Professional Certificate in Density Functional Theory Calculations

Pseudopotentials And Basis Sets

Pseudopotential is a central concept in modern electronic-structure calculations. It is an effective potential that replaces the all-electron Coulomb interaction of the atomic nucleus and its tightly bound core electrons with a smoother, computationally cheaper representation. By removing the need to describe the rapid oscillations of core-electron wavefunctions, pseudopotentials allow the use of relatively modest basis sets while preserving the accuracy of valence-electron properties. In density functional theory (DFT) the most common types are norm-conserving, ultrasoft, and the projector-augmented wave (PAW) method. Each has distinct theoretical foundations, practical advantages, and limitations.

The term core electron refers to electrons that occupy inner shells and are chemically inert for most processes. Because their wavefunctions are highly localized near the nucleus, they contribute little to chemical bonding but dominate the computational cost if treated explicitly. By contrast, valence electrons are those that participate directly in bonding, transport, and optical phenomena. A well-designed pseudopotential reproduces the scattering properties of the true atomic potential for valence states while “freezing” the core states.

Norm-conserving pseudopotentials (NCPP) were introduced by Hamann, Schlüter, and Chiang in the 1970s. The defining feature is that the pseudo wavefunction’s norm within a chosen cutoff radius matches that of the all-electron wavefunction. This condition ensures that the charge enclosed by the pseudo orbital is identical to the true orbital, preserving transferability across different chemical environments. NCPPs are typically generated using either the local density approximation (LDA) or the generalized gradient approximation (GGA) for the exchange-correlation functional, and they can be constructed for scalar-relativistic or fully relativistic treatments. Because the pseudo wavefunctions are still relatively hard near the nucleus, NCPPs often demand a larger plane-wave cutoff energy (e.G., 80–120 Ry for light elements) compared with softer alternatives.

Ultrasoft pseudopotentials (USPP) relax the norm-conservation requirement, allowing the pseudo wavefunctions to be smoother and consequently reducing the plane-wave cutoff dramatically (often to 30–50 Ry). The trade-off is the introduction of an augmentation charge and a non-local operator that restores the correct electron density. USPPs are especially advantageous for transition metals and heavy elements where the d- and f-electrons produce very sharp features in the all-electron potential. However, the additional augmentation terms increase the complexity of the formalism and may affect the convergence of forces and stresses if not carefully converged.

The projector-augmented wave (PAW) method, developed by Blöchl and later refined by Kresse and Joubert, combines the advantages of all-electron calculations with the efficiency of pseudopotentials. In PAW, the true wavefunction is expressed as a sum of a smooth pseudo part plus a set of atom-centered partial waves and corresponding projectors. The method yields exact results (within the chosen exchange-correlation functional) if the partial-wave basis is complete, yet in practice a modest number of partial waves suffices for high accuracy. PAW datasets are commonly distributed as “PAW potentials” and can be used with plane-wave codes such as VASP, GPAW, and Quantum ESPRESSO. A key benefit is that the same PAW dataset can be employed for both total-energy and electronic-structure analyses (e.G., Density of states, band structures) without re-optimizing the basis.

A related term is effective core potential (ECP), which is often used interchangeably with pseudopotential in quantum-chemistry contexts. ECPs are typically expressed in a Gaussian or Slater-type functional form, making them compatible with localized basis sets (see below). While the underlying philosophy mirrors that of pseudopotentials—replacing core electrons with an effective potential—ECPs are usually tailored for molecular calculations and may include scalar-relativistic corrections or spin-orbit coupling explicitly.

The construction of a pseudopotential involves several steps. First, an all-electron reference calculation is performed for the isolated atom, often using a high-quality radial solver. Next, a cutoff radius r_c is chosen for each angular momentum channel (s, p, d, …). Inside r_c, the true radial function is replaced by a smooth analytic form that satisfies continuity of the function and its first few derivatives at the boundary. The non-local projector operators are then derived from the differences between the true and pseudo wavefunctions. Finally, the pseudopotential is tested for transferability by applying it to a series of test compounds, checking properties such as bond lengths, bulk moduli, and ionization energies against experimental data or all-electron calculations.

