

New developments in the theory and interpretation of X-ray spectra based on fast parallel calculations

J. J. Rehr* and A. L. Ankudinov

Department of Physics, University of Washington, Seattle, WA 98195-1560, USA. E-mail: jjr@phys.washington.edu

There has been dramatic progress over the past decade both in theory and in *ab initio* calculations of X-ray absorption fine structure. Significant progress has also been made in understanding X-ray absorption near-edge structure (XANES). This contribution briefly reviews the developments in this field leading up to the current state. One of the key advances has been the development of several *ab initio* codes such as *FEFF*, which permit an interpretation of the spectra in terms of geometrical and electronic properties of a material. Despite this progress, XANES calculations have remained challenging both to compute and to interpret. However, recent advances based on parallel Lanczos multiple-scattering algorithms have led to speed increases of typically two orders of magnitude, making fast calculations practicable. Improvements in the interpretation of near-edge structure have also been made. It is suggested that these developments can be advantageous in structural biology, *e.g.* in post-genomics studies of metalloproteins.

Keywords: extended X-ray absorption fine structure spectroscopy (EXAFS); X-ray absorption near-edge structure (XANES); Lanczos multiple-scattering algorithms; parallel processing.

1. Introduction

Advances in theory over the past decade have revolutionized the technique of extended X-ray absorption fine-structure spectroscopy (EXAFS) for local structure determinations. Indeed, the basic theory is now well understood, as discussed in a recent review (Rehr & Albers, 2000). Significant progress has also been made in understanding X-ray absorption near-edge structure (XANES), *i.e.* the structure within about 30 eV of the threshold. These theoretical advances have led to the development of *ab initio* codes for XAS calculations. Some commonly used codes include *CONTINUUM* (Natoli *et al.*, 1980), *EXCURVE* (Binsted, Campbell *et al.*, 1991), *FEFF* (Rehr *et al.*, 1991; Zabinsky *et al.*, 1995; Ankudinov, Ravel *et al.*, 1998), *GNXAS* (Filipponi & Di Cicco, 1995) and *WIEN98* (Blaha *et al.*, 1990). The EXAFS code development was revolutionary in that it provided accurate theoretical standards that eliminated the need for tabulated phase shifts (Teo & Lee, 1979; McKale *et al.*, 1986), thus simplifying and improving the experimental analysis.

Despite this progress, a fully quantitative theory of XANES has been elusive because of many-body effects such as the treatment of the core hole, inelastic losses and multiplet effects. Moreover, XANES calculations have remained computationally challenging, typically taking many cpu hours on current machines. However, recent code developments based on Lanczos multiple-scattering algorithms on parallel computers (Ankudinov *et al.*, 2002) have led to speed increases, typically of two orders of magnitude, making fast calculations practicable. Improvements in the interpretation of near-edge structure have also been made based on electronic (*e.g.* charge counts) and structural information. We suggest that these developments may enable an automated XANES analysis of biological

structures and may be used in BioXAS applications such as post-genomics studies of metalloproteins.

2. Key developments in XAS theory

2.1. One-electron approximation

The basic multiple-scattering (MS) theory of XAS (Lee & Pendry, 1975; Rehr & Albers, 2000) is now well established. Formally, the X-ray absorption coefficient μ from a core level i and a given X-ray energy $\hbar\omega$ is given by Fermi's Golden Rule,

$$\mu(E) \simeq \sum_f |\langle i | \mathbf{A} \cdot \mathbf{p} | f \rangle|^2 \delta(E - E_f), \quad (1)$$

where $E = \hbar\omega + E_i$ is the photoelectron energy, $\mathbf{A} \cdot \mathbf{p}$ is the coupling to the X-ray field and the sum is over unoccupied final states $|f\rangle$. Most practical calculations are based on the dipole approximation and the reduction of the Golden Rule to a one-electron approximation. In addition, current work usually assumes the final-state rule, in which the states $|f\rangle$ are calculated in the presence of an appropriately screened core hole, and all many-body effects and inelastic losses are lumped into a complex-valued optical potential. This theory is the basis for *FEFF* and many other codes. However, many-body effects, *e.g.* the screening of the X-ray field (Zangwill & Soven, 1980) and multiplet effects (de Groot, 1994), are sometimes important. A challenge for the future is to improve these approximations.

