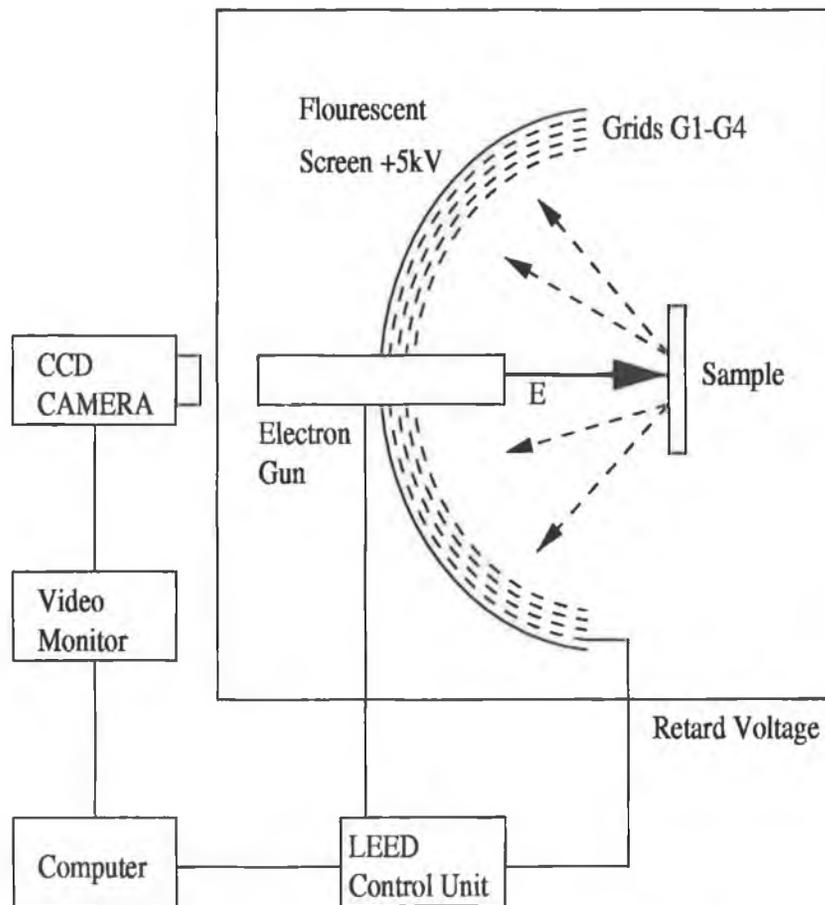


Figure 2.4: Experimental setup for a LEED

variable negative potential which is used to reject the majority of the inelastic electrons. The LEED pattern which is produced on the fluorescent screen is recorded using either a video or camera.

2.4 LEED as a qualitative probe

As shown in section 2.2 the diffraction pattern produced from a surface is a (scaled) representation of the reciprocal lattice. Therefore by examining these

diffraction spots, information on the symmetry and order of the surface can be obtained. The simple analysis of spot positions and spot symmetry is by far the most common use of LEED. In sample preparation under UHV conditions a sharp LEED pattern is regarded as indicating an atomically ordered clean surface. The LEED spots can show whether the surface is reconstructed and may also indicate the presence of steps on a surface. When the surface of a material differs from that of the bulk either through reconstructions or the presence of an adsorbate on the surface, it is convenient to label this difference between the bulk and surface layers. Given the bulk vectors \hat{a}_b and \hat{b}_b the surface vectors \hat{a}_s, \hat{b}_s can be related to the bulk (or substrate) vectors by

$$\hat{a}_s = M_{11}\hat{a}_b + M_{12}\hat{b}_b \quad (2.10)$$

$$\hat{b}_s = M_{21}\hat{a}_b + M_{22}\hat{b}_b \quad (2.11)$$

where M_{ij} are coefficients of a matrix,

$$\mathbf{M} = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix} \quad (2.12)$$

This matrix notation is useful because the determinant of the \mathbf{M} is the ratio of the surface mesh to the bulk mesh and provides a classification scheme for surface structures, as follows

- (a) If the determinant M is an integral and all matrix components are integral, the two meshes are simply related.
- (b) If the determinant M is a rational fraction, the true surface mesh is larger than either the bulk or apparent surface mesh. The size of the true surface mesh is dictated by the distance over which the two meshes come into coincidence. This is sometimes referred to as a coincidence lattice.

- (b) If the determinant M is irrational. Then the two meshes are incommensurate and no true surface mesh exists.

A more popular, but less versatile, representation of surface meshes is Wood notation. Wood notation defines the ratio of the lengths of the unit meshes together with the angle through which they are rotated. For an adsorbate A on the hkl surface of a material X having translation vectors such that

$\hat{a}_s = m\hat{a}_b$ and $\hat{b}_s = n\hat{b}_b$ with a unit mesh rotation of ϕ the structure is represented in wood notation as,

$$X(hkl) - (m \times n) - R\phi - A \quad (2.13)$$

where $R\phi$ indicates that the surface mesh is rotated through an angle ϕ with respect to the bulk mesh.

2.5 LEED as a quantitative probe

The location of beams in the diffraction pattern gives general information about the surface but for exact structural information a study of the intensity of the diffracted beams as a function of the incident beam energy must be performed. The Intensity vs. Energy curves are commonly referred to as IV curves. When an electron is incident on a crystal, the strong potential fields scatter the electron. The electron can be forward scattered, that is into the crystal or it can be backscattered which reverses its direction away from the crystal. LEED relies heavily on backscattering, as this is obviously what is being observed on the fluorescent screen, but it is the forward scattering that really affects the LEED intensities. If forward scattering were weak compared to backscattering the result would be a diffraction pattern and an intensity Vs energy profile which result from the simple

Bragg condition. However if forward scattering is strong an incident electron beam will be scattered into several strong forward travelling beams as well as the incident beam. From this simple argument we see that strong forward scattering will have the effect of producing additional peaks other than the Bragg peaks and could also effect the Bragg peak positions due to interference affects. It is the latter case of strong forward scattering that dominates in LEED and as a result a simple analysis of the intensities of the diffracted beams with energy will not yield any useful information regarding surface structure. Consider an electron beam incident on the first layer at the surface of a crystal. When the beam is scattered , interference between the scattered waves will result in a set of diffracted beams at points determined by the reciprocal lattice,

$$g_{hk} = ha_1 + ka_2 \quad (2.14)$$

One set of diffracted beams will be propagating backward (the reflection set) another will be propagated forward with the unscattered incidence beam (the transmission set). The wavefield incident on the second layer will now be comprised of the forward scattered beams from the first layer, which will be subsequently forward and backscattered by the second layer. The process is indicated in figure 2.5.

The total backscattered intensity therefore comprises of backscattered beams from several surface layers. Any diffracted beam s_{nm} will pick up contributions from from the scattering of each of member of the beam set from each layer. Any diffracted beam s'_i can be described by the sum of the scattering from each plane of an incident beam,

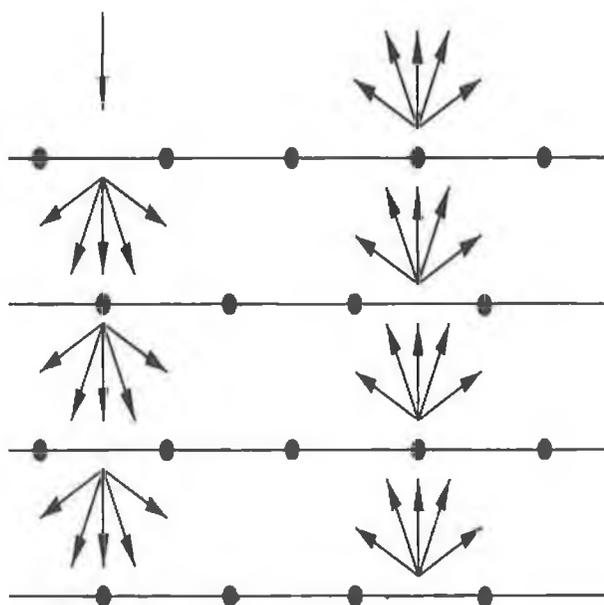


Figure 2.5: Diffraction from a bulk crystal. Note how scattering from a plane produces both a reflected (backscattered) and transmitted (forward scattered) beam set.

$$s'_i = m_1 \cdot s_1 + m_2 \cdot s_2 + \cdots + m_N \cdot s_N \quad (2.15)$$

Each element of the matrix m , referred to as the scattering matrix, describes the change in both the amplitude and phase a given layer causes to a incident beam. The changes are represented by complex numbers with the real part describing the amplitude change and the imaginary parts describing the phase changes. Therefore given an incident beam set S , and a diffracted beam set S' , these may be related by a square matrix M , the scattering matrix, such that

$$S' = M \cdot S \quad (2.16)$$

Consequently a complete description of the scattering from a plane of periodic scatterers may be given by two complex($N \times N$) matrices, one describing the 'reflected set' and the other describing the 'transmitted' set.

When an electron enters a crystal it gains energy due to the potential inside the crystal referred to as the 'inner potential'. The size of the inner potential is approximately equal to the sum of the work function and fermi energy. This effectively shifts the energy of the diffraction peaks in energy. In addition to this energy shift electrons incident on a crystal also suffer energy losses due to scattering from plasmons and phonons in the crystal. The general effect of this inelastic scattering is the exponential decay of the beam intensity as it travels into the crystal. This results in a broadening of the diffraction peaks. This energy loss is represented by introducing an imaginary component V_{oi} to the inner potential, typically equal to 4eV for most materials. As seen above LEED diffraction spectra cannot be described in terms of a single scattering model and in order to correctly interpret the LEED IV curves we need to use dynamic or multiple scattering calculations.

2.6 Tensor LEED

There are several techniques and methods used for calculating scattering matrices and summing the forward and backscattered electron beams, explanations of which would be beyond the scope of this thesis. These calculations usually take a lot of CPU time, e.g. anything from 5mins to 24hrs depending on the complexity of the structural model. Tensor LEED was developed in an effort to reduce the computational effort involved by allowing the user to examine several structural models which are simply variations on some base model [5]. Tensor LEED examines how the scattering matrices of a basic model structure are affected by small changes in positions of the atoms in the structure. A knowledge of how the scattering matrices vary with atomic movement allows approximate IV curves to be generated for a range of structures. The approximated IV curves are usually only valid for atomic movements within 0.4\AA of the base structure.

