POLARIZATION FOR BACKGROUND REDUCTION IN EDXRF ---
THE TECHNIQUE THAT DOES INDEED WORK

Richard W. Ryon
Lawrence Livermore National Laboratory, ret.
P.O. Box 808, Livermore, CA 94551-0808

ABSTRACT

As with all electromagnet radiation, polarization of X-rays is a general phenomenon. Such polarization has been known since the classic experiments of Barkla in 1906. The general implementation of polarization to X-ray analysis had to await the fixed geometry of energy-dispersive systems. The means of optimizing these systems is shown in this review paper. Improved detection limits are the result.

INTRODUCTION

I first heard of using polarized X-rays while talking with several colleagues in the lobby of the Brown Palace Hotel in downtown Denver (where the Denver X-ray Conference was held in the early and mid 1970’s). Discussion centered on a paper several people had heard by John Young, Ron Vane, and Patrick Lenehan1) a few months before in 1973. These researchers described their experiments in producing and using polarized X-rays for background reduction. It transpired that these researchers had been inspired by the section on polarized X-rays in the classic Compton and Allison book2), first published in 1935. If X-rays were scattered twice in orthogonal directions, intensity vanished due to polarization effects. Wouldn’t this be a good way to suppress the original exciting radiation from an X-ray tube that contributes to background in a fluorescent spectrum? See Figure 1. Shortly later, Tom Dzubay, B.V. Jarrett, and Joe Jaklevic published their widely referenced paper “Background Reduction in X-ray Fluorescence Spectra using Polarization3).” They demonstrated a five-fold reduction in Compton scattering of the source radiation and an improvement in peak-to-background ratios. They said, “The inherent

Figure 1. Modern energy disperse X-ray spectrometers that take advantage of polarization effects use the orthogonal geometry of Compton’s early experiments for investigating polarization (detector perpendicular to the first scatter plane, along the y-axis on the right).
This document was presented at the Denver X-ray Conference (DXC) on Applications of X-ray Analysis.

Sponsored by the International Centre for Diffraction Data (ICDD).

This document is provided by ICDD in cooperation with the authors and presenters of the DXC for the express purpose of educating the scientific community.

All copyrights for the document are retained by ICDD.

Usage is restricted for the purposes of education and scientific research.

DXC Website
– www.dxcicdd.com

ICDD Website
- www.icdd.com
decrease in counting rate with increased peak to background ratio constitutes the principal limitation of the polarization method. This can be compensated to a degree by an increase in the X-ray tube output ---.” Reading between the lines, one might rephrase this to say, “While interesting in concept, intensity reduction due to geometric and scatter inefficiency doom the use of polarized X-rays.” This widely-held judgement inspired the subtitle for this paper.

Other researchers at Lawrence Livermore Lab followed the lead of Dzubay and colleagues. David Camp teamed with Leon Kaufman at the University of California Medical Center to measure non-radioactive tracers in medical applications using a polarized X-ray source. A paper by Rich Howell and Bill Pickles gave some mathematical guidance in the form of an equation showing the dependence of signal-to-noise ratio on the angular divergence of the collimators. It was at this point that Dave Camp persuaded me to become involved, and I found internal funding for a small effort. Around this time, I found papers on polarization by Peter Wobrauschek and colleagues at the Atom Institute in Vienna, and related papers using the same orthogonal geometry by P. Stanzenieks and Eva Selin-Lindgren at Chalmers University in Göteborg, Sweden and a related paper by Bisgård, Laursen, and Nielsen in Copenhagen, Denmark. There were many other papers, but it is these people who influenced me the most.

DISCUSSION OF DESIGN CONSIDERATIONS

After the first explorations, the issue regarding practicality raised in the Dzubay paper evolved into several unanswered question about how best to design a spectrometer that made use of the demonstrated advantages of a polarized system:

1.) Is there an optimum collimation? As the aspect ratio (width/length) decreases, polarization increases and the fraction of the primary beam scattered into the detector decreases. However, the fluorescence signal intensity also decreases. Is the polarization method indeed doomed by geometric inefficiency and the physics of producing a polarized beam, or can these competing influences on the quality of spectra be optimized?

2.) When optimized, does polarization offer advantages over similarly optimized other methods of fluorescence excitation, such as direct beam, filtered beam, or secondary fluorescer?