When selecting a pseudopotential for a specific DFT calculation, several practical considerations arise. The choice of exchange-correlation functional used during generation should match the functional employed in the simulation to avoid systematic errors. Relativistic effects become important for elements beyond the first-row transition metals; scalar-relativistic pseudopotentials incorporate mass-velocity and Darwin terms, while fully relativistic versions also include spin-orbit coupling. The computational cost is tied to the plane-wave cutoff needed to converge the total energy, forces, and stress; thus, USPP or PAW datasets are often preferred for large supercells or high-throughput studies. Conversely, NCPPs may be required when high precision is needed for properties sensitive to the core region, such as hyperfine parameters or core-level spectroscopy.

In addition to plane-wave representations, many DFT codes employ localized basis sets. The two dominant families are Gaussian-type orbitals (GTOs) and numeric atomic orbitals (NAOs). Both families rely on a set of functions centered on atoms that expand the Kohn-Sham orbitals. The choice of basis set determines the balance between accuracy and computational efficiency, and the vocabulary surrounding basis sets is extensive.

A basis function is a mathematical function used to represent an electronic wavefunction. In the context of DFT, the Kohn-Sham orbitals $\psi_i(r)$ are expressed as linear combinations of basis functions $\chi_\mu(r)$: $\Psi_i(r) = \Sigma_\mu c_{i\mu} \chi_\mu(r)$. The coefficients $c_{i\mu}$ are obtained by solving the generalized eigenvalue problem $Hc = ESc$, where H and S are the Hamiltonian and overlap matrices, respectively. Basis functions can be classified by their functional form, spatial extent, and flexibility.

Gaussian-type orbitals are defined as products of a Gaussian radial part $e^{-\alpha r^2}$ and spherical harmonics. Their principal advantage is the analytical evaluation of multi-center integrals, which dramatically speeds up calculations in quantum-chemistry packages such as Gaussian, ORCA, and NWChem. A Gaussian basis set is typically described by a notation such as “6-31G(d)”. The first number (6) indicates the number of primitive Gaussians combined to form a contracted function for the core; the second number (31) describes the split-valence scheme (three primitives for the inner valence, one for the outer valence); the “(d)” denotes the addition of polarization functions that allow the electron density to deform anisotropically. Variants such as “cc-pVnZ” (correlation-consistent polarized valence n-zeta) systematically increase the size of the basis by adding more functions, thereby approaching the complete-basis-set limit.

Numeric atomic orbitals (NAOs) are obtained by solving the radial Schrödinger equation for an isolated atom, possibly with a confinement potential that forces the orbital to vanish beyond a chosen cutoff radius. The resulting functions are tabulated on a radial grid and interpolated during calculations. NAO-based codes such as SIESTA and FHI-aims combine these atom-centered functions with a flexible set of “multiple-ζ” (double-ζ, triple-ζ) and polarization functions. The term “ζ” (zeta) denotes the number of basis functions used to describe each atomic orbital. A single-ζ basis contains one function per valence orbital, a double-ζ adds a second, more diffuse function, and so on. Adding polarization functions (e.G., D-functions on a carbon atom) improves the description of angular flexibility, while “diffuse” functions with small exponent α extend the basis far from the nucleus, which is essential for anions and excited states.

The concept of basis set superposition error (BSSE) is crucial when interpreting interaction energies. BSSE arises because each fragment in a molecular complex can “borrow” basis functions from its partner, artificially lowering the total energy. The counterpoise correction, introduced by Boys and Bernardi, evaluates the energy of each fragment in the full basis of the complex and subtracts this contribution to obtain a more reliable interaction energy. In plane-wave calculations, BSSE is negligible because the basis is delocalized and identical for all fragments, but in localized-basis approaches it can be significant, especially for weakly bound systems.

Another key term is linear-combination-of-atomic-orbitals (LCAO). This phrase simply describes the expansion of Kohn-Sham orbitals in a set of atomic-like functions, which is the foundation of most localized-basis codes. The LCAO method allows for systematic improvement by increasing the number of functions per angular momentum channel (i.E., Moving from single-ζ to triple-ζ) and by adding polarization and diffuse functions. In practice, the choice of LCAO basis is often guided by a trade-off between computational cost and the desired accuracy for a given property (e.G., Geometry, vibrational frequencies, band gaps).

The notion of completeness is central to both pseudopotential and basis-set discussions. A basis set is said to be complete when any square-integrable function can be represented arbitrarily accurately by a linear combination of its elements. In the limit of an infinite plane-wave cutoff or an infinite series of Gaussian primitives, the basis becomes complete, and the calculated properties converge to the exact DFT solution (subject to the approximations of the functional). In real calculations, one monitors convergence by increasing the cutoff energy (for plane waves) or the cardinal number of the Gaussian basis (e.G., Moving from cc-pVDZ to cc-pVQZ) and observing the change in total energy, forces, or other observables.