2.7 Comparison of Theory to Experiment

The IV spectra resulting from different structural models are usually significantly different and so visual comparison of the theoretical and experimental IV curves is sufficient to distinguish between correct and incorrect models. However when fine tuning the parameters to obtain an accurate structural model this method is quite inadequate. Jona and Zanazzi [7] were the first to tackle this problem with the introduction of the so called reliability or R-factor, which is a measure of the discrepancy between experimental and theoretical IV beams. The R-factor was proposed as a common guide to the quality of a structural model. The first R-factors used were adapted from those used in x-ray diffraction analysis, but these

did not apply well to LEED spectra. However several other R-factors were later proposed, most notably by Pendry [9] and Van Hove [7] which although equally valid, undermined the universality that Jona proposed. The various R-factors proposed emphasise different aspects of the match between the experimental and theoretical IV curves. The basis of the Zanazzi-Jona R-factor, R_{zj} , is to emphasise peak position rather than peak heights, because peak positions are highly structurally dependent whereas peak heights are due to non-structural parameters such as vibrational amplitudes etc. The Pendry R-factor, R_p , probably the most popular R-factor, emphasises peak positions but also emphasises the importance of overlapping peaks and points where the intensity falls to zero in the spectra. Zero points are obviously important as they are a result of destructive interference and as such are related to the structure of the crystal. A comparison of the two methods carried out by Clarke [10], showed that the Zanazzi-Jona R-factor is much less sensitive than the Pendry R-factor to structural change. He showed that a 1 percent change in a single structural parameter leads to a 1.6% change in R_{zj} and a 4.8% change in R_p . He also found that a 1% change in the inner potential changed R_{zj} by 1% and R_p by 0.5%. The requirements of a good R-factor are that it is chiefly sensitive to peak positions. It should not be at all sensitive to absolute intensities, but should pay some attention to relative intensities especially features that are close in energy. It should also be a simple function not requiring more than a first derivative. Pendry deals with this problem as follows. He first assumes a LEED IV curve to be composed of a series of Lorentzian peaks located at various energies E_j , and of various various intensities a_j , such that

$$I(E) \simeq \sum \frac{a_j}{(E - E_j)^2 + V_{oi}^2} \quad (2.17)$$

Their widths are dictated by the imaginary part of the electron self energy V_{oi} . In most materials V_{oi} is approximately independent of energy above 30eV, and takes a value of approximately -4eV. For a series of roughly spaced peaks, the criterion of insensitivity to the amplitudes is met by using the logarithmic derivative

$$L(E) = \frac{1}{I} \frac{dI}{dE} \quad (2.18)$$

which from equation 2.17 reduces to

$$L(E) \simeq \sum \frac{-2(E - E_j)}{(E - E_j)^2 + V_{oi}^2} \quad (2.19)$$

The logarithmic derivative, L , has removed the relative intensities of the Lorentzian peaks, a_j , and in their place there are peaks of amplitude,

$$L = \pm 1/|V_{oi}| \quad \text{at} \quad E = E_j \pm |V_{oi}| \quad (2.20)$$

This completely eliminates the relative amplitudes of the peaks and if peaks overlap to some extent then their relative intensities do affect L . This works well except when the intensity, I , is zero, so a new function is defined which gives equal emphasis to intensity zeros and Lorentzian peaks. This new function, Y , which shall be referred to as a modified logarithmic derivative, takes the form,

$$Y(E) = L^{-1}/(L^{-2} + V_{oi}^2) \quad (2.21)$$

It can be shown that $Y(E)$ takes a maximum value of

$$Y_{max} = \pm \frac{1}{2}|V_{oi}| \quad \text{when} \quad L = \pm 1/|V_{oi}| \quad (2.22)$$

Which for a series of lorentzian peaks occurs when

$$E = E_j \pm |V_{oi}| \quad (2.23)$$

The new reliability factor R_p is defined as,

$$R_p = \frac{\int (Y_e - Y_t)^2 dE}{\int (Y_e^2 + Y_t^2)} \quad (2.24)$$

where Y_e and Y_t are the modified logarithmic derivatives of the experimental and theoretical curves respectively. Pendry also proposed a method for estimating the error in the R-factor, R_p and the error in the best model. The estimate of the error comes from a general statistical treatment of a LEED IV curve. Given an IV curve with N well separated peaks, the ratio of the variance ($var R_N$) to the mean R-value (R_N) is,

$$\frac{var R_N}{R_N} \propto \frac{1}{\sqrt{N}} \quad (2.25)$$

This is obviously a reasonable assumption to make, i.e. increasing the number of features ,(or IV curves) in an R-factor analysis will increase the accuracy of the result. Given an IV curve of energy range δE there can be N well resolved peaks where,

$$N = \frac{\delta E}{2|V_{oi}|} \quad (2.26)$$

This permits us to define a double reliability factor, RR , such that

$$RR = \frac{var R}{\bar{R}} = \sqrt{\frac{8|V_{oi}|}{\delta E}} \quad (2.27)$$

The statistical error in the R-factor minimum is then,

$$\text{var} R_{min} = RR \times R_{min} \quad (2.28)$$

2.8 Surface Composition

Surface structure analysis is primarily concerned with the study of clean surfaces or the effects of an adsorbate on the surface. It is therefore essential to be able to verify that a surface is free of contaminants and to be able to measure the exact amount of a material adsorbed on the surface. The techniques used in surface science to study surface composition are Auger Electron Spectroscopy, and X-ray Photoelectron Spectroscopy. These techniques will be outlined in the following sections.

2.8.1 Auger Electron Spectroscopy

The basic Auger process starts with an incident electron (or photon) causing ionisation of a core level electron. The electron vacancy or hole created in the core level may be neutralised by an electron transition from an electron level of lower binding energy. A quantum of energy, (ΔE) , the difference between the core hole and the electron falling into it is transferred to a third electron which escapes as an Auger electron. The process is described graphically in figure 2.6. The kinetic energy of an Auger electron for the process shown in figure 2.6 is ,

$$E_{kin} = E_k - E_{L_1} - E_{L_{2,3}} \quad (2.29)$$

$E_k - E_{L_1}$ is the quantum of energy released by an electron falling from the L_1 shell to the core hole in the K shell and $E_{L_{2,3}}$ is the binding energy of the electron in the $L_{2,3}$ shell. This Auger transition is assigned the term KL_1, L_{23} . The

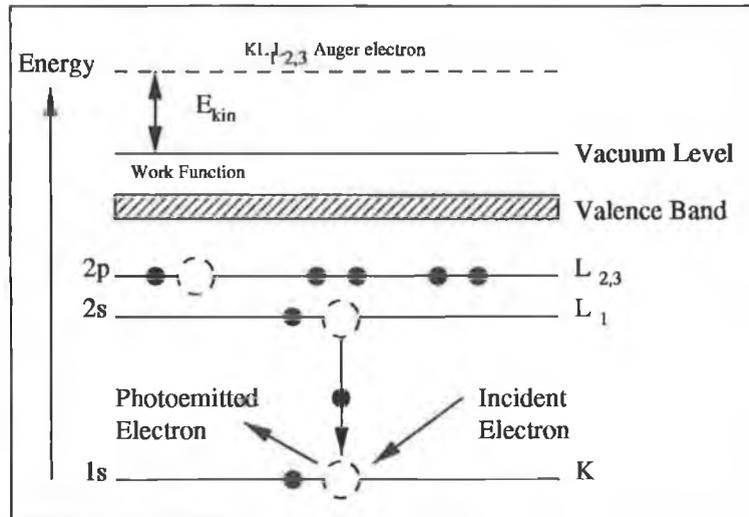


Figure 2.6: Energetics of the Auger process

probability of relaxation by Auger emission is the dominant process for core holes with binding energies below about 2keV for K shell ionisation. The kinetic energy of the Auger electron is independent of the energy of the ionising particle or the incident radiation giving rise to the initial core hole. The kinetic energy of the Auger electron is characteristic solely of the binding energies of the electrons within the atom. Hence, Auger electrons may be used for chemical identification.

Auger excitation is usually carried out using electron sources due to the relative ease of producing sufficiently energetic beams of high intensity. The use of electron beams to produce the initial ionisation process is advantageous in that the incident beam can be focused thus giving good spatial resolution. For the light elements (atomic number $Z < 20$) Auger emission is more probable than X-ray emission for a K-shell initial-state hole and for $Z < 15$ it is almost the exclusive process. For higher Z , Auger processes dominate for initial state holes in outer shells.

2.8.2 X-ray Photoelectron Spectroscopy

Photoelectron spectroscopy is, in principle, a particularly simple process. A photon of energy $h\nu$ (h is planks constant and ν is the frequency) penetrates the surface and is absorbed by an electron with a binding energy E_B below the vacuum level, which then emerges with a kinetic energy E_{kin} ,

$$E_{kin} = \nu - E_B - \phi \quad (2.30)$$

The process is described in figure 2.7. The kinetic energies of the ejected electrons have an energy spread characteristic of the binding energies and as a result can be used to identify atomic species. Any photon whose energy exceeds the work function of a solid can be used for photoelectron spectroscopy but the most common sources used in the lab are the K_α lines of Al and Mg at 1486.6 and 1253.6 eV respectively.

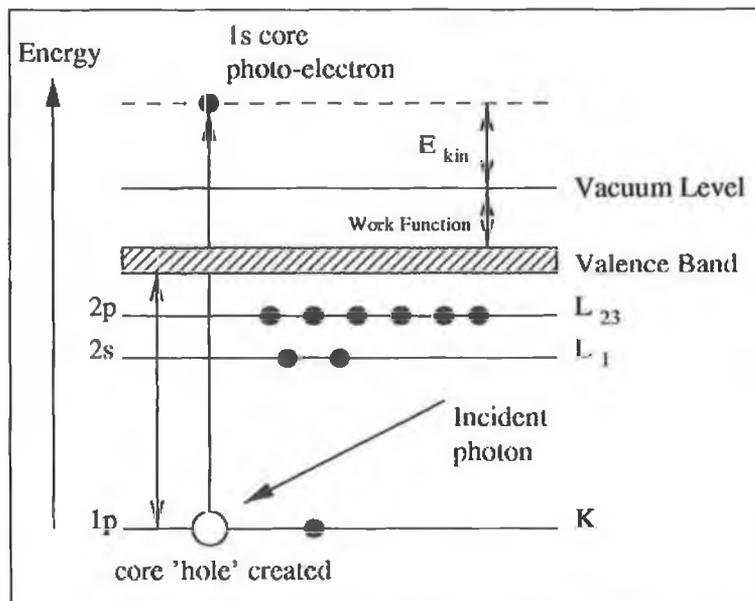


Figure 2.7: An example of the photoemission process, where an incident X-ray photon transfers its energy to a 1s core level electron

References

- [1] Wood E.A. *J. Appl. Phys.* **35** (1964) 1306
- [2] M. Prutton *Introduction to surface physics* Oxford University Press
- [3] D.P. Woodruff, T. A. Delchar *Modern techniques of surface science* Cambridge University Press
- [4] J.C. Vickerman ,*Surface Analysis - The Principal Techniques* Wiley
- [5] P.J. Rous *Prog.Surf.Sci.*,**39** (1992) 3
- [6] Luth Hans *Surfaces and Interfaces of Solids* Springer-Verlag (1993)
- [7] M.A. VanHove, W.H. Weinberg,C.M. Chan *Low-Energy-Electron-Diffraction* Springer Verlag,(1986)
- [8] M.A. Van Hove, S.Y. Tong *Surface Crystallography by LEED* Springer-Verlag (1979)
- [9] J.B. Pendry, *Low Energy Electron Diffraction* Academic Press (1974)
- [10] L.J. Clarke, *Surface Crystallography : An introduction to Low Energy Electron Diffraction* Wiley (1985)

Chapter 3

LEED study of the Ni(110) and Ni(111) surfaces

3.1 Introduction

In this chapter experimental LEED data taken for the Ni(110) and Ni(111) surfaces is compared to theory. A detailed analysis of the structure of clean Ni surfaces is particularly important because it provides a basis for studies of the LEED spectra of chemically adsorbed ordered overlayers of atoms or molecules on Ni single crystal surfaces. I will first outline the experimental and computational issues pertinent to the analysis and the results of the structural analysis will then be presented and discussed.