3.) Can anything be gained by using a bundle of collimators? This is the technique used in mammography to cut down on the blurring of the radiographic shadow caused by scattering. Such bundled small collimators were proposed in the Dzubay paper, but nobody to my knowledge ever investigated the suggestion. By integrating over a large number of highly polarized but weak beams, is there a net gain in detection limits in a given length of time?

When considering these questions, we should understand that peak-to-background ratios do not tell the entire story, but signal-to-noise ratios do. Signal-to-noise is defined by

\[ S/N = \text{Signal}/(\text{Background} + \text{Signal})^{1/2}. \]  \hspace{1cm} \text{Eqn. 1}

S/N converts directly to minimum detection limits

\[ \text{MDL} = 3 \times \text{concentration} / (S/N) \]  \hspace{1cm} \text{Eqn. 2}
for a standard specimen. If background could be totally eliminated (say, by using a 100% polarized beam for excitation), the signal to noise ratio becomes \((\text{signal counts})^{1/2}\). We need a large S/N, and that requires a strong signal if we cannot wait forever to accumulate that signal!

To answer the first question regarding optimum collimation, let us consider what happens to the fluorescent signal and the background as the solid angle \((\Omega)\) of the collimators is changed. (A rigorous analysis would consider scatter angles, but these are closely related to the more easily considered collimator angles.) We shall use the simple case where all the collimators are the same length and diameter, where noise comes entirely from random fluctuations in the background, and that background comes entirely from scattered source radiation. Then, both the fluorescent peak intensity and the scattered intensity after three collimators will be proportional to \(\Omega^3\). However, there is an additional factor for the scattered radiation, namely, \((1-P)\) where \(P\) is the polarization. Now, the polarization is directly related to the solid angle:

\[
P = 1 - \left(\frac{4}{\pi}\right) \Omega .
\]

Therefore,

\[
\text{Signal}_{\text{polarized}} \propto \Omega^3 \quad \text{Eqn. 4}
\]

\[
\text{Background}_{\text{polarized}} \propto \Omega^3 (1-P) \propto \Omega^3 \left(\frac{4}{\pi}\right) \Omega \propto \Omega^4 \quad \text{Eqn. 5}
\]

and

\[
\text{S/N}_{\text{polarized}} = \frac{\text{Signal}}{\sqrt{\text{Background}}} \propto \frac{\Omega^3}{(\Omega^4)^{1/2}} \propto \Omega \quad \text{Eqn. 6}
\]

This direct relationship between signal-to-noise and solid angle for Barkla polarizers does not offer optimization in the sense that there is a maximum in a quadratic or higher order equation, or when there are inversely related factors. It just says that we need large solid angles.

To answer the second question regarding comparison to other methods of excitation, we can follow the same procedure as above. Excitation using a secondary fluorescer also requires three collimators. (It should be noted that an optimized system would use orthogonal geometry, to suppress background radiation from the X-ray tube due to polarization effects.) Thus, we have

\[
\text{Signal}_{\text{secondary}} \propto \Omega^3 \quad \text{Eqn. 7}
\]

\[
\text{Background}_{\text{secondary}} \propto \Omega^3 \quad \text{Eqn. 8}
\]

\[
\text{S/N}_{\text{secondary}} = \frac{\text{Signal}}{\sqrt{\text{Background}}} \propto \frac{\Omega^3}{(\Omega^3)^{1/2}} \propto \Omega^{3/2} \quad \text{Eqn. 9}
\]

In orthogonal geometry, the X-ray tube spectrum is suppressed. However, there is no factor operating to suppress the characteristic lines from the secondary fluorescer, and they appear as strong elastically and inelastically scattered peaks, giving rise to a functionally lower signal to noise ratio than in the polarized case.

For direct excitation, or filtered direct excitation, there are only two collimators and we have

\[
\text{Signal}_{\text{direct}} \propto \Omega^2 \quad \text{En. 10}
\]

\[
\text{Background}_{\text{direct}} \propto \Omega^2 \quad \text{Eqn. 11}
\]

\[
\text{S/N}_{\text{direct}} = \frac{\text{Signal}}{\sqrt{\text{Background}}} \propto \frac{\Omega^2}{(\Omega^2)^{1/2}} \propto \Omega \quad \text{Eqn. 12}
\]
I shall hold off further discussion of these functional relationships for the moment, except to note that in every case it is best to seek the strongest signal by opening up the collimators to give large solid angles.

