Information-theoretic classifications and quantifications of the symmetries and pseudo-symmetries in two-dimensional experimental data from crystals and quasicrystals

P. Moeck

Department of Physics, Portland State University, Portland/OR, USA
pmoeck@pdx.edu

Keywords: Symmetry, Pseudo-Symmetry, Information Theory

Recently developed information-theory-based methods enable objective classifications and quantifications of the crystallographic [1-5] and non-crystallographic [6] symmetries in noisy experimental data that are deemed to be translation periodic or quasi-periodic in two dimensions (2D). These classifications are objective because they are based solely on the experimental data via the fulfillment or violation of numerical inequalities that are based on pair-wise ratios of geometric Akaike Information Criterion (G-AIC) values for non-disjoint geometric models of the data. Because the models are in minimal supergroup to maximal subgroup relationships with each other as far as their symmetries are concerned, an a priori estimate of the noise level in the experimental study is not required. Confidence level can be assigned for the selection of a geometric model that features the symmetry of a minimal supergroup over a non-disjoint model that features a maximal subgroup. G-AIC values are in essence geometric bias corrected sums of squared residuals of the difference between the experimental data and geometric models of that data. The type of microscope or diffraction apparatus that has been used for the recording of the experimental data is immaterial for the application of the new methods. Pseudo-symmetries [7] can reliably be distinguished from genuine symmetries [1], even in the presence of large amounts of generalized noise [2]. Generalized noise includes all effects of (unavoidably) imperfect recordings of experimental data, all kinds of rounding effects and numerical approximations by any kinds of data processing algorithms, and all structural defects in crystalline and quasicrystalline real-world material samples. When there are many noise sources and the effects of none of these sources dominate, the resulting generalized noise is approximately Gaussian distributed (by generalizations of the central limit theorem of statistics [6]). Such an approximate distribution is the precondition for the application of G-AIC framework [8]. After an information theoretic symmetry classification has been made, one obtains a good a posteriori estimate of the noise level of the experimental study as a byproduct [1].

Conditional symmetry model probabilities, i.e. so called geometric Akaike weights, within model sets [2, 3] can also be calculated on the basis of the G-AIC values of the individual geometric models of the experimental data. These weights represent the probability that a certain geometric model of the experimental data is the Kullback-Leibler best model in the selected model set. The mathematical feature that probabilities need to be multiplied when one wants to obtain joint probabilities aids the distinction between genuine symmetries and pseudo-symmetries in experimental data [3] on a quantitative basis. The prevailing common practice in materials science is, by stark contrast, to make crystallographic symmetry classifications on the basis of subjective judgments whenever an unknown crystalline or quasicrystalline sample is involved. Those classifications are bound to be misleading or false on occasions, especially when the images or diffraction patterns feature a comparatively large amount of generalized noise, metric specializations [7], and/or pseudo-symmetries.

Note that the information-theoretic methods deliver only probabilistic crystallographic or non-crystallographic symmetry classifications as it is fundamentally unsound to assign abstract mathematical concepts such as a single 2D Bravais lattice type, a crystallographic or non-crystallographic projected Laue class, a point symmetry group, or a plane symmetry group with 100 % certainty to a real-world image or diffraction pattern from a crystal or quasicrystal. The new methods quantify deviations from symmetries, which can be interpreted as providing "error-bars" on symmetry measurements. Experimental atomic-resolution transmission electron microscope images (both in parallel illumination and the scanning probe mode) as well as scanning tunneling microscope images serve as examples for the demonstration of the image-based symmetry classification and quantification methods. Experimental selected-area electron diffraction spot patterns and precession electron diffraction patterns serve as examples for the demonstration of the diffraction-pattern based counterparts [4, 5] of these methods.


Department of Physics, Portland State University, Portland/OR, USA
pmoeck@pdx.edu

Keywords: Symmetry, Pseudo-Symmetry, Information Theory

Recently developed information-theory-based methods enable objective classifications and quantifications of the crystallographic [1-5] and non-crystallographic [6] symmetries in noisy experimental data that are deemed to be translation periodic or quasi-periodic in two dimensions (2D). These classifications are objective because they are based solely on the experimental data via the fulfillment or violation of numerical inequalities that are based on pair-wise ratios of geometric Akaike Information Criterion (G-AIC) values for non-disjoint geometric models of the data. Because the models are in minimal supergroup to maximal subgroup relationships with each other as far as their symmetries are concerned, an a priori estimate of the noise level in the experimental study is not required. Confidence level can be assigned for the selection of a geometric model that features the symmetry of a minimal supergroup over a non-disjoint model that features a maximal subgroup. G-AIC values are in essence geometric bias corrected sums of squared residuals of the difference between the experimental data and geometric models of that data. The type of microscope or diffraction apparatus that has been used for the recording of the experimental data is immaterial for the application of the new methods. Pseudo-symmetries [7] can reliably be distinguished from genuine symmetries [1], even in the presence of large amounts of generalized noise [2]. Generalized noise includes all effects of (unavoidably) imperfect recordings of experimental data, all kinds of rounding effects and numerical approximations by any kinds of data processing algorithms, and all structural defects in crystalline and quasicrystalline real-world material samples. When there are many noise sources and the effects of none of these sources dominate, the resulting generalized noise is approximately Gaussian distributed (by generalizations of the central limit theorem of statistics [6]). Such an approximate distribution is the precondition for the application of G-AIC framework [8]. After an information theoretic symmetry classification has been made, one obtains a good a posteriori estimate of the noise level of the experimental study as a byproduct [1].

Conditional symmetry model probabilities, i.e. so called geometric Akaike weights, within model sets [2, 3] can also be calculated on the basis of the G-AIC values of the individual geometric models of the experimental data. These weights represent the probability that a certain geometric model of the experimental data is the Kullback-Leibler best model in the selected model set. The mathematical feature that probabilities need to be multiplied when one wants to obtain joint probabilities aids the distinction between genuine symmetries and pseudo-symmetries in experimental data [3] on a quantitative basis. The prevailing common practice in materials science is, by stark contrast, to make crystallographic symmetry classifications on the basis of subjective judgments whenever an unknown crystalline or quasicrystalline sample is involved. Those classifications are bound to be misleading or false on occasions, especially when the images or diffraction patterns feature a comparatively large amount of generalized noise, metric specializations [7], and/or pseudo-symmetries.

Note that the information-theoretic methods deliver only probabilistic crystallographic or non-crystallographic symmetry classifications as it is fundamentally unsound to assign abstract mathematical concepts such as a single 2D Bravais lattice type, a crystallographic or non-crystallographic projected Laue class, a point symmetry group, or a plane symmetry group with 100 % certainty to a real-world image or diffraction pattern from a crystal or quasicrystal. The new methods quantify deviations from symmetries, which can be interpreted as providing “error-bars” on symmetry measurements. Experimental atomic-resolution transmission electron microscope images (both in parallel illumination and the scanning probe mode) as well as scanning tunneling microscope images serve as examples for the demonstration of the image-based symmetry classification and quantification methods. Experimental selected-area electron diffraction spot patterns and precession electron diffraction patterns serve as examples for the demonstration of the diffraction-pattern based counterparts [4, 5] of these methods.