2.2. Real-space Green's function (RSGF) formalism

An important formal development in XAFS theory is the RSGF approach (Schaich, 1973; Lee & Pendry, 1975), which is equivalent to a real-space analog of the KKR band structure method (Schaich, 1973). The need to calculate final states explicitly in the Golden Rule is generally a computational bottleneck and can only be carried out efficiently for highly symmetric systems such as molecules or crystalline solids. However, many systems of interest, *e.g.* biological structures such as metalloproteins, lack such symmetry. Thus, instead of actually calculating final states, it is preferable to re-express μ in terms of the photoelectron Green's function or propagator in real space,

$$\mu(E) \simeq -\pi^{-1} \text{Im} \langle i | \hat{\boldsymbol{\epsilon}} \cdot \mathbf{r}' G(\mathbf{r}', \mathbf{r}, E) \hat{\boldsymbol{\epsilon}} \cdot \mathbf{r} | i \rangle, \quad (2)$$

where $\hat{\boldsymbol{\epsilon}}$ is the X-ray polarization. In MS theory, the propagator $G(\mathbf{r}', \mathbf{r}, E) = \sum_{L,L'} R_L(\mathbf{r}') G_{L,L'} R_{L'}(\mathbf{r})$, so the expression for μ can be reduced to a calculation of atomic dipole-matrix elements $M_L = \langle i | \hat{\boldsymbol{\epsilon}} \cdot \mathbf{r} | L \rangle$ for final-state angular momentum and a full propagator matrix $G_{L,L'} = G_{L,L'}^c + G_{L,L'}^{\text{sc}}$, where G^c is the central atom part and G^{sc} includes all contributions from the environment. The matrix $G_{L,L'}$ can be re-expressed formally as a sum over all MS paths that a photoelectron can take away from the absorbing atom and back (Lee & Pendry, 1975). For XANES, this sum may be calculated exactly by matrix inversion (Durham *et al.*, 1982; Natoli *et al.*, 1980),

$$G^{\text{sc}} = \exp(i\delta) (\mathbf{I} - G^0 T)^{-1} G^0 \exp(i\delta'), \quad (3)$$

where T is the scattering t -matrix, while for EXAFS the expansion in an MS series gives rise to a path expansion

$$G^{\text{sc}} = \exp(i\delta') (G^0 T G^0 + G^0 T G^0 T G^0 + \dots) \exp(i\delta). \quad (4)$$

Since $G_{L,L'}$ naturally separates into intraatomic contributions from the central atom and from MS, one obtains $\mu = \mu_0(1 + \chi)$ and hence the structure in μ depends both on the atomic background μ_0 and on the MS signal χ . This result is consistent with the experimental definition of XAFS $\chi = (\mu - \mu_0)/\Delta\mu_0$, where $\Delta\mu_0$ is the jump in the

smooth atomic-like background. Another key development for EXAFS is an efficient and accurate curved-wave scattering theory. Because of curved-wave effects, exact MS calculations are impractical at high energies and can only be carried out with the path expansion for low-order MS paths (Gurman *et al.*, 1986). Nevertheless, the MS expansion can generally be expressed as a sum over MS paths R of the form

$$\chi(k) = S_0^2 \sum_R \frac{|f_{\text{eff}}(k)|}{kR^2} \sin(2kR + \Phi_k) \exp(-2R/\lambda_k) \exp(-2\sigma^2 k^2). \quad (5)$$

To overcome the computational bottleneck, we devised an efficient method now known as the Rehr–Albers (RA) scattering-matrix formalism. This yields curved-wave calculations of the effective scattering amplitude $f_{\text{eff}}(k)$ in terms of a separable representation of the free propagator $G^0(E)$ (Rehr & Albers, 1990). Equation (5) is essentially identical in forms to the famous XAFS equation of Sayers *et al.* (1971), except that all quantities are redefined to include curved-wave and many-body effects implicitly. Here $k = [2(E - E_0)]^{1/2}$ is the wavenumber measured from threshold E_0 , λ_k is the XAFS mean free path and σ is the r.m.s. fluctuation in the effective path length $R = R_{\text{path}}/2$. Because of the path-dependent phase shift Φ_k , theoretical calculations are essential to analyze experimental XAFS data beyond the nearest neighbors, due to the difficulty of obtaining experimental standards. For XANES, exact propagators are needed; however, it transpires that the RA approach still provides an efficient algorithm (Manar & Brouder, 1995) for calculating G^0 , which is implemented in *FEFF8*.