The conclusion that we need large solid angles bodes ill for the third question regarding bundles of collimators. However, we can have a large number of individual collimators and the area in the view of the collimators can remain large, so there may be hope. The packing density for cylinders in a cylinder is

\[ \text{Density} = \frac{\pi}{2 \sqrt{3}} = 0.9069 \text{ Eqn. 13} \]

from which the number of cylindrical collimators that can fit into a larger cylinder for idealized close packing when the number is very large can be calculated:

\[ N = 0.9069 \left( \frac{R}{r} \right)^2 \text{ Eqn. 14} \]

where \( R \) is the radius of the large cylinder and \( r \) is the radius of the small collimators. The idealized number never quite fits. For instance, when \( N \) is 10, the idealized number is 13.2. However, we can use Equation 14 in the parametric analysis to see the shape of things. The analysis is the same as for a single collimator, except that we multiply all intensities by \( N \), the number of such collimators. We therefore have

\[ \text{Signal} (N \text{ polarized}) \propto N \Omega^3 \text{ Eqn. 15} \]
\[ \text{Background} (N \text{ polarized}) \propto N \Omega^3 (1-P) \propto N \Omega^3 \left( \frac{4}{\pi} \right) \Omega \]
\[ \propto N \Omega^4 \text{ Eqn. 16} \]
\[ \text{S/N} (N \text{ polarized}) \propto N \Omega^3 \left( \frac{\Omega_{\text{large}}}{\Omega} \right)^{1/2} \propto N^{1/2} \Omega \text{ Eqn. 17} \]

The above equations emphasize that everything is the same as the single polarized beam, except that \( N > 1 \) for the present case. However, the number of collimators and the solid angle of those collimators are related. We can therefore reduce the number of parameters. From the definition of solid angles when divergences are small and equation 14, it follows directly that

\[ N = 0.9069 \left( \frac{\Omega_{\text{large}}}{\Omega} \right) \text{ Eqn. 18} \]

where \( \Omega_{\text{large}} \) is the solid angle for the cylinder in which the small collimators fit. Then, equations 15, 16, and 17 become

\[ \text{Signal} (N \text{ polarized}) \propto N^{-2} \left( 0.9069 \Omega_{\text{large}} \right)^3 \text{ Eqn. 19} \]
\[ \text{Background} (N \text{ polarized}) \propto N^{-3} \left( 0.9069 \Omega_{\text{large}} \right)^4 \text{ Eqn. 20} \]
\[ \text{S/N} (N \text{ polarized}) \propto N^{-1/2} \left( 0.9069 \Omega_{\text{large}} \right) \text{ Eqn. 21} \]

Thus, we see the best measurement is made when \( N = 1 \), that is, with a single collimator and not with a bundle. (This simplified analysis ignores the total reflection phenomenon from the interior walls of each tube, which will increase the throughput.) Perhaps this is why the original suggestion of long ago dropped into oblivion.

It might be worth noting in passing that the above parametric approach illustrates why scatter-suppression collimators in front of the film in mammography does work. The shadow caste in radiography is independent of collimation, as each point in the object maps directly onto the film when a (ideal) point source of radiation is used. However, scatter can come from anywhere, and discriminating against off-axis scatter will improve S/N. We have

\[ \text{Signal} (\text{radiograph}) \propto \text{constant} \text{ Eqn. 22} \]
The image improves with smaller collimators.

A parameter that has not been considered to this point is the power of the X-ray tube. When doing analyses near detection limits, scatter of source radiation into the detector dominates the count rate, and analyte fluorescence is negligible. Therefore, as collimation is tightened to discriminate against scatter, tube power can be increased proportionately to scatter decrease without exceeding the count rate of the detector. If power is increased inversely to the scatter, we have

\[
\begin{align*}
\text{Signal} \propto \Omega^3 / \Omega^4 = \Omega^{-1} \\
\text{Scatter} \propto \Omega^4 / \Omega^4 = 1 \\
\text{S/N} \propto \Omega^1 \\
\text{Signal} \propto N^{-2} \Omega_{\text{large}}^3 / N^{-3} \Omega_{\text{large}}^4 = N \Omega_{\text{large}}^{-1} \\
\text{Background} \propto N^{-3} \Omega_{\text{large}}^4 / N^{-3} \Omega_{\text{large}}^4 = 1 \\
\text{S/N} \propto N \Omega_{\text{large}}^{-1}
\end{align*}
\]

The above results are gathered together for easy comparison in Table 1 below.