3.2 Experimental Details

The experiments were conducted using a standard ultra high vacuum chamber equipped with a range of facilities for sample preparation and surface charac-

terisation together with a computer controlled LEED diffractometer. The base pressure of the chamber was typically 1×10^{-10} mbar. The Ni crystals were cut by spark erosion from a crystal initially orientated by X-ray Laue alignment. The surface was then polished using progressively finer grades of diamond paste to produce a mirror finish. After insertion into vacuum the crystals were cleaned and polishing damage was removed by repeated cycles of sputtering with 1keV Ar^+ ions and subsequent annealing to 900K. The temperature was monitored using a chromel-alumel thermocouple in contact with the sample. The cleaning cycles were repeated until no carbon, oxygen, or sulphur (sulphur is used in the growth of Ni crystals) were detectable using x-ray photoelectron spectroscopy and the LEED indicated a sharp 1×1 diffraction pattern.

Quantitative LEED intensities were recorded from 60eV to 500eV at room temperature using the Omicron LEEDstar video system. Normal incidence was verified by comparison of symmetry equivalent diffraction beams. The intensities were accumulated over nine runs to improve the signal to noise ratio. For the Ni(111) surface the IV curves for 17 diffraction beams were measured. The symmetry equivalent beams were then averaged reducing the beam set to five symmetry inequivalent beams, namely the (0,1), (1,0), (2,0), (0,2) and (1,1). For the Ni(110) surface 25 beams were collected which were averaged to produce a reduced beam set of nine symmetry inequivalent beams the (0,1), (1,0), (1,1), (0,2), (2,0), (2,1), (1,2), (3,0), and (3,1)

3.3 Computational Issues

The LEED calculations were performed on a Sun UltraSparc1 using the symmetrized automated tensor LEED (SATLEED) package [9, 10]. The phase shifts

were calculated using the Barbieri/Van Hove phase shift package [9, 10]. Different scattering phase shifts were generated for both the Ni(111) and Ni(110) surfaces. Nine phase shifts were used in the calculations. The number of phase shifts used depends on two quantities. Firstly the energy range used; The higher the energy the more phase shifts have to be included in the calculation. Secondly the higher the atomic number the larger the number of phase shifts required. Previous studies on the Nickel surfaces have been performed with 5 phase shifts so 9 is more than adequate. The bulk Debye temperature was taken to be 450K. The real and imaginary parts of the inner potential were initially assumed to be $V_{0r} = 10\text{eV}$ and $V_{0i} = -4\text{eV}$. The real and imaginary parts of the inner potential were then adjusted as part of the optimisation procedure. LEED intensities were calculated at 2eV intervals. The interlayer spacing of the top three layers for the Ni(110) and Ni(111) surfaces were allowed to vary as part of the optimisation procedure.

3.4 Results and Discussion

Plotted in Figures 3.1 to 3.5 are the best fits of theory to experiment for the Ni(110) and Ni(111) surfaces.

The fitting of experiment to theory was performed using the Davidson Fletcher Powell search algorithm. The algorithm was started at several random points in the parameter space to ensure that the minimum found was the true minimum. The quality of the fit between experiment and theory was evaluated using the Pendry R-factor, the properties of which are explained in chapter 2. The interlayer spacing determined from the fits together with the overall R-factor are displayed in table 3.1. The Debye temperature was optimised at 435K and the imaginary part of the inner potential was found to be -3.75eV. The real part of the inner potential

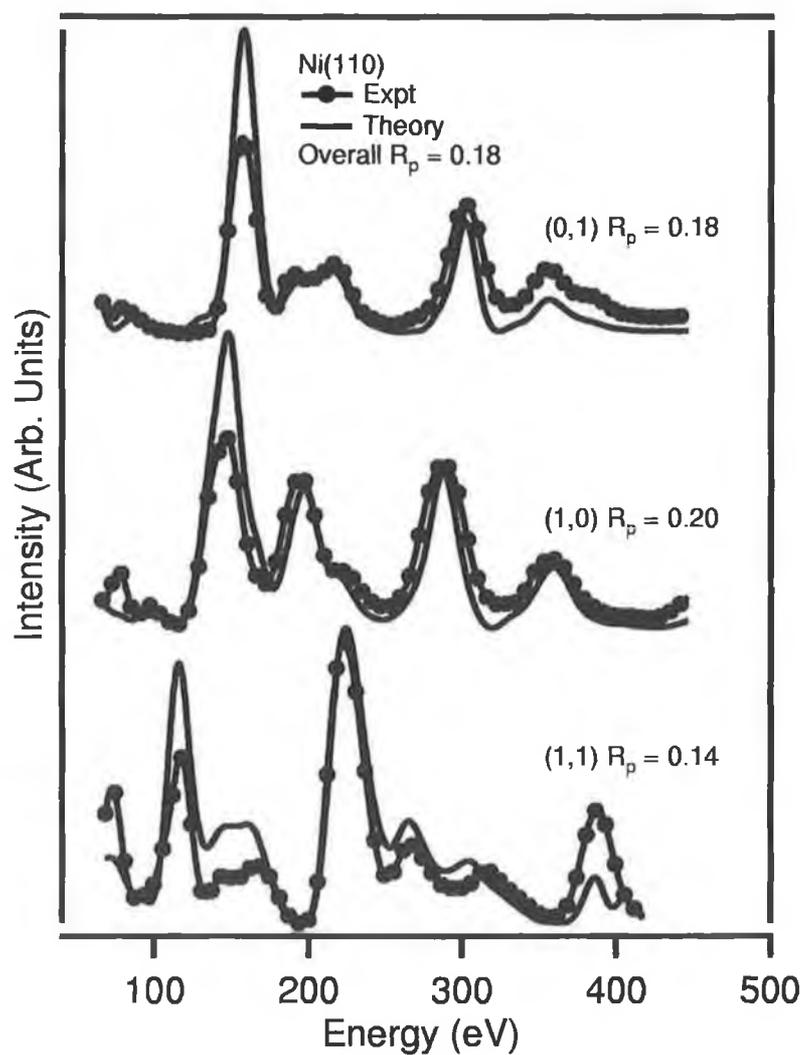


Figure 3.1: Graph of the best fit theory to expt IV curves for the (1,0), (0,1) and (1,1) diffraction beams for the Ni(110) surface

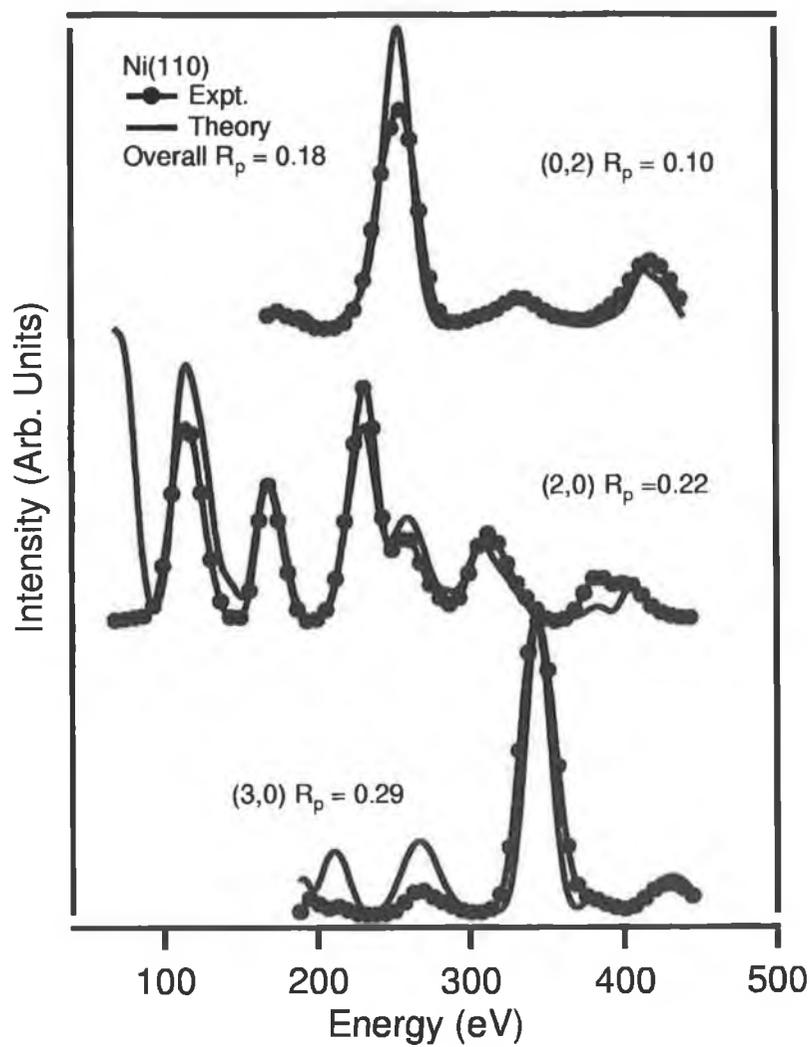


Figure 3.2: Graph of the best fit theory to expt IV curves for the (0,2),(2,0) and (3,0) diffraction beams for the Ni(110) surface