2.3. Fast parallel Lanczos XANES calculations

Owing to the need for matrix inversion in full MS calculations, which scale in time as the cube of system size, XANES calculations are much more time-consuming than EXAFS. Indeed, XANES calculations become computationally intractable in the EXAFS regime or when (*e.g.* for low- Z atoms) the mean free path is very long. Thus one of the challenges in XANES theory is to increase the computational speed. Promising methods include the recursion method (Filipponi, 1991; Ankudinov *et al.*, 2002), repartitioning (Fujikawa, 1993) and iterative approaches (Wu & Tong, 1999), which can provide substantial improvements to the conventional LU (lower-upper) decomposition. In particular, we have found that many modern Lanczos algorithms are stable and can improve computational speed, typically by a factor of three to five. These improvements are now incorporated in the most recent version (8.2) of *FEFF8*. An advantage of the Lanczos approach is that it naturally interpolates between the full and finite MS limits. Thus the approach clarifies how the MS expansion converges with respect to energy and when a path expansion is valid and hence provides a clear way to differentiate between the extended and near-edge regimes.

However, much more dramatic reductions in cpu time can be obtained from parallel computational algorithms. The time for such algorithms scales as $A + B/N$, where N is the number of processors, and hence can provide one to two orders of magnitude further improvement (Ankudinov *et al.*, 2002). Parallelization has been implemented in *FEFF8* with the MPI (message-passing-interface) protocol (Gropp *et al.*, 1994). As a result, XANES calculations, even for very large systems of the order of 10^3 atoms, can now be carried out in about 1 cpu-hour on large parallel computers, *e.g.* systems with 32–64 processors and MPI.

2.4. Scattering potentials, losses and disorder

The scattering potentials at high energies are well approximated by the Matheiss prescription, *i.e.* an overlapped atomic charge density and the muffin-tin approximation. This approximation, however, is often inadequate for XANES where chemical effects are important, and hence self-consistent (SCF) calculations are usually necessary. Muffin-tin corrections are also sometimes needed. The SCF approach implemented in *FEFF8* yields an accurate estimate of the Fermi energy E_F , eliminating an important fitting parameter from XAFS analysis. Another crucial ingredient is the electron mean free path and self-energy shifts. A major difference between ground-state electronic structure calculations and excited states is the need in the latter for a complex energy-dependent ‘self-energy’ $\Sigma(E)$ to account for inelastic losses. Since $k^2/2$ then has an imaginary part $-\text{Im} \Sigma$ (in atomic units) and the terms in G^{sc} vary as $\exp(2ikR)$, each MS path in (5) has a damping factor $\exp(-2R/\lambda_k)$, where the XAFS mean free path is given approximately by $\lambda_k \simeq k/(|\text{Im} \Sigma| + \Gamma/2)$ and Γ is the inverse core-hole lifetime. The self-energy is essentially a screened exchange interaction and is the analog of the exchange-correlation potential V_{xc} of density functional theory. The amplitude reduction factor S_0^2 arises from intrinsic losses in the creation of the core hole, *i.e.* the multielectron shake-up and shake-off excitations (Rehr *et al.*, 1978), and is typically between 0.7 and 0.9. Recently, however, a quasi-boson formalism has been developed to calculate such losses (Campbell *et al.*, 2002). The interference terms between extrinsic and intrinsic losses tend to suppress excitations near the threshold, which may explain why the existence of sharply defined multielectron peaks in XANES has been controversial (Filipponi & Di Cicco, 1996).

The effects of disorder are also crucial in XAFS, as a static-structure approximation yields large errors in XAFS amplitudes. An important development for treating such disorder is the cumulant expansion (Crozier *et al.*, 1988; Dalba & Fornisini, 1997). The *FEFF* codes include Gaussian Debye–Waller factors $\exp(-2\sigma^2 k^2)$ for each MS path. The thermal contributions to this factor can often be fit to a correlated Debye model (Beni & Platzman, 1976). The first cumulant $\sigma^{(1)}$ is the thermal expansion while the third $\sigma^{(3)}$ characterizes the anharmonicity or asymmetry in the pair distribution function. If $\sigma^{(3)}$ is neglected in the analysis, bond distances obtained from EXAFS are too short. Improved treatments of XAFS Debye–Waller factors have been developed that go beyond the Debye approximation (Poiarkova & Rehr, 1999) and permit fits of Debye–Waller factors to local spring constants. Such treatments are important in highly anisotropic materials, such as biological structures. Another approach is to parameterize the N -particle distribution as in *GNXAS* (Filipponi & Di Cicco, 1995). Molecular-dynamics approaches are promising (McCarthy *et al.*, 1997) as a less phenomenological approach, but accurate *ab initio* treatments remain a challenge for the future.