<table>
<thead>
<tr>
<th></th>
<th>S/N @ const pwr</th>
<th>S/N @ pwr ( \propto \text{scat}^{-1} )</th>
<th>pwr compensation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Polarized</td>
<td>( \Omega )</td>
<td>( \Omega^{-1} )</td>
<td>( \Omega^4 )</td>
</tr>
<tr>
<td>Polarized, bundle</td>
<td>( N^{1/2} \Omega_{\text{large}} )</td>
<td>( N \Omega_{\text{large}}^{-1} )</td>
<td>( N^3 \Omega_{\text{large}}^{-4} )</td>
</tr>
<tr>
<td>Secondary fluorescer</td>
<td>( \Omega^{3/2} )</td>
<td>1</td>
<td>( \Omega^{-3} )</td>
</tr>
<tr>
<td>Direct, filtered direct</td>
<td>( \Omega )</td>
<td>1</td>
<td>( \Omega^2 )</td>
</tr>
</tbody>
</table>

Table 1. Signal-to-noise dependency on collimator solid angles when power is constant and when it is unlimited (inversely proportional to scatter).

We see that the signal-to-noise ratio for the secondary fluorescer increases somewhat more rapidly with solid angle than for the other excitation methods, giving it a functional advantage, but starts from a lower point. In all cases, it is best to use the largest possible collimator apertures when power is limited. However, polarized systems clearly benefit when power is increased to compensate for reduced intensity in a more highly collimated system. The bundled collimator has an additional advantage when there is plenty of power.
From the forgoing, a question arises: just how much power increase is necessary to give polarized systems an advantage? The answer is seen in the functional dependencies in Table 1. As collimation is tightened to produce higher polarization, X-ray tube power is increased inversely with scatter to compensate for intensity reduction. Consider what power increase is necessary to increase S/N by a factor of 2. For the polarized case with single collimators, the solid angle is halved. The power must be increased by a power of 4, becoming $2^4 = 16$. For the polarized case with collimator bundles, the number $N$ must be increased by a factor of 2 to produce an increase in S/N of 2. (This is equivalent to halving the solid angle, equation 14.) In this case, the power need only be raised by a power of 3, which is $2^3 = 8$. There is indeed some advantage to using a cluster of collimators when tube power is considered: the area remains fixed as polarization increases. For the other two cases, there is no dependence of S/N on solid angle, as power and intensity are balanced. The discriminating factor is that of polarization.

Power increases of factors of 8 and 16 are rather steep to increase S/N by a factor of 2, but such increases are within the range of practicality. However, there is clearly a practical limit on how much help polarization can be in improving detection limits.

It is not possible to say from the foregoing which method of excitation will be superior in practice and in each particular case. For instance, we know that a monochromatic source has a small background under the fluorescent peaks. This infers that a nearly monochromatic source from secondary fluorescence or diffraction has a clear advantage. However, excitation efficiency of elements with absorption edges far removed from the source energy is greatly diminished. Therefore, a polychromatic, polarized source may be advantageous for general use. Another consideration is noise generated by a strong monochromatic source in the detector itself by Compton scatter and incomplete charge collection; a monochromatic source does not translate into negligible background. A monochromatic, polarized source (produced by diffraction at $2\theta = \pi/2$) would not have this detrimental effect. Well-designed Monte Carlo simulations can help to refine answers, and also significantly help in optimizing an engineered system. We shall discuss this simulation technique a little later. For now, we shall turn to experimental science; the above analysis offers guidance to experiments.