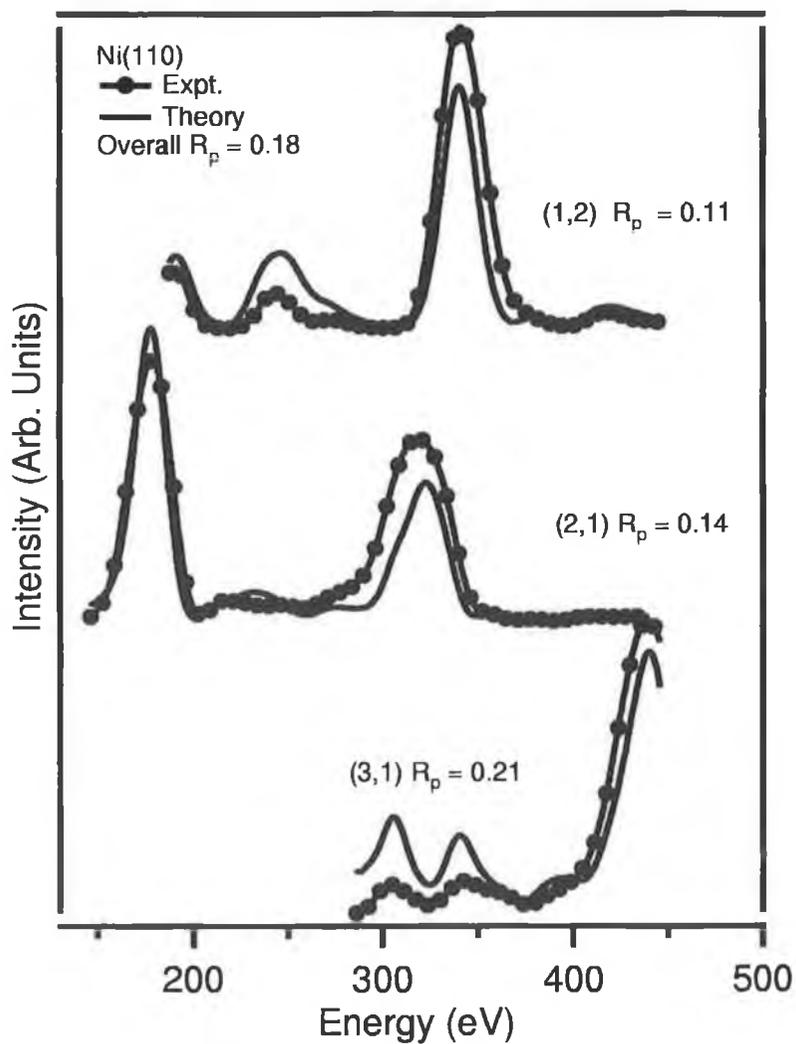


Figure 3.3: Graph of the best fit theory to expt IV curves for the (2,1),(1,2) and (1,3) diffraction beams for the Ni(110) surface

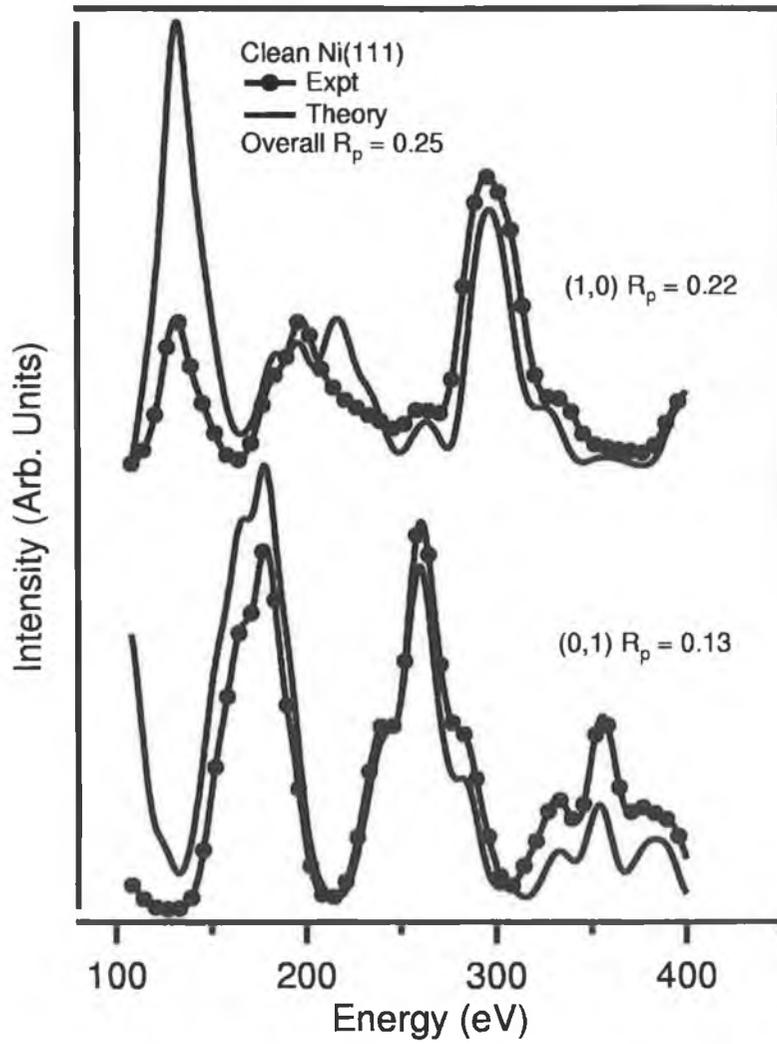


Figure 3.4: Graph of the best fit theory to expt IV curves for the (1,1),(2,0) and (0,2) diffraction beams for the Ni(111) surface

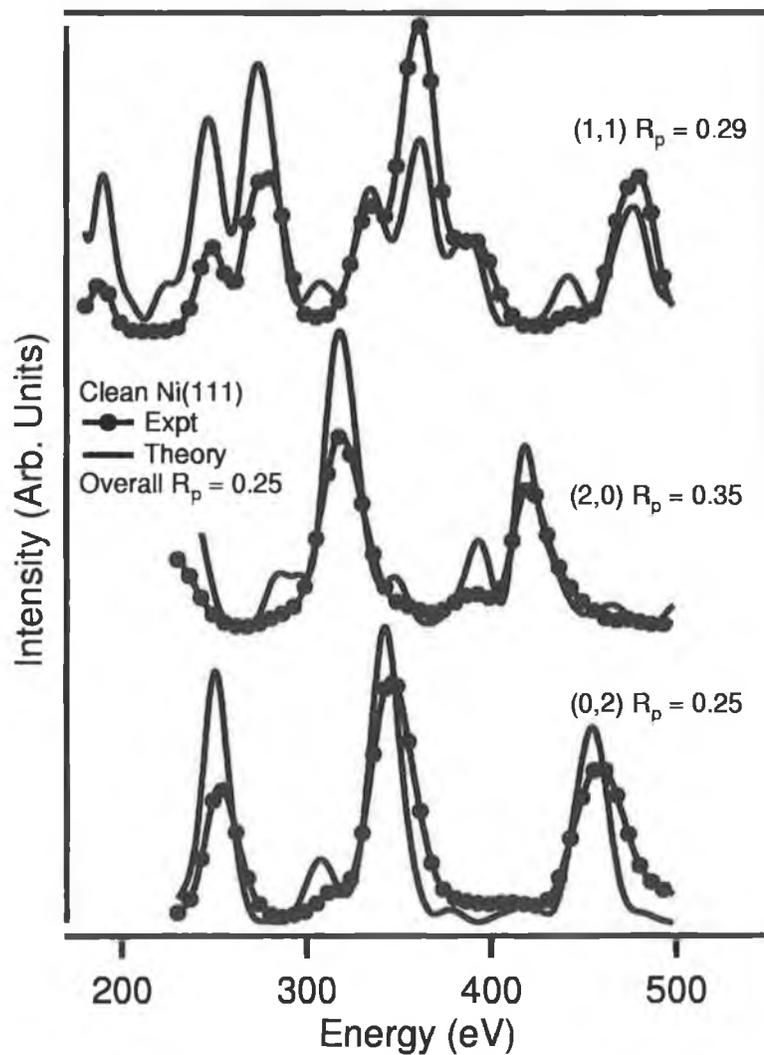


Figure 3.5: Graph of the best fit theory to expt IV curves for the (1,1),(2,0) and (0,2) diffraction beams for the Ni(111) surface

Parameter	Ni(110)	Ni(111)
R_p	0.18	0.25
d_{12}	1.15	2.03
d_{23}	1.28	2.03
d_{34}	1.24	2.03

Table 3.1: Best-fit interplanar spacings and Pendry R-factors, R_p obtained for the Ni(110) and Ni(111) surfaces. d_{12}, d_{23} and d_{34} represent the distances from the first to second, second to third and third to fourth Ni layers respectively.

was also adjusted during the fitting procedure. Although the real part of the inner potential contains information about the work function and fermi energy of the material it is very rarely quoted. This is because in computerised setups it is common for a fixed energy difference to develop between the actual electron gun energy and the computer recorded energy. This is merely a calibration problem but it is rarely fixed !!!!.

Previous studies of the Ni(111) surface have been carried out using between three and five diffracted beams for the analysis [1–3]. These studies concluded that the surface layers are not contracted but remain at their bulk values, i.e. 2.03Å. This result is consistent with our result and the general trend observed in fcc(111) surfaces for which little or no relaxation of the surface layer is generally expected or observed [4–6]. The results for the Ni(110) surface are also in excellent agreement with previous studies [1, 7, 8]. The Ni(110) surface structure is found to exhibit a damped oscillatory relaxation of the interlayer spacings of $\Delta d_{12} = -7.3\%$, $\Delta d_{23} = +2.4\%$ and $\Delta d_{34} = -0.8\%$. The results for this study are compared with previous results in table 3.2

The results of the analysis although consistent with previous work did not produce exceptionally good R-factors. There are both experimental and theoret-

Parameter	This Work	Adams [7]	Feidenhans [8]
Δd_{12}	-7.3%	-8.7%	-4.8%
Δd_{23}	+2.4%	+3.0%	+2.4%
Δd_{34}	-0.8%	-0.5%	Bulk

Table 3.2: Comparison of the results of this study of the Ni(110) surface with previous studies of this surface. Δd represents the difference between the interlayer spacing found from experiment and the bulk Ni(110) value.

ical reasons for this. An essential part of any LEED experiment is the ability to achieve normal incidence to an accuracy of 0.1° . Normal incidence is achieved by adjusting the crystal until the symmetry equivalent beams are identical. This comparison is also performed by estimating the Pendry R-factor of equivalent beams. The R-factor resulting from comparison of symmetric beams should be, according to a general rule of thumb < 0.02 . In figure 3.6 and 3.7 the raw data for the (0,1) and (1,0) diffraction beams for the Ni(111) surface are shown. It can be clearly seen that there are both intensity and peak position differences between the beams which is a clear indication that the crystal is not aligned at normal incidence. This stems from a problem with the manipulator upon which the crystal was mounted and the fact that the chamber was not properly shielded from stray magnetic fields. These stray fields can deflect the incident electron beam. The symmetrized tensor LEED codes use a technique known as RFS (renormalised forward scattering) to calculate the diffraction beam intensities. This method basically involves passing the electron beam forward and back through the crystal until the diffraction beam intensities converge to some value. The RFS method only works for interplanar spacings greater than 1.2\AA . The Ni(110) crystal has an interplanar spacing of 1.24\AA which is on the limit of the techniques capabilities and so can suffer convergence problems.