3. Quantitative interpretation of XAS

In independent developments, robust EXAFS analysis procedures have been developed based on the high-order MS path approach. These include novel automated background-removal methods (Bridges *et al.*, 1995) and fitting codes (Newville, 2001; George, 1999; Filipponi & Di Cicco, 1995; Binsted, Campbell *et al.*, 1991; Binsted & Hasnain, 1996) and other codes described in the International XAFS Society catalog (http://ixs.iit.edu/ISX/catalog/XAFS_Programs), which permit accurate refinements of structural parameters from XAFS data.

Considerable effort has also gone into the interpretation of XANES data (Brown *et al.*, 1977; Mansour *et al.*, 1984). However, the quantitative analysis of XANES is still not well developed and

remains a challenge. There is also a need for a reliable inverse method of extracting chemical and geometrical structure from XANES. On the other hand, there has been significant recent progress (Della Longa *et al.*, 2001). Moreover, although the XANES signal depends sensitively on the geometrical structure, its shape directly reflects the excited-state electronic structure in a material. The reason is that the local projected density of states (LDOS) ρ has a form analogous to XAFS, *i.e.* $\rho = \rho_0(1 + \chi)$, and hence $\rho \simeq \gamma\mu$, where $\gamma = \rho_0/\mu_0$ is a smooth atomic ratio. This and analogous relations have recently been exploited to interpret charge counts from XANES and spin and orbital moments from XMCD (Nesvizhskii *et al.*, 2001). Challenges for the future include the need for improved error-analysis methods (Krappe & Rossner, 2000) and increased automation. These developments are crucial to an efficient analysis of the many metalloproteins to be studied in post-genomics research.

4. Conclusions

RSGF theory and codes such as *FEFF* now make possible a general treatment of XAS, encompassing both XAFS and XANES as well as a number of other X-ray spectroscopies. The availability of quantitative theory is, of course, central to a quantitative interpretation of XAS spectra in terms of local geometrical structure and electronic structure, such as LDOS and spin and orbital moments. The current state of XAFS is now highly quantitative and widely used. Although significant progress has been made in XANES theory, the development of a fully quantitative treatment remains challenging. Improved treatments of the scattering potential going beyond the muffin-tin approximation and a better many-body theory, which includes accurate treatments of core-hole, multiplet and other effects, are all necessary. Nevertheless, rapid progress is now being made, and XANES is slowly realizing its promise as a tool for elucidating chemical and electronic structure. In particular, the development of fast parallel calculations of XANES promises to revolutionize XANES analysis. These developments should thus enable automated post-genomics studies of metalloproteins and many other BioXAS systems.

We gratefully acknowledge the contributions to this work from our co-workers, students and collaborators, including R. C. Albers, C. Bouldin, L. Campbell, L. Hedin, A. Nesvizhskii, M. Newville, B. Ravel and J. Sims, and especially the members of the SSRL NIH Biotechnology Resource, G. George, H. Harris, B. Hedman and K. Hodgson. This work was supported in part by the US Department of Energy grants DE-FG06-97ER45623 and DE-FG03-98ER45718 and by NIH grant RR 01209, and was facilitated by the DOE Computational Materials Sciences Network.