A related practical term is cutoff energy, denoted E_cut. In plane-wave approaches, the kinetic-energy cutoff defines the maximum |G|^2/2 (where G is a reciprocal lattice vector) of plane waves included in the expansion. Higher E_cut values increase the number of basis functions, improve the representation of rapid spatial variations, and reduce the Pulay stress in geometry optimizations. Typical values range from 30 Ry for soft USPPs to >120 Ry for hard NCPPs. Convergence tests are essential: one should plot total energy versus E_cut and look for a plateau where the change is below a predefined threshold (e.g., 1 meV per atom).

When discussing projector functions in the PAW formalism, the term augmentation sphere is often used. Each atom is associated with a spherical region within which the all-electron and pseudo wavefunctions differ. Inside this sphere, the PAW method reconstructs the true electron density by adding contributions from the partial waves weighted by the projector coefficients. The radius of the augmentation sphere is typically chosen to enclose the region where the core states are significant, but not so large as to cause overlap with neighboring atoms, which would compromise orthogonality.

The term non-local operator appears in both NCPP and USPP formulations. In contrast to the local part of the pseudopotential, which depends only on the radial distance r, the non-local part acts differently on each angular momentum channel. Mathematically, it is expressed as Σ_l |χ_l⟩V_l⟨χ_l|, where |χ_l⟩ are projector functions and V_l are channel-specific potentials. This construction ensures that scattering properties for each l are reproduced accurately. The non-locality is essential for capturing the correct phase shifts of valence electrons as they encounter the core region.

In addition to the basic categories, specialized pseudopotentials exist for particular applications. Hybrid functionals such as HSE06 or PBE0 incorporate a fraction of exact Hartree-Fock exchange. When employing hybrid functionals, one often prefers PAW datasets because the exact exchange term requires an accurate description of the wavefunction near the nucleus. Dispersion-corrected pseudopotentials incorporate empirical van der Waals terms (e.G., DFT-D3) to improve the description of long-range interactions. While the dispersion correction is typically added as a separate energy term, some modern PAW potentials embed the necessary parameters to ensure consistency with the underlying functional.

A practical challenge in pseudopotential work is the management of ghost states. Ghost states are spurious low-energy solutions that arise when the pseudopotential is too soft or poorly constructed, leading to unphysical occupation of non-existent states. Manifestations include negative eigenvalues for unoccupied bands or irregular convergence behavior. Detecting ghost states involves inspecting the projected density of states or performing test calculations on simple molecules. If ghost states appear, one may need to use a harder potential, increase the plane-wave cutoff, or switch to a different pseudopotential library.

The term pseudopotential library refers to a curated collection of pre-generated potentials for many elements. Prominent examples include the PSlibrary, GBRV, VASP PAW, and SG15 sets. Each library adopts specific generation protocols (e.G., Choice of reference configuration, relativistic treatment, exchange-correlation functional) and provides documentation on recommended cutoffs and known issues. Users should consult the accompanying “pseudopotential table” to select the appropriate dataset for their element, oxidation state, and computational method.

Turning to basis sets, the concept of contracted Gaussian functions is essential. A contracted function combines several primitive Gaussians with fixed coefficients, forming a more chemically intuitive orbital that mimics the shape of an atomic subshell. Contraction reduces the number of variational parameters, speeding up self-consistent-field (SCF) iterations. However, excessive contraction can limit flexibility, particularly for describing polarization or charge-transfer effects. Therefore, many modern basis sets employ a “segmented” contraction scheme, where each contracted function is built from a distinct set of primitives, allowing selective decontraction of valence shells when needed.

The notion of basis set completeness error (BSCE) quantifies the difference between the energy obtained with a finite basis and the (theoretical) complete-basis limit. In many studies, the BSCE is estimated by extrapolating results from a series of systematically larger bases, such as the correlation-consistent series. For example, the extrapolation formula E_n = E_∞ + A n^{−3} (where n is the cardinal number) can be used to estimate the infinite-basis energy E_∞. This approach is valuable for benchmarking and for assessing the reliability of calculated reaction energies or barrier heights.