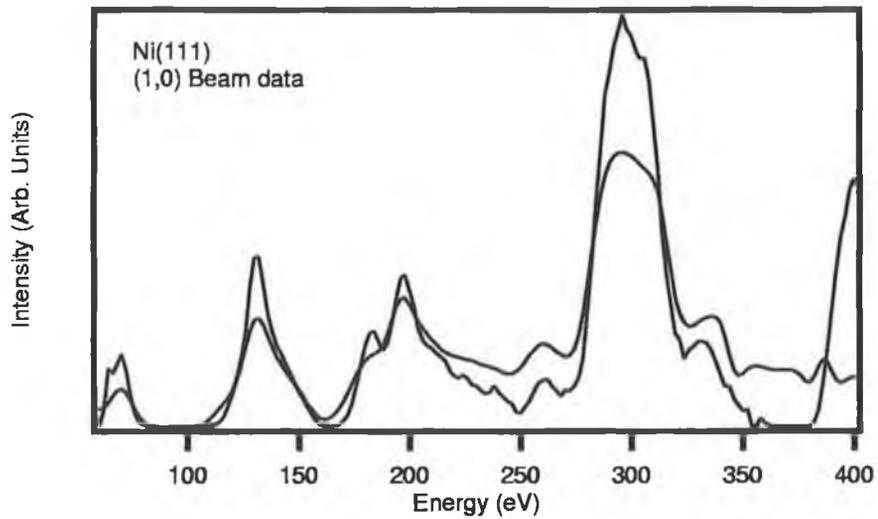


Figure 3.6: Graph of the raw data for the (1,0)

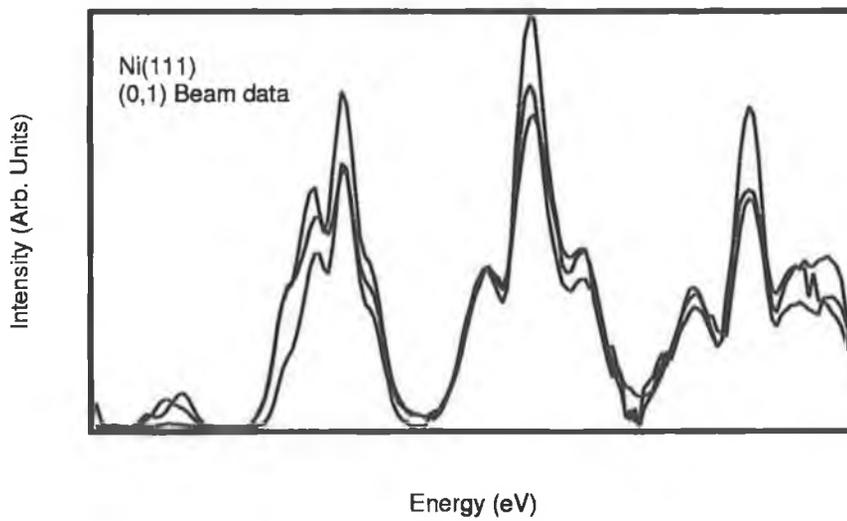


Figure 3.7: Graph of the raw data for the (0,1)

3.5 Conclusion

In this chapter we demonstrated that LEED is a powerful technique for the study of surface structures due to the well developed theory of multiple scattering. The Ni(111) surface showed no change in the surface interlayer spacings compared to the bulk whereas the Ni(110) showed an oscillatory relaxation of the top three layers. The results are consistent with previous works but the results could be improved with better experimental conditions and an analysis of the structure using an alternative LEED code e.g. Layer Doubling which should improve or even solve the convergence problems.

References

- [1] Demuth JE, Marcus PM, Jepsen DW *Phys. Rev. B* **11** (1975) 1460
- [2] Christman K, Ertl G, Schober O *Surf. Sci.* **40** (1973) 61
- [3] Kaukasoina P, Lindroos M, Diehl RD, Fisher D, Chandavarkar S, Collins IR
J. Phys: Cond. Matt. **5** (1993) 2875
- [4] Chae KH, Lu HC, Gustafsson T *Phys. Rev. B* **54** 19 (1996) 14081
- [5] Stairis P, Lu HC, Gustafsson T *Phys. Rev. B* **72** 22 (1994) 3574
- [6] Nascimento VB, de Castilho CMC, Soares EA, de Carvalho VE, Bittencourt
C, Woodruff DP *to be published*
- [7] Adams DL, Petersen LE, Sorensen CS *J. Phys. C: Solid State Phys* **18** (1985)
1753
- [8] Feidenhans IR, Sorensen CS, Stensgaard I *Surf. Sci.* **134** (1983) 329
- [9] M.A. VanHove, W.H. Weinberg, C.M. Chan, *Low-Energy-Electron-Diffraction* (1986), Springer Verlag
- [10] M.A. Van Hove, S.Y. Tong *Surface Crystallography by LEED* (1979)
- [11] J.B. Pendry, *Low Energy Electron Diffraction*

Chapter 4

An Evolutionary Strategy for LEED

4.1 Introduction

Mathematical Optimisation is the formal title given to the branch of computational science concerned with finding the best solution for problems in which the quality of an answer can be expressed as a numerical value. Such problems arise in all areas of mathematics, the physical, chemical and biological sciences, engineering, economics and management to name a few and the range of techniques designed to solve these problems is equally as wide. Optimisation techniques can be broadly split into two classes local and global. Local techniques are generally applied where it is known that only one best solution exists. Global search methods are concerned with problems in which several good solutions exist but only one best solution exists.

Low Energy Electron Diffraction (LEED) has undergone many improvements over the last number of years. Increases in speed in the analysis of LEED data have been the result of two distinct efforts. The development of perturbative techniques such as Tensor LEED, Kinematic LEED, and Linear LEED has led to huge

increases in computational speed compared to full dynamical LEED calculations. A parallel effort has been in the development of search strategies for global minimisation of the reliability factor (R-factor) that quantifies the agreement between the experimental and theoretical data. It is the aim of this chapter to present a new approach to this problem.

Clean metal surfaces, like the Ni(110) and Ni(111) surfaces discussed in chapter 2 have been extensively studied in the past. It is well documented that the structures of clean metal surfaces generally undergo some relaxation of the outer most layers but the resulting structure is not significantly different from that of the bulk crystal. The number of parameters involved in a clean metal structural determination is also quite small and so Tensor LEED or a simply grid search can be performed to find a structural solution reasonably quickly. Reconstructed surfaces are significantly more complex and R-factor minimisation becomes a difficult problem due to strong coupling between the structural parameters involved and the existence of many local minima in the R-factor parameter space. For many years the basic method of finding the global minimum was to use an exhaustive grid search. Improvements came with the application of steepest decent methods [1,2]. These methods worked well in finding minima but to find the true global minimum the search has to be started from several randomly chosen parameters, the lowest minimum found being taken as the global minimum. This is the so called "multiple launch" approach. A review of standard search strategies and their application to LEED can be found in references [1,2]. Two general classes of methods have been developed to solve such global minimisation problems. These are simulated annealing (SA) [3] and what could be called biological algorithms. Biological algorithms aim to find the global minimum by mimicking to some extent biological

processes. Examples of these are genetic algorithms, evolutionary strategies and Neural networks.

4.2 Review of Current Global Search Methods

4.2.1 Simulated Annealing

The simulated annealing algorithm is a technique that has attracted a lot of attention in the last few years as it is a robust optimisation technique suitable for large scale optimisation problems, especially problems in which the desired global minimum is located among many shallower local minima. The simulated annealing algorithm is what is called a probabilistic hill-climbing algorithm. This means that during the search process the moves that increase the R-factor (uphill moves) are accepted in addition to moves which decrease the R-factor (downhill moves). This is the central point that enables the search algorithm to locate the global minimum among all the other local minimum. At the heart of the SA algorithm is the Metropolis criterion which controls the acceptance probability of every "uphill" step in the search. Starting from an initial point \hat{x}_i a random step dx is chosen leading to a new point \hat{x}_j . The change in the R-factor $\Delta R = R(\hat{x}_j) - R(\hat{x}_i)$ is evaluated. If the function change is negative ($\Delta R < 0$) i.e. a better point has been found, the move will be accepted and \hat{x}_j will be a better approximation to the minimum than \hat{x}_i . If the change is positive ($\Delta R > 0$) a poorer point has been found. This point will be accepted with a probability given by the Boltzmann distribution $P(\Delta R) = e^{-\frac{\Delta R}{T}}$. The acceptance of bad points is essential to finding the global minimum as otherwise the algorithm can get trapped in a local minimum.

The probability is controlled by the dimensionless parameter T , an artificial

temperature. T is initially set high to allow large steps around the parameter space and is gradually decreased during the search. The SA algorithm is basically a search based on randomly chosen steps accepted or rejected according to the Metropolis criteria together with a gradual reduction of the temperature T . The main issue regarding the performance of simulated annealing is to lower the temperature as fast as possible while ensuring that one does not get trapped in a local minimum. The basic outline of the simulated annealing algorithm is presented in figure 4.1

4.2.2 Genetic Algorithms

Genetic Algorithms essentially simulate the natural evolution of living organisms. At the beginning the user of a GA has to devise a way to encode possible solutions of the optimisation problem into bit strings, each bit string is called a "chromosome". A "fitness value" has to be defined as a measure of the chromosomes performance in the problem. The goal is to maximise this fitness.

Initially a randomly chosen "population" of chromosomes is created. The chromosomes in this first "generation", are chosen randomly in the sense that they are distributed about the parameter space of the problem. Using a selection rule that combines chance and a preference for chromosomes with a high fitness, several pairs of chromosomes are selected as "parents" for the creation of the next generation. The chromosomes (bit strings) of the parents are split at some point and their parts combine to produce two new chromosomes in a process called crossover. Each parent therefore hands some information to the next generation. Occasionally some of the bits in a chromosome are randomly inverted thus simulating "mutation" in nature. In this way a new generation is produced

