A new method of point-to-curve distance calculations

L. W. May, S. Wuenschel, B. Stein, and S. J. Yennello

One method of particle identification in heavy ion collisions is ΔE vs. E. These plots are often analyzed by linearization as described in [1]. This requires the determination of distance between any given data point and a curve of known form. A new method was developed for determining the exact distance between a point and a curve and is outlined below.

Three other methods of distance determination have been previously used: horizontal distance, vertical distance and distance along a 45-degree line. The challenge lies in that a method is needed that works over the whole range of a line of extreme curvature. The development of the new Point-to-Curve method is an attempt to reach a more accurate determination of distance globally over all regions. The horizontal and vertical methods simply take distances along paths parallel to the axes that intersect the curve, while the 45-degree method does the same, but tracks along a path that occurs at the angular bisection of the two axes.

There is a mathematical similarity between the 45-degree line and the Point-to- Curve methods. The method of using a 45-degree line to find the closest distance between a point and line is done by parameterization of the equation for the curve into a vector in terms of x and y values. For instance, a curve denoted by the equation $y = 3x^2 - 6x + 10$ would become the vector $r = (x, 3x^2 – 6x + 10)$. The value for the point in question, $p = (x1, y1)$, is then subtracted from the vector r giving $r – p = (x-x1, 3x^2 – 6x + 10 – y1)$. Now, since the slope of a 45-degree line is exactly 1 and slope is defined as the change in y divided by the change in x, then for a 45-degree line $\Delta y = \Delta x$ and $\Delta y – \Delta x = 0$. As $r – p$ is a representation of the change in y and change in x, then the x-value of $r – p$ is subtracted from the y-value of the vector and the resulting equation is set to zero. The resulting equation is solved for x and the values given should be solutions to the 45-degree line approximation. Since several possible roots can be given as solutions, constraints can be placed on the output such that only the value within the desired range will be found. The other solutions can be discarded because the equation solved is a general equation over the entire Euclidean space. The final x-value obtained is the x location of the intersection of the curve and the 45-degree line. This value can then be used to find the y-value of intersection and then a simple distance formula, $d^2 = (x1 – 2)^2 + (y1 – y2)^2$ can be used to determine the distance between the point and the curve along a 45-degree line.

Like the 45-degree line method, Point-to-Curve uses the parameterization of the curve and the value of the point to fix $r – p$. Since $r – p$ is a measure of the deviation in x and y between the point and the curve, this value can be squared by taking the dot product $(r – p) \cdot (r – p)$. This equation is minimized yielding points on the chosen curve, one of which will be the minimum for calculating the distance between the point and curve. The minimization is done by setting the derivative of the dot product equal to zero and solving for x. These values can be both real and imaginary, but as imaginary values give non-physical results, these values can be discarded. The same location constraints from the 45-degree line method can be applied to the solutions and the one remaining solution is the x-intersection of the curve and the line drawn from the point to the curve with the smallest possible distance. This x-intersection value can then be used to find the y-intersection value from the equation of the curve and the distance formula applied to the intersection values and the value of the point to give the actual distance between the point and the curve.

The computationally intensive step in both of these methods is in the factorization of the polynomial. Polynomial factorization has been demonstrated as an NP-complete problem, meaning that the time it takes to factor the polynomial completely scales exponentially with the size of the polynomial. This is not a problem for polynomials of order 4 or less, since these can be solved analytically using known algorithms. A problem could arise if high order functions are used to model the line. However, for the types of plots common for ΔE vs. E and CsI fast-slow data this is not prohibitive as the use of spline fits allows the order of the resulting polynomial to be kept low enough (no greater than 5 or 6) that the time to factor the polynomial is quite low. The factoring algorithm used in this case is that found in the GSL library packages linked in the ROOT [2] analysis package.

The Point-To-Curve method is more theoretically robust than the other three methods and should yield an exact analytical answer to the question of distances while the other methods give approximations. While the Point-To-Curve method could potentially use more computational time, all tests on actual data have shown that at the level the method is currently used, there is no noticeable time difference in the calculations.

[1] S. Wuenschel *et al.*, *Progress in Research*, Cyclotron Institute, Texas A&M University (2007-2008), p.II-28.

[2] Rene Brun and Fons Rademakers, ROOT - An Object Oriented Data Analysis Framework, Proceedings AIHENP'96 Workshop, Lausanne (September 1996); Nucl. Instrum. Methods Phys. Res. **A389**, 81 (1997). See also http://root.cern.ch/.