A frequently encountered term in the context of NAO-based codes is numerical integration grid. Since NAOs are defined on a radial mesh, evaluating the exchange-correlation energy and potential requires integration over real space. The grid is typically constructed as a product of radial shells (often using an exponential or logarithmic spacing) and angular points (e.G., Lebedev or spherical-design grids). The density of the grid influences the convergence of total energies and forces; a too-coarse grid can lead to “egg-box” artifacts where the energy oscillates with atomic displacement relative to the grid. Grid convergence is therefore checked by increasing the number of radial points or angular directions until the energy change falls below a target threshold.

In plane-wave calculations, the term k-point sampling denotes the discretization of the Brillouin zone. Since the plane-wave basis is periodic, the electronic states are labeled by crystal momentum k. Accurate integration over the Brillouin zone requires a sufficiently dense mesh of k-points, often generated using Monkhorst-Pack or Gamma-centered grids. The density of the k-point mesh interacts with the plane-wave cutoff: A coarse k-point grid may mask basis-set incompleteness, while an excessively fine grid can waste computational resources if the cutoff is the limiting factor. Convergence tests typically involve increasing both the cutoff and the k-point density and monitoring the total energy and stress.

A practical issue for large-scale calculations is the parallel scalability of pseudopotential and basis-set implementations. Plane-wave codes scale well with the number of processors because the Fourier transforms and matrix-vector multiplications can be distributed efficiently. In contrast, Gaussian-basis codes face challenges due to the dense two-electron integral tensors, which scale roughly as N^4 (with N the number of basis functions). Recent developments such as density fitting (also called resolution-of-the-identity, RI) and Cholesky decomposition reduce the scaling to N^3 or lower, enabling hybrid functional calculations on medium-size systems. Users must be aware of these algorithmic differences when planning high-throughput DFT studies.

The term spin-orbit coupling (SOC) appears in relativistic pseudopotentials. SOC arises from the interaction between an electron’s spin and its motion in the electric field of the nucleus. In the PAW framework, SOC can be included either by generating fully relativistic datasets where the projector functions carry spinor character, or by adding an on-site SOC term to the Hamiltonian after the scalar-relativistic calculation. SOC is essential for materials containing heavy elements (e.G., Bi, Pb, Au) and for phenomena such as topological insulators, Rashba splitting, and magnetic anisotropy.

The phrase frozen core approximation is often used when describing pseudopotentials. It indicates that the core electron density is held fixed throughout the calculation, and only the valence electrons are allowed to relax self-consistently. This approximation dramatically reduces the size of the Hamiltonian matrix and the number of SCF cycles needed for convergence. However, for properties that involve core-level excitations (e.G., X-ray absorption spectroscopy) or for systems where core-valence interactions are strong, an all-electron approach or a “semi-core” pseudopotential (which treats shallow core states as valence) may be required.

A subtle but important vocabulary item is partial-wave cutoff. In PAW, each angular momentum channel has an associated cutoff radius and a set of partial waves (typically a “reference” wave and an “excited” wave). The quality of the PAW reconstruction depends on how many partial waves are included; adding a second partial wave per channel improves the description of high-energy unoccupied states, which is important for calculating accurate band gaps or optical spectra. The term “dual-basis” sometimes appears in literature to denote a PAW dataset that includes both a smooth pseudo basis and a set of all-electron partial waves.

When dealing with periodic systems, the concept of cell-size convergence is crucial. Even with a highly accurate pseudopotential and basis set, finite-size effects can arise from the periodic images of defects, surfaces, or molecules. For charged defects, a compensating background charge is introduced, and correction schemes (e.G., Makov-Payne) must be applied. The choice of a sufficiently large supercell mitigates spurious interactions; convergence tests involve increasing the cell dimensions and monitoring properties such as formation energies or defect levels.

Another term related to basis sets is basis-set pruning. In large molecular systems, one may remove high-energy functions (e.G., Tight d-functions) from regions far from the active site to reduce computational cost while retaining accuracy where it matters most. Pruning strategies are often guided by chemical intuition or automated algorithms that assess the contribution of each function to the total energy. Care must be taken to avoid introducing artificial discontinuities in the potential energy surface.