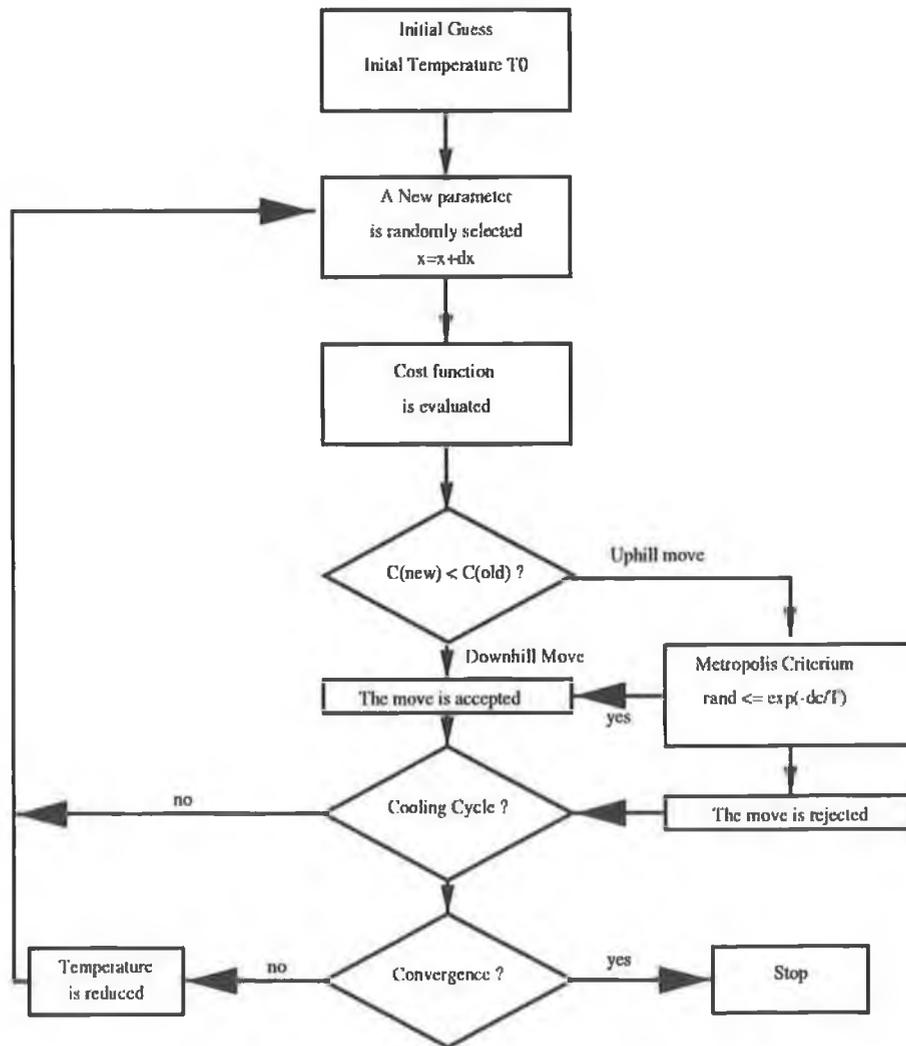


Figure 4.1: A general overview of the Simulated Annealing algorithm

Nth Generation		Nth + 1 Generation	
R_p	bit code	R_p	bitcode
0.1258	010110001101111000001	0.0351	<i>011000001010101000001</i>
0.1422	<i>011000001000101001100</i>	0.1113	011000101000111001010
0.1541	011000001000011000011	0.1142	011000001000111001010
0.1554	010010001101111000001	0.1218	0110101010011111000001
0.1668	011000001000101001111	0.1258	010110001101111000001
0.1730	011000001101101000001	0.1422	011000001000101001100
0.2042	010000001101111000001	0.1422	011000001000101001100
0.2146	011010101000101010000	0.1478	010001000110111100000
0.2515	011010100111001000010	0.1505	011010001000011000010
...
0.2584	010010001010101000011	0.3630	010010001000101001100
0.2612	011100101000111001010	0.4476	010010000111001010010

Table 4.1: Example of the crossover procedure used in Genetic Algorithms The chromosome fragments printed in italics and bold letters in the Nth generation recombine to form the italic and bold printed chromosomes in the Nth+1 generation Note that mutation has occurred in the formation of the Nth+1 italic chromosome

that replaces the previous generation. The process of producing new generations is then repeated until a termination criteria is fulfilled. For more details about Genetic algorithms see [11]

4.3 Review of Global Search Methods in LEED

Rous [3] was the first to apply true global minimization techniques to LEED with the application of the simulated annealing technique to the Ir(110)-(2x1) surface [5]. This surface has since been used as a benchmark for any global search technique as it contains 41 local minima besides the global minimum. In the Ir(110)-(2x1) [5] surface structure there are 11 parameters but for the structural analysis only three parameters, the interlayer spacing between the first and second

layers, the spacing between the second and third layers and the row pairing in the second layer were used. The search was conducted over a $1\text{\AA} \times 1\text{\AA} \times 1\text{\AA}$ parameter space in 0.05\AA steps. This is equivalent to some 8000 structural evaluations being needed for an exhaustive grid search.

Rous [5] found that application of an unoptimized simulated annealing algorithm resulted in 4000 structures being needed on average to find the global minimum, a 50% saving in computational time when compared with an exhaustive search strategy. As SA is a stochastic method the number of structures which have to be evaluated to locate the global minimum will depend on the choice of an initial starting point. The most common way to evaluate a search algorithm is to run the simulation a number of times and find the average number of runs required to find the global minimum. Rous, however, evaluated the efficiency of global search methods in a different manner. He asked how many function evaluations are required to always find the minimum or to find the minimum with a given probability. To evaluate the efficiency of the SA algorithm a statistical distribution of the number of structural evaluations needed to find the global minimum for 1000 random starting points, i.e. structures was generated. From this distribution Rous [3] determined that although 4000 function evaluations were needed on average to locate the global minimum the probability of finding the global minimum, P_{global} , with 4000 structural evaluations was 0.9. Optimization of all parameters involved in the SA enabled him to find the minimum with a probability P_{global} of 0.9 with 800 structure evaluations while 2500 structures were needed to attain a P_{global} close to 1.

An alternative global strategy based on an R-factor dependent gaussian distribution of structures was developed by Kottcke and Heinz [6]. This method has

the distinct advantage of using the more computationally efficient Tensor LEED thereby improving the required computational times. However this method is inefficient when using low numbers of parameters and is restrictive in parameter ranges. For example this method was applied to the Ir(110)-(2x1) with only the top two interlayer spacings varied in 0.05 steps over a range of $\pm 0.4\text{\AA}$. They found that 330 structural evaluations were needed to find the global minimum whereas an exhaustive grid search would only take 256 structures. The efficiency of the method increased as the number of parameters increased but the ranges over which the parameters were varied also decreased making it difficult to assess the quality of the search. The problem with using Tensor LEED for optimization is that it is first and foremost an approximation method. Tensor LEED can generate the general shape of the parameter space and in this way it can locate the minimum, but the R-factors produced can differ significantly. For example if the best fit R-factor is at a distance 0.2 from the start point the Tensor LEED algorithm will find the minimum and estimate the R-factor as 0.3. If the search is restarted on the minimum the same result is found but the accuracy improves and the R-factor drops to 0.25. This makes it difficult to use in a global search because not only does the method have to contend with a parameter space with several minima but the algorithm will produce spurious R-factors for trial structures.

While research work into Evolutionary Strategies detailed in this thesis was underway, Doll and Van Hove [4] independently applied Genetic Algorithms to the problem. They found that for the Ir(110)-(2x1) [5] structure they could locate the global minimum with an average of 314 structural evaluations, with 800 structures being needed to achieve a P_{global} indistinguishable from 1 (0.9998). This is a considerable improvement on the Simulated Annealing method [3]. It was

suggested by Doll and Van Hove [4] that an Evolutionary Strategy (ES) might improve further on the Genetic Algorithm approach. An extremely simple and promising Evolutionary Strategy is the Differential Evolution method developed by Storn and Price [7]. This method is very fast and remarkably simple and is described in the next section.

4.4 Differential Evolution

4.4.1 Overview

Differential Evolution (DE) is a global search strategy which in its simplest form works as follows. Given a number of parameters D , one initially creates an array of NP , D dimensional real valued vectors. Each member of the vector corresponds to a parameter in the LEED analysis. These parameter vectors are randomly chosen and should uniformly cover the entire parameter space. This array of vectors is referred to as a target population and each subsequent target population created is referred to as a generation. The R-factors of each of the parameter vectors of this initial target population are evaluated. For a particular target vector, three vectors are selected at random from the target population. A mutant vector is created from these by adding the weighted difference of two of the vectors to the third. From this mutant vector a trial vector is created by randomly mixing parameters from the mutant vector and the target vector in a crossover procedure. The R-factor for this trial vector is then evaluated and if it has a fitness better than that of the target vector it will replace the target vector in the next population. This procedure is repeated for every member of the current target population. In this way a new target population, or generation is created. Typically this proced-

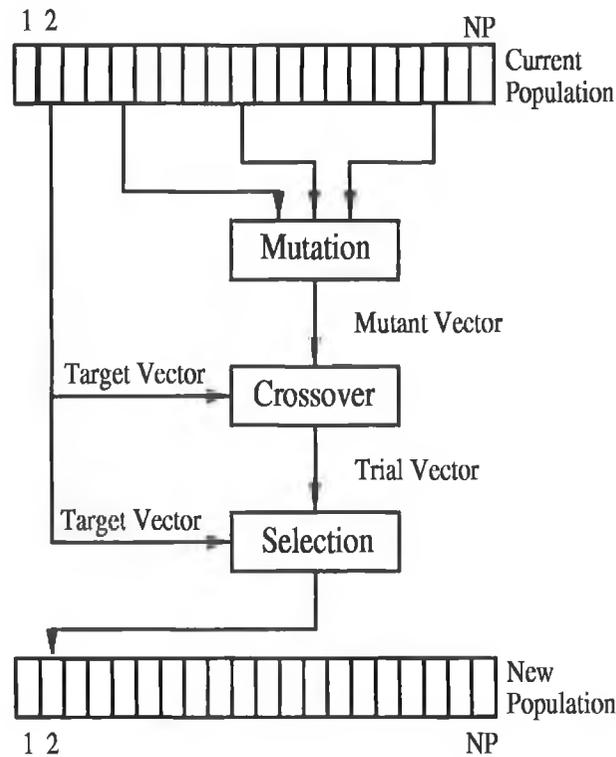


Figure 4.2: A general overview of the Differential Evolution algorithm

ure is repeated until the target population converges towards some optimum set of parameters. The process is outlined diagrammatically in figure 4.2 and in the next section each step of the process shall be discussed in more detail.

4.4.2 Population Initialisation

The most common approach is to generate a uniform random distribution of values about the parameter space, or alternatively a gaussian distribution about some estimated "best" parameter set could be used. Both methods have been investigated here the results of which will be discussed later.

4.4.3 Mutation

In the introduction to this section an outline was given of the simplest DE method, whereby three vectors were used to generate a target vector. An alternative method for generating target vectors is to add the weighted differences of four randomly chosen vectors to the best member of a population (2). (The best member is the member of the population with the lowest R-factor) . In stricter terms, for each target vector x_i of dimension D where $i=1,2 \dots NP$ a mutant vector v_i can be generated by either

$$v_i = x_1 + F \cdot (x_2 - x_3) \quad (4.1)$$

$$v_i = x_{best} + F \cdot (x_1 + x_2 - x_1 - x_2) \quad (4.2)$$

where x_1, x_2, x_3, x_4 are four different vectors randomly chosen from the population. F is a real factor that controls the size of the vector difference, and is empirically chosen to be greater than 0 and less than 2.0. The most common values used are between 0.5 and 1.0. However the value of F can be fine tuned to optimise the performance of the DE algorithm for a particular cost function. The two mutation schemes described above shall be referred to as DE method (1) and DE method (2) respectively. Both methods were investigated and the effects of initial population choice were investigated for method (2) and are discussed later.

4.4.4 Crossover

To increase diversity in the trial population a crossover operation is introduced such that a trial vector u_i is formed by randomly interchanging the individual parameters of the target vector, x_i and mutant vector v_i according to,

$$u_{ij} = \begin{cases} v_{ij} & \text{if random no. } < \text{CR or } j = \text{a random vector index} \\ x_{ij} & \text{if random no. } \geq \text{CR and } j \neq \text{a random vector index} \end{cases} \quad (4.3)$$

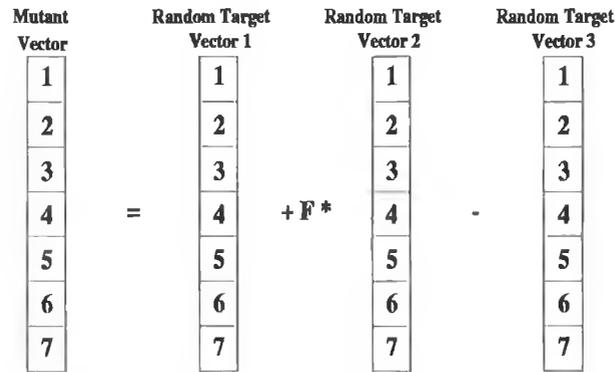
CR is the crossover probability [$0 \geq \text{CR} \leq 1$] which is determined by the user. The choice of CR depends on the DE method used. The use of the randomly chosen index $j = 1, 2, \dots, D$ to determine whether crossover occurs is to ensure that the trial vector u_i , contains at least one parameter from the mutant vector v_i .