References

- Ankudinov, A., Ravel, B., Rehr, J. J. & Conradson, S. (1998). *Phys. Rev. B*, **58**, 7565–7576.
- Ankudinov, A. L., Conradson, S. D., Mustre de Leon, J. & Rehr, J. J. (1998). *Phys. Rev. B*, **57**, 7518–7525.
- Ankudinov, A. L., Rehr, J. J., Bouldin, C., Sims, J. & Hung, H. (2002). *Phys. Rev. B*, **65**, 104107-1–104107-11.
- Beni, G. & Platzman, P. M. (1976). *Phys. Rev. B*, **14**, 1514–1520.
- Binsted, N., Campbell, J. W., Gurman, S. J. & Stephenson, P. C. (1991). *EXCURVE*. SERC Daresbury Laboratory, Warrington, UK.
- Binsted, N. & Hasnain, S. S. (1996). *J. Synchrotron Rad.* **3**, 185–196.
- Binsted, N., Strange, R. W. & Hasnain, S. S. (1991). *Biochemistry*, **31**, 12117–12125.
- Blaha, P., Schwarz, K., Sorantin, P. & Trickey, S. B. (1990). *Comput. Phys. Commun.* **59**, 399–415.
- Bridges, F., Booth, C. H. & Li, G. G. (1995). *Physica B*, **208/209**, 121–124.
- Brown, M., Peierls, R. E. & Stern, E. A. (1977). *Phys. Rev. B*, **15**, 738–744.
- Campbell, L. W., Hedin, L., Rehr, J. J. & Bardyszewski, W. (2002). *Phys. Rev. B*, **65**, 064107-1–064107-13.
- Crozier, E. D., Rehr, J. J. & Ingalls, R. (1988). *X-ray Absorption: Principles, Applications, Techniques of EXAFS, SEXAFS and XANES*, edited by D. C. Koningsberger & R. Prins, pp. 375–384. New York: Wiley.
- Dalba, G. & Fornasini, P. (1997). *J. Synchrotron Rad.* **4**, 243–255.
- Della-Longa, S. D., Arcovito, A., Girasole, M., Hazemann, J. L. & Benfatto, M. (2001). *Phys. Rev. Lett.* **87**, 155501-1–155501-4.
- Durham, P. J., Pendry, J. B. & Hodges, C. H. (1982). *Comput. Phys. Commun.* **25**, 193–205.
- Filippini, A. (1991). *J. Phys. Condens. Matter*, **3**, 6489–6507.
- Filippini, A. & Di Cicco, A. (1995). *Phys. Rev. B*, **52**, 15122–15134.
- Filippini, A. & Di Cicco, A. (1996). *Phys. Rev. B*, **53**, 9466–9467.
- Fujikawa, T. (1993). *J. Phys. Soc. Jpn.* **62**, 2155–2165.
- George, G. (1999). *EXAFSPAK. SSRL XAFS Analysis Package*. Unpublished.
- Groot, F. M. F. de (1994). *J. Electron Spectrosc.* **67**, 529–622.
- Gropp, W., Lusk, E. & Skjellum, A. (1994). *Using MPI: Portable Parallel Programming with the Message-Passing Interface*. Cambridge, MA: MIT Press.
- Gurman, S. J., Binsted, N. & Ross, I. (1986). *J. Phys. C*, **19**, 1845–1861.
- Krappe, H. & Rossner, H. (2000). *Phys. Rev. B*, **61**, 6596–6610.
- Lee, P. A. & Pendry, J. B. (1975). *Phys. Rev. B*, **11**, 2795–2811.
- McCarthy, M. I., Schenter, G. I., Chacon-Taylor, M. R., Rehr, J. J. & Brown, G. E. Jr (1997). *Phys. Rev. B*, **56**, 9925–9936.
- McKale, A. G., Knapp, G. S. & Chan, S.-K. (1986). *Phys. Rev. B*, **33**, 841–846.
- Manar, F. & Brouder, Ch. (1995). *Physica B*, **208/209**, 79–80.
- Mansour, A. N., Cook, J. W. Jr & Sayers, D. E. (1984). *J. Phys. Chem.* **88**, 2330–2334.
- Natoli, C. R., Misemer, D. K., Doniach, S. & Kutzler, F. W. (1980). *Phys. Rev. A*, **22**, 1104–1108.
- Nesvizhskii, A. I., Ankudinov, A. L. & Rehr, J. J. (2001). *Phys. Rev. B*, **63**, 094412-1–094412-5.
- Newville, M. (2001). *J. Synchrotron Rad.* **8**, 96–100.
- Poiarkova, A. V. & Rehr, J. J. (1998). *Phys. Rev. B*, **59**, 948–957.
- Rehr, J. J. & Albers, R. C. (1990). *Phys. Rev. B*, **41**, 8139–8149.
- Rehr, J. J. & Albers, R. C. (2000). *Rev. Mod. Phys.* **72**, 621–654.
- Rehr, J. J., Mustre de Leon, J., Zabinsky, S. I. & Albers, R. C. (1991). *J. Am. Chem. Soc.* **113**, 5136–5140.
- Rehr, J. J., Stern, E. A., Martin, R. L. & Davidson, E. R. (1978). *Phys. Rev. B*, **17**, 560–565.
- Sayers, D. E., Stern, E. A. & Lytle, F. W. (1971). *Phys. Rev. Lett.* **27**, 1204–1205.
- Schaich, W. L. (1973). *Phys. Rev. B*, **8**, 4028–4032.
- Teo, B.-K. & Lee, P. A. (1979). *J. Am. Chem. Soc.* **101**, 2815–2832.
- Wu, H. & Tong, S. Y. (1999). *Phys. Rev. B*, **57**, 1659–1660.
- Zabinsky, S. I., Rehr, J. J., Ankudinov, A. L., Albers, R. C. & Eller, M. J. (1995). *Phys. Rev. B*, **52**, 2995–3009.
- Zangwill, A. & Soven, P. (1980). *Phys. Rev. A*, **21**, 1561–1572.