The self-consistent field (SCF) loop is the iterative procedure that solves the Kohn-Sham equations. In each iteration, the electron density ρ(r) is constructed from the current set of orbitals, the effective potential V_eff[ρ] is built, and the Kohn-Sham Hamiltonian is diagonalized (or solved via iterative methods). Convergence criteria typically involve the change in total energy, the norm of the density difference, and the residual forces. When using ultrasoft or PAW pseudopotentials, the SCF algorithm must also update the augmentation charges and ensure that the orthonormality constraints are satisfied. Mixing schemes (e.G., Pulay, Broyden) accelerate convergence, but the optimal parameters can depend strongly on the hardness of the pseudopotential and the size of the basis.

A specialized term is double-grid technique. In plane-wave codes that employ PAW, the charge density is often represented on a finer grid than the wavefunctions, because the augmentation charges contain high-frequency components. The double-grid approach stores the smooth part of the density on a coarse grid (saving memory) and the augmentation part on a finer grid (ensuring accuracy). This technique is essential for achieving high-precision forces without excessive computational cost.

The local-density approximation (LDA) and generalized gradient approximation (GGA) are the most common exchange-correlation functionals. When generating a pseudopotential, one must decide which functional to adopt because the shape of the effective potential depends on the exchange-correlation energy density. A pseudopotential generated with LDA will generally give slightly different lattice constants and bulk moduli than one generated with GGA, even if the same functional is used in the subsequent calculation. Consequently, many practitioners maintain separate libraries for LDA-compatible and GGA-compatible pseudopotentials.

The phrase transferability test describes the validation step where a pseudopotential is applied to several chemically distinct environments (e.G., Atoms, molecules, solids) to assess whether it reproduces reference data within acceptable error margins. Common test sets include the equilibrium bond lengths of diatomic molecules, the cohesive energies of elemental crystals, and the ionization potentials of atoms. A well-transferable pseudopotential should yield errors no larger than a few percent across the test suite.

In the context of basis sets, the term basis-set superposition correction (BSSE) often appears alongside the counterpoise method. For a dimer AB, the BSSE is estimated as ΔE_BSSE = E_A^AB + E_B^AB − (E_A^A + E_B^B), where E_A^AB is the energy of fragment A computed in the full basis of the AB complex. Subtracting ΔE_BSSE from the raw interaction energy yields a more reliable estimate of the true binding energy. This correction is especially important when using small or heavily contracted Gaussian bases.

A more advanced concept is all-electron PAW. Although PAW is typically considered a pseudopotential method, it can be formulated to retain all electrons by setting the augmentation spheres to include the full core region and by providing a complete set of partial waves. This approach yields results indistinguishable from traditional all-electron methods (e.G., FLAPW) while keeping the computational cost comparable to standard PAW calculations. All-electron PAW is useful for benchmarking and for studying properties that depend sensitively on core states, such as hyperfine couplings.

The term kohn-sham eigenvalue refers to the energies ε_i obtained from solving the Kohn-Sham equations. While these eigenvalues are often interpreted as band energies, only the highest occupied eigenvalue has a rigorous physical meaning (equal to the negative of the ionization potential in exact DFT). The other eigenvalues are useful for constructing density of states, optical spectra, and for serving as starting points in many-body perturbation theory (e.G., GW). The accuracy of these eigenvalues depends on both the pseudopotential quality and the completeness of the basis set.

A practical hurdle when using pseudopotentials for transition metals is the treatment of semi-core states. Semi-core electrons (e.G., 3P in 3d metals) are more tightly bound than valence electrons but can participate in bonding under certain conditions (e.G., High pressure, oxidation). To capture these effects, one can employ a pseudopotential that treats the semi-core electrons as valence (often labeled “pseudopotential with p-semicore”). This increases the number of valence electrons and typically requires a higher plane-wave cutoff, but it yields more reliable predictions of structural and electronic properties.

The phrase hard versus soft pseudopotential captures the trade-off between accuracy and computational cost. Hard pseudopotentials have small cutoff radii and require large plane-wave cutoffs to converge, but they are more transferable across diverse chemical environments. Soft pseudopotentials, with larger cutoff radii, converge quickly but may be less reliable for systems where the valence electrons experience strong core-penetration. Users often start with a soft potential for exploratory calculations and switch to a harder one for final, high-precision results.

In Gaussian-basis calculations, the term basis-set incompleteness error (BSIE) quantifies the residual error after a basis set has been chosen. Unlike BSSE, which is an artifact of overlapping fragments, BSIE is intrinsic to the finite nature of the basis. Strategies to reduce BSIE include adding diffuse functions, increasing the cardinal number, and employing explicit correlation methods (e.G., MP2-F12) that accelerate basis-set convergence.