4.4.5 Selection

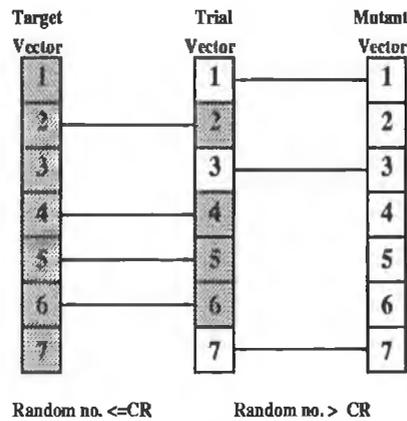
To decide whether the trial vector should become a member of the next generation, the R- factors of the trial and the target vector are compared. The vector with the lowest R- factor is kept and becomes a member of the next generation. In the following analysis the Pendry R-factor was used to compare IV-spectra but in practice any R-factor can be applied.

4.5 Ir(110)-(2x1) Missing Row - Row Pairing Model

As outlined in 4.3 the Ir(110)-(2x1) [5] has been used to test the efficiency of global search algorithms in LEED. In order to compare Differential Evolution [7] to alternative algorithms the Ir(110)-(2x1) [5] structure was also used in this work. Theoretical IV-curves were generated for the known structure of Ir(110)-(2x1) [5] and used and used as pseudo-experimental data. Six integral order and five half order spots over a range 45 eV to 250 eV were used. The Differential Evolution method was applied to the standard Van Hove and Tong codes [10] The two forms of the DE method, (1) and (2), mentioned above were tried and



(a)



(b)

Figure 4.3: Schematic Diagram of the mutation and crossover processes in the Differential Evolution algorithm. (a) A mutant vector i is formed by choosing three random target vectors and adding the weighted difference of two of them to a third vector (b) A trial vector is generated by swapping the parameters in the target vector with parameters in the mutant vector. Swapping is determined by the crossover probability CR . A random number is generated for each parameter in the vector. If this number is $\leq CR$, then the parameters are swapped

the effects of initial population choice on speed of convergence was examined for DE method (2). For DE method (1) crossover probabilities, CR between 0.1 and 0.5 were found to work best. The data presented here is for a population of 20, a crossover probability CR of 0.4 and mutation value F of 0.5. For DE method (2) crossover probabilities between 0.5 and 1.0 were found to work best. The data presented here is for a population of 50 a crossover probability, CR of 0.8 and the mutation value, F was chosen to 0.5. The performance of the algorithm depends to some extent on the choice of CR and F. The crossover and mutation probabilities displayed were chosen based on previous experience with the method [9]. Following the method of references [3,4,6] three parameters, the interlayer spacing between the first and second layers, the second and third layer spacing and the row pairing in the second layer were varied over a $1\text{\AA} \times 1\text{\AA} \times 1\text{\AA}$ parameter space.

4.6 Results

To examine the efficiency of the method, the statistical distribution of the number of structures evaluations required to obtain the global minimum must be examined. To generate such a distribution the method was executed for 100 random initial populations. The algorithm was terminated when the best parameter in the population was within 0.05 of the actual parameter values which are known a priori. (This criteria was used in both previous optimisation methods [3,4]). The number of evaluations carried out is plotted in histogram form in figures 4.5,4.6 and 4.7. Fitting a normal distribution to these histograms, the average number of structural evaluations needed, n_{av} can be determined along with the variance of the distribution σ_n . The probability of finding a global minimum after a certain

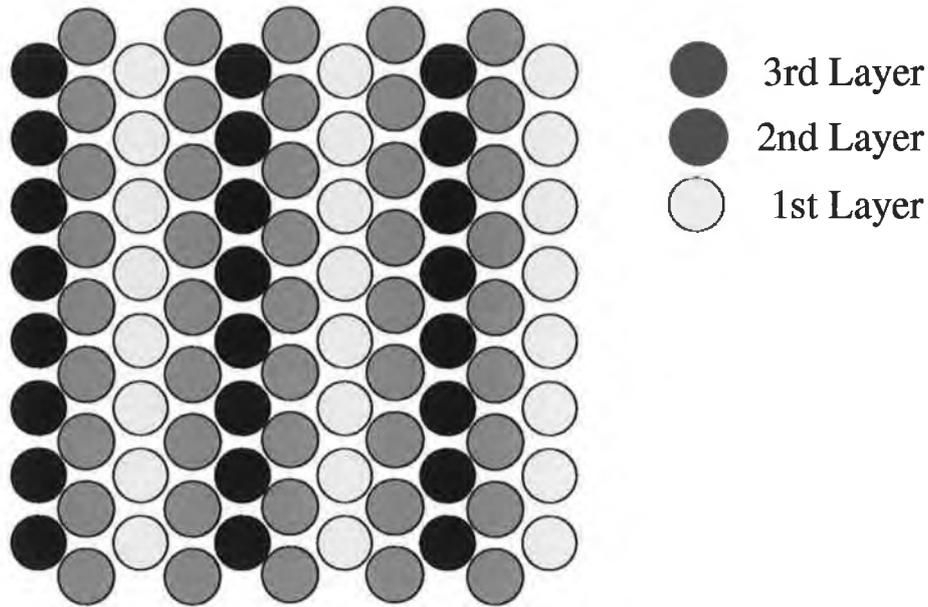


Figure 4.4: A top view of the Ir(2x1) surface

number of structures, n_{tot} is found from

$$P_{global}(n_{tot}) = \frac{1}{\sqrt{2\pi}\sigma_n} \int_{-\infty}^{n_{tot}} \exp\left(-\frac{1}{2} \frac{(x - n_{av})^2}{\sigma_n^2}\right) dx \quad (4.4)$$

From the plots shown in figures 4.5 - 4.7 it was found that DE method (1) required $n_{av} \pm \sigma = 186 \pm 86$ structures on average. This corresponds to a probability of finding the global minimum, P_{global} of 0.998 for 500 structures. It was found that DE method (2) required on average $n_{av} \pm \sigma = 189 \pm 121$ structures which corresponds to a probability, P_{global} of 0.998 for 600 structures. Differential Evolution Method (2) with a gaussian deviate population required $n_{av} \pm \sigma = 143 \pm 126$ structures which corresponds to a probability, P_{global} of 0.998 for 600 structures.

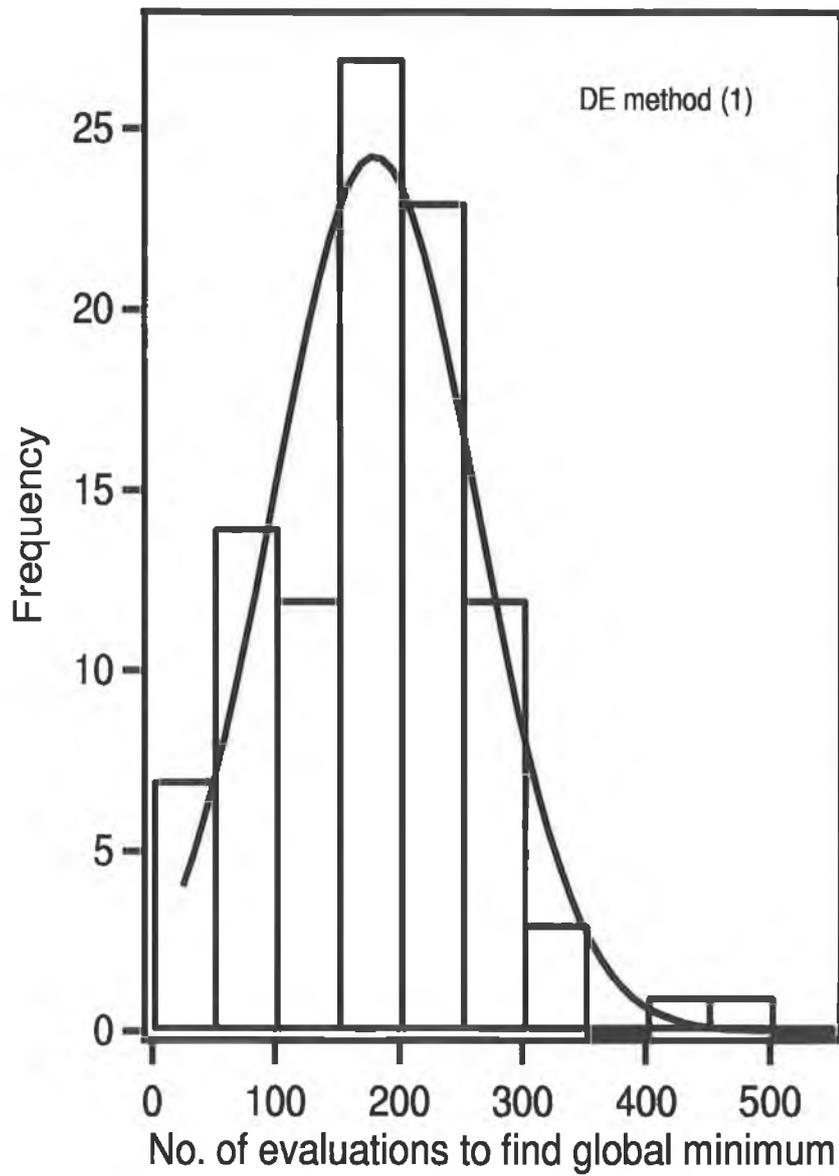


Figure 4.5: Histogram resulting from DE method (1) with a uniform initial population