I had an apparatus constructed in which collimator lengths and apertures could readily be changed in order to study Barkla polarizers and other methods of excitation. It was found that with boron carbide polarizers with enough thickness to give decent intensities, and with organic specimens of reasonable thickness, polarization was limited to around 85-90% due to multiple scatter. Hence, collimators were adjusted to give geometrically defined polarization at the same level as this limitation. Maximum counting rates could be achieved at these solid angles with our high wattage X-ray tubes. These series of experiments are summarized in Figure 2 taken from a paper presented at the 1976 Denver X-ray Conference. It is clear that for a wide range of elements, the strong suit of energy disperse analysis, that a polarized source has a distinct advantage. This is due to the bremsstahlung component of the tube radiation. This continuum radiation provides efficient excitation (as in direct excitation) while polarization reduces the background (unlike direct excitation). For elements with absorption edges near the characteristic radiation of the X-ray source, there is little difference among the excitation methods. The 85% polarized Mo tube radiation had similar detection limits as the yttrium secondary source for
elements with absorption edges near the Y characteristic lines, but about 3 times lower for Fe and lighter elements.

After this initial work, I had the good fortune to go to Vienna to work with Peter Wobrauschek and Hannes Aiginger, to use their equipment and learn from their experience and expertise. The result was an empirical demonstration that polarization does indeed offer advantages for general analysis⁷, and that there is a place for both Barkla and Bragg polarizers. Bragg polarizers have certain advantages over Barkla polarizers. Scatter is directional instead of (more-or-less) isotropic so that intensity does not fall off as rapidly with distance. Fine-focus X-ray tubes with their higher specific heat loads can be used. Lastly, Johann and Johansson focusing geometries can be used to increase solid angles while maintaining a high degree of polarization. The problem is to find strongly diffracting lines with \(2\theta = \pi/2\) at higher energies. Quintin Johnson compiled an extensive list of possible crystals for various characteristic lines⁸. Among his suggestions was garnet; this family of materials can have a range of cell constants and is therefore adaptable to several anode characteristic lines, such as Cu, Zr, Mo, Rh, and Ag. It can be seen in Figures 3a and 3b that Bragg polarizers are effective when good reflections are available for the X-ray tube characteristic lines.

A nice comparison between Bragg and Barkla polarizers was subsequently made at the University of Bremen by Birgit Kanngiesser, Burkhard Beckhoff, and Walter Swoboda⁹. They compared highly ordered and disordered graphite. The conclusion is that there is little difference between Bragg and Barkla polarizers when using Mo K\(\alpha\) radiation at the relatively large solid angles required for decent counting rates. This is the result of the predominance of Compton scattering. The diffraction line has a very narrow angular range and thus its integrated reflectivity is relatively low. At lower energies, there are reflections that pick up the bremsstrahlung from the X-ray tube, giving diffraction lines considerably stronger than the Compton scatter. Measured detection limits at higher energies are virtually the same for highly ordered pyrolytic and ordinary graphite; at lower energies, the crystalline form yields lower detection limits because of the efficient excitation by diffraction lines. Please see Figure 4 for the exciting spectra, and Figure 5 for detection limits of solutions. At the present state of development, Bragg polarization is
Effective only at lower energies where there are good, strong diffraction lines. At higher energies, Barkla polarizers provide more intensity when integrated over the relatively large solid angles required for decent counting rates.

We are concerned here with polarized X-rays derived from conventional X-ray sources. It should be at least briefly mentioned that synchrotron light sources are ideal sources for polarized X-rays. Because electrons in storage rings are centripetally accelerated in a plane, the radiation emitted is polarized in this plane. Synchrotron radiation can therefore be used as described above to minimize background from scatter. However, synchrotron radiation is also ideally suited to total reflection X-ray fluorescence (TXRF). TXRF has an inherently low background when the sample mass on the support is small because the excitation radiation penetrates only a few nanometers into the substrate. However, when there is a residue, say from salts or organic materials, there is significant scatter. By placing the detector in...
the plane of the ring to take advantage of the polarization, little scatter is detected. The Vienna group of Peter Wobrauschek, Christina Streli, and colleagues has proven these ideas in practice\textsuperscript{10}. They obtain detection limits in the femtogram (10\textsuperscript{-15}) range.