An important vocabulary item for PAW users is partial-wave projector. The projector functions ⟨p_i| are designed to be orthogonal to the smooth pseudo wavefunctions and to extract the coefficients needed for the augmentation. In practice, a set of projectors is generated for each angular momentum channel, often using the same radial functions as the partial waves but with a different normalization. The quality of the projectors directly influences the accuracy of forces and stress tensors, especially when the system undergoes large deformations.

When performing geometry optimizations, the term Pulay stress correction may arise. In plane-wave calculations with a finite cutoff, the total energy depends on the cell volume due to the discretization of the reciprocal lattice. This leads to an artificial stress (the Pulay stress) that can bias the optimized lattice constant. By converging the cutoff energy sufficiently or by applying a post-hoc correction based on the derivative of the energy with respect to the cutoff, one can mitigate this effect.

The concept of effective potential encompasses the sum of the external potential (from nuclei), the Hartree (electrostatic) potential, and the exchange-correlation potential. In pseudopotential calculations, the external potential is replaced by the pseudopotential, which includes both a local term and a non-local term. The effective potential is the quantity that enters the Kohn-Sham Hamiltonian and determines the shape of the Kohn-Sham orbitals.

A specialized term is meta-GGA functional. Meta-GGA functionals, such as SCAN, depend not only on the electron density and its gradient but also on the kinetic-energy density. When generating pseudopotentials for meta-GGA calculations, one must ensure that the reference atomic calculation uses the same functional form, because the kinetic-energy density influences the shape of the effective potential. Meta-GGA–compatible pseudopotentials are increasingly available and enable more accurate predictions of lattice constants and formation energies.

The phrase basis-set linear dependence describes a situation where two or more basis functions become nearly collinear, leading to an ill-conditioned overlap matrix. This issue is common in large Gaussian bases with many diffuse functions, where the exponents become very small and the functions overlap extensively. Linear dependence manifests as very small eigenvalues of the overlap matrix, causing numerical instability in SCF convergence. Remedies include removing redundant functions, increasing the threshold for linear dependence detection, or using a more compact contraction scheme.

In the context of plane-wave calculations, the term egg-box effect refers to artificial variations in total energy as atoms are displaced relative to the underlying real-space grid. This effect arises because the discretized representation of the potential and density can introduce spurious periodicity. The egg-box error diminishes with higher plane-wave cutoffs and finer FFT grids. Users can assess the magnitude of the effect by translating a crystal by a fraction of the grid spacing and monitoring the energy variation.

A key term for high-throughput DFT is standardized workflow. Such a workflow typically includes a sequence of steps: (1) Selection of a pseudopotential from a curated library, (2) convergence testing for cutoff energy and k-point density, (3) geometry relaxation with appropriate convergence criteria, (4) static calculation for accurate total energy, and (5) post-processing for properties of interest (e.G., Band gap, elastic constants). Standardization ensures reproducibility across many calculations and facilitates data mining.

The notion of spin-polarized calculation is relevant when dealing with magnetic materials. In spin-polarized DFT, the electron density is split into up- and down-spin components, $\rho\uparrow(r)$ and $\rho\downarrow(r)$, and the Kohn-Sham equations are solved separately for each spin channel. Pseudopotentials can be generated for either spin-unpolarized or spin-polarized reference atoms; however, the pseudopotential itself does not depend on spin, because the core does not carry a net magnetic moment. The exchange-correlation functional, however, must be capable of handling spin dependence (e.G., LSDA, spin-GGA).

An advanced term is non-collinear magnetism. In this framework, the spin direction is allowed to vary continuously in space, rather than being constrained to a global up/down axis. Implementing non-collinear magnetism requires spinor wavefunctions and a pseudopotential that includes spin-orbit coupling. The resulting Hamiltonian contains off-diagonal spin terms, and the SCF procedure must converge both charge and magnetization densities.

The phrase core-level shift appears when discussing spectroscopy. Core-level binding energies can be calculated using the Δ-SCF method, where a core electron is removed (or excited) and the total energy difference is taken. Accurate core-level shifts demand an all-electron description of the core region, which can be achieved with PAW datasets that include the relevant core states as valence or with specialized ”core-hole” pseudopotentials that mimic the presence of a vacancy.