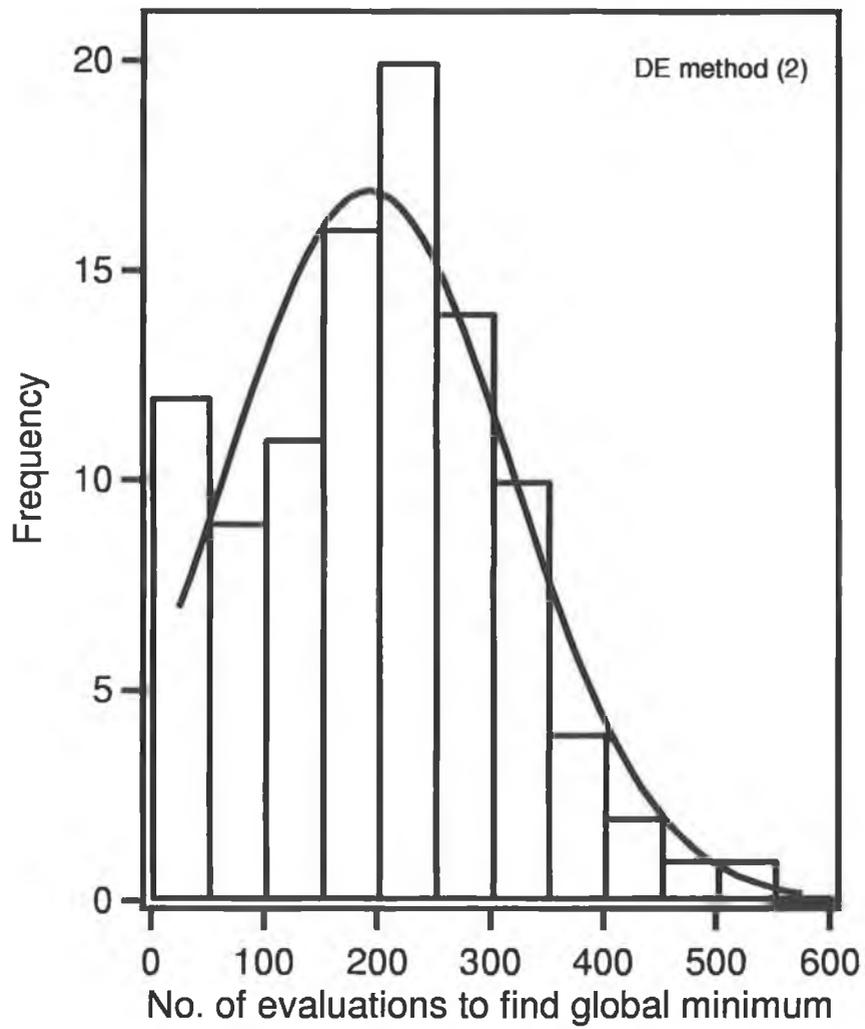


Figure 4.6: Histogram resulting from DE method (2) with a uniform initial population

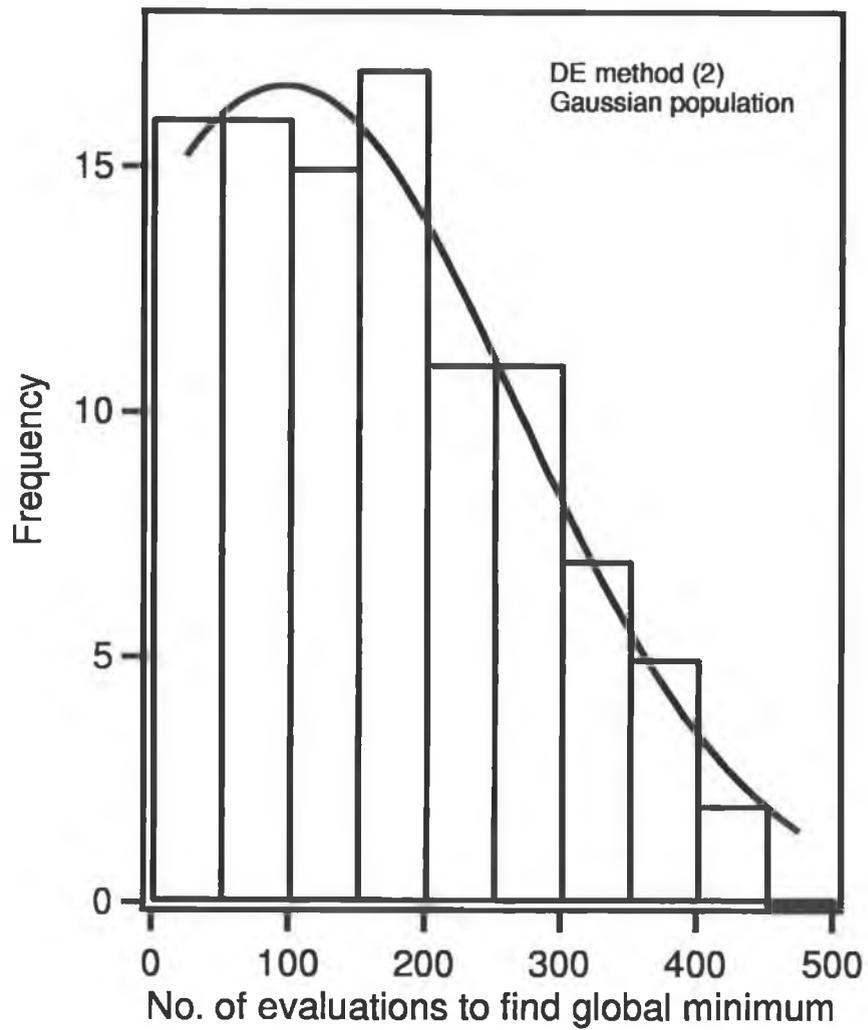


Figure 4.7: Histogram resulting from DE method (2) with a gaussian deviate initial parameter population

Method used	No. of structures	P_{global}
Exhaustive Search	8000	1
Simulated Annealing	2500	0.998
Genetic Algorithm	800	0.998
DE method(1)	500	0.998
DE method(2)	600	0.998

Table 4.2: Comparison of search strategies as applied to the Ir(110)(2x1) surface over a $1\text{\AA} \times 1\text{\AA} \times 1\text{\AA}$ parameter space

4.7 Discussion

The two Differential Evolution methods differ in their performance. The standard Differential Evolution method (1) requires only a small initial population, (typically 5 times the number of parameters) and several generations, in this case 25, to find the global minimum. Increasing the size of the population for this method does not produce a marked increase in the speed of convergence. Differential Evolution method (2) however requires a large initial population, (typically 20 times the number of parameters), but does not require several generations, 12 generations being needed on average in this case. The results compare quite favorably when compared with previous global search methods [3,4] Table 4.2.

The DE methods were 4 times faster than the Simulated Annealing method and 1.33 times faster than the Genetic algorithm

A gaussian initial population was chosen to see if some prior knowledge of a system could increase the speed of convergence. Clean metal surface relaxation, for example, does not vary greatly from the guessed (bulk-like) structure. Therefore in this situation it would be reasonable to distribute the initial parameter estimates around their bulk values. It was also hoped that the gaussian population could complement the 'fingerprinting' technique [8]. Fingerprinting is

a technique whereby LEED intensity spectra measured from an unknown surface structure are compared to previously studied structures on the same or similar materials. It has been shown that a good initial guess at a structure can be obtained from this method. However, the results were disappointing. Although the method resulted in a smaller average number of trials there was quite a wide distribution which resulted in a larger number of structures being required to achieve a global minimum to a probability of 0.998 when compared to DE method (1).

The majority of the calculations were done on a standard Pentium 100MHz machine running the Linux operating system. All code was compiled using gnu Fortran and on average it took between 18-24 hours to find a global minimum (6 phase shifts were used in the psuedo-expermental data and in the search process). These execution times could definitely be improved using a good workstation and an optimized compiler. (Doll and Van Hove quote 1.25 hours for a genetic algorithm to evaluate 800 structures). The results presented here show that stochastic search methods, such as the Differential Evolution algorithm, greatly reduce the number of structures to be evaluated and permit the evaluation of structural parameters on modest computational facilities, and in a relatively short time. Accumulation of the statistical data given these computational facilities was lengthy and time considerations limited further investigations of the scaling behavior of the method with increasing parameters and of fully optimizing the crossover probability, CR and mutation rate F. Storn [7] has shown that Differential Evolution can provide a 50% increase in speed over a conventional genetic algorithm so it is reasonable to assume the values used in this analysis are quite good although perhaps not fully optimized. It has been noted that genetic algorithms and other stochastic methods can find the general location of a global

minimum extremely quickly but take quite a long time to locate the actual minimum. Rous [3] suggests a hybrid search strategy, by which a global search method is employed for a coarse search and then a descent method is used to find the minimum exactly. A possible solution to this would be to use the mutation rate of DE in a similar way to the artificial temperature, T , in Simulated Annealing. In the DE method the size of the mutation rate could be varied depending on the spread of parameter values. So for example a large spread could result in a large initial mutation rate but as the population converges the size of the mutation rate could decrease such that smaller steps are taken speeding convergence and enabling the actual global minimum to be determined quite quickly. This idea has not been tested and is an area that requires further research. The main advantage to the DE method is its sheer simplicity. It requires little additional code to the standard Van Hove Tong codes and it is extremely robust.

4.8 Conclusion

In this chapter an Evolutionary Strategy, Differential Evolution method has been applied to the global search problem in LEED and has been found to be an extremely effective method for global optimization. It is hoped that application of the algorithm to more complex systems will provide information on how the algorithm scales with the number of parameters and prove itself a valuable tool for LEED structural analysis.

References

- [1] M.A. Van Hove, W. Woritz, H. Over, P.J. Rous, A. Wander, A. Barbieri, N. Materer, U. Starke, G.A. Somorjai, *Surf. Sci. Reports*, **19** (1993) 191
- [2] P.J. Rous, M.A. VanHove, G.A. Somorjai *Surf. Sci.* **226** (1990) 15
- [3] P.J. Rous, *Surf. Sci.* **296** (1993) 358
- [4] R. Doll, M.A. VanHove, *Surf. Sci.* **355** (1996) L393
- [5] C.M. Chan, M.A. Van Hove, *Surf. Sci.* **171** (1986) 226
- [6] M. Kottcke, K. Heinz, *Surf. Sci.* **376** (1997) 352
- [7] R. Storn, K. Price, *Jn. Global Optimization* **11** (1997) 341
- [8] H. Over, M. Gierer, H. Bludau, G. Ertl, S.Y. Tong, *Surf. Sci.* **314** (1994) 243
- [9] A.A. Cafolla, *Surf. Sci.* **314** (1994) 243
- [10] M.A. Van Hove, S. Y. Tong *Surface Crystallography by LEED*, Heidelberg, 1979
- [11] D.E. Goldberg *Genetic Algorithms is search, optimization and machine learning* Addison-Wesley (1989)

Chapter 5

Summary and Conclusion

5.1 Summary

In this thesis the general theory and application of LEED to the study of surface structure has been outlined. As an example of the use of LEED results on the Ni(110) and Ni(111) surfaces have been presented. The clean Ni(111) surface exhibits no surface relaxation which is to be expected of close packed surfaces. The Ni(110) surface which is the most open of the face centred cubic, fcc, metal faces shows an oscillatory relaxation of the top three Ni layers.

In LEED theoretical curves are compared to experiment by reliability factors or R-factors. For complex systems multiple solutions can exist to a problem with only one solution being the true solution. To deal with this problem an Evolutionary Strategy called Differential Evolution was applied to the problem and results were presented. Differential evolution was found to be extremely fast at finding the solution compared to alternative methods such as Simulated Annealing and Genetic Algorithms. Overall the work presented in this thesis has demonstrated that LEED is a highly accurate technique for determining the structure of surfaces

and coupled with a global search algorithm such as Differential Evolution can be a powerful tool for the analysis of surface structures.