Traditionally, \textit{in vivo} measurements have used radioisotopic sources and two-axis geometry. The very large Compton peaks tend to overwhelm the counting electronics and bury the lines of interest in background from long Compton tails. These measurements would clearly benefit from well-designed polarized systems, and Monte Carlo simulations are the way to efficiently optimize systems before starting to turn metal in the fabrication shop. Robin Gardner and colleagues at the University of North Carolina pioneered the use of Monte simulations in XRF. This technique has recently been applied to the \textit{in vivo} measurement of lead in bone (environmental exposure) and platinum in pituitary gland (chemotherapy). A recent application by H.H. Al-Ghorabie Fayez in Saudi Arabia is the measurement of cadmium in kidney\textsuperscript{11}. He used the public-domain program EGS4, which was adapted in 1993 by Namito to include polarization. In a parametric analysis of collimators, excitation voltage, polarizer materials and angles, an optimized system was defined. The results are shown in Figure 6. The simulated system is predicted to have better detection limits than those reported in classic papers by Ahlgren, Mattsson, and Christoffersson. The proof is anticipated.

A recent paper by Joachim Heckel demonstrates the power of Monte Carlo simulations to optimize the design of a general-purpose XRF spectrometer that uses orthogonal beams in order to take
advantage of polarization effects\textsuperscript{12).} The results extend, augment, and clarify earlier experimental works and their theoretical underpinnings and otherwise are in conformity with this body of knowledge. For example, it is confirmed that boron carbide is the preferred Barkla polarizer below about 25 keV, and aluminum oxide is preferred for energies above this range and below about 50 keV. Also demonstrated is that polarizers and specimens can be too thick, resulting in depolarization by multiple scattering. The use of filters between the source and polarizer is advantageous as expected. When implemented in hardware, detection limits for heavy metals in organic materials are at and below 1 µg/g.

Heckel employs a useful relationship for the Barkla scattering factor:

\[
R_{\text{Barkla}} = 0.5 \times (1 - \cos \delta) \times (\sigma / \mu) \times (1 - \exp(- \mu \rho d)) \quad \text{Eqn. 29}
\]

where

- \( \delta \): angular beam width of scattered radiation
- \( \sigma \): scattering cross section of the target material
- \( \mu \): mass attenuation coefficient of the target material
- \( \rho \): density of the target material
- \( d \): thickness of the target material

This equation can be used to compare the integral reflectivity \( R_{\text{Bragg}} \) for diffraction peaks. Compared to Bragg reflections with rocking curves \( \delta \leq 1^\circ \), Barkla beam widths in the range \( \delta \approx 10^\circ \) are used for adequate counting rates. The results to date conform with the view that Bragg reflections are best used to advantage at lower energies, less than around 5 keV.

**CONCLUSION**

From my personal perspective and the foregoing exposition, I have concluded that general-purpose EDXRF spectrometers should be designed with tri-axial, orthogonal geometry. The polarization effect can be exploited to reduce spectral background and improve detection limits. In those cases where secondary fluorescers or Bragg reflectors have advantage over Barkla scatterers, they can simply be substituted. Maximum flexibility to tailor analyses for a broad range of elements or for specific elements is thereby attained. Table 2 below summarizes the excitation methods available with the preferred geometry.

| Characteristic of Excitation Methods available with tri-axial, orthogonal geometry |
|---------------------------------------------------------------|------------------|-------------|-----------------|
| **Feature**                                                   | Secondary Fluorescers | Bragg-Polarized | Barkla-Polarized |
| Scatter Efficiency                                           | Moderate to high   | Moderate to high | Low             |
| Integrated reflectivity                                      | High              | Low to moderate | Moderate to High |
| Monochromaticity                                             | Yes, shifted by fluorescence | Yes, no change | No; Compton-shifted; Broad band at wide solid angles |
| Self-collimating                                             | No                | Yes, in one dimension | No             |
| Useful energy ranges                                         | Useful for most energies | Best for low energies due to availability of d-spacings for diffraction | Best for moderate to high energies due to absorption |
| Range of elements analyzed                                  | Narrow            | Narrow         | Broad           |

Table 2. Compare excitation methods.
REFERENCES

1. J.C Young, R.A.Vane, and J.P. Lenehan, “Background Reduction by Polarization in Energy Disperse X-ray Spectrometry,” Western Regional Meeting of the American Chemical Society, San Diego, California (October, 1973)