A practical term is pseudopotential file format. Different codes adopt distinct conventions: For example, the UPF format is used by Quantum ESPRESSO, the POTCAR format by VASP, and the PAW XML format by GPAW. The file contains the radial functions, projector definitions, and metadata such as the recommended cutoff energy, the exchange-correlation functional used, and relativistic flags. Users must ensure compatibility between the pseudopotential file and the DFT code version.

In Gaussian-basis packages, the term basis-set file refers to a text file that lists the exponents and contraction coefficients for each shell. Common formats include the ”.Gbs” file used by Gaussian and the ”.Basis” file used by NWChem. The file also specifies the angular momentum (s, p, d, f) and any polarization or diffuse augmentations. Maintaining a library of basis-set files allows rapid switching between different levels of theory.

The concept of finite-difference grid arises in real-space DFT codes (e.G., Octopus, PARSEC). Instead of expanding wavefunctions in plane waves or localized orbitals, the Kohn-Sham equations are discretized directly on a uniform or adaptive grid. Pseudopotentials in this context are often represented as tabulated radial potentials that are interpolated onto the grid. The grid spacing determines the effective basis set: Finer grids provide higher accuracy but increase memory usage. Convergence tests involve reducing the grid spacing until the total energy changes by less than a prescribed tolerance.

A specific challenge for heavy elements is the inclusion of scalar-relativistic effects. These effects arise from the mass-velocity and Darwin terms in the Dirac equation and modify the effective potential felt by valence electrons. Pseudopotentials that incorporate scalar-relativistic corrections are labeled accordingly (e.G., “SR-PAW” or “relativistic USPP”). Neglecting these corrections can lead to significant errors in bond lengths and cohesive energies for elements beyond the first-row transition metals.

When dealing with periodic surfaces, the term slab model is ubiquitous. A slab consists of a finite number of atomic layers separated by vacuum. Pseudopotentials must be chosen to describe both the surface atoms and any adsorbates accurately. Convergence with respect to slab thickness and vacuum size must be verified, as insufficient vacuum can cause spurious interaction between periodic images, altering adsorption energies.

The notion of charge-density convergence is central to all DFT calculations. The SCF loop iterates until the change in charge density between successive steps falls below a target threshold (e.G., 10^{-6} E). In plane-wave codes, the density is represented on a reciprocal-space grid; in localized-basis codes, it is reconstructed from the density matrix. Tight convergence is essential for reliable forces, especially when calculating phonon spectra or performing molecular dynamics.

A useful term is preconditioning. In iterative diagonalization methods (e.G., Conjugate-gradient, Davidson), preconditioning improves the convergence rate by approximating the inverse of the Hamiltonian’s diagonal part. For plane-wave codes, the kinetic energy operator serves as an effective preconditioner, while for Gaussian-basis codes, a diagonal approximation of the Fock matrix is often employed.

The phrase density functional perturbation theory (DFPT) describes a linear-response approach to compute phonons, dielectric constants, and electron-phonon coupling without resorting to finite differences. DFPT relies on the same pseudopotentials and basis sets used in the ground-state calculation, but the perturbative equations require additional response functions (e.G., The derivative of the Kohn-Sham potential with respect to atomic displacements). Accurate DFPT results depend on well-converged plane-wave cutoffs and k-point meshes, as well as on pseudopotentials that faithfully reproduce the response of the core region.

A term that appears in the context of hybrid functionals is exact-exchange screening. Hybrid functionals like HSE06 introduce a screened Coulomb interaction to limit the range of the Hartree-Fock exchange term, reducing computational cost. When using PAW, the exact-exchange operator must be evaluated on both the smooth pseudo wavefunctions and the all-electron partial waves. Efficient implementations use the auxiliary-density matrix method or the adaptively compressed exchange operator to accelerate calculations.

The expression basis-set extrapolation refers to the practice of estimating the infinite-basis limit from a series of finite calculations. In Gaussian-basis work, one may apply the two-point extrapolation formula E_n

$= E_\infty + A e^{-\alpha n}$, where n denotes the cardinal number (D, T, Q). In plane-wave calculations, extrapolation can be performed by fitting the total energy as a function of $1/E_{cut}$, assuming a linear dependence for sufficiently high cutoffs. Extrapolation is valuable when high-precision energies are needed but computational resources limit the attainable cutoff.

A practical vocabulary item for PAW is augmentation charge density.