Distance of Sample Measurement Points to Prototype Catalog Curve

Poul G. Hjorth\textsuperscript{1} and Mirza Karamehmedović\textsuperscript{2}
Department of Mathematics, Building 303, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark.

John Perram\textsuperscript{3}
Maersk-McInney Institute for Production Technology, University of Southern Denmark, DK-5230 Odense, Denmark.

Robert Piché\textsuperscript{4}
Department of Mathematics, Tampere University of Technology, FIN-33101 Tampere, Finland.

Ulrik Ullum\textsuperscript{5}
Research and Technology, Grundfos Management A/S, DK-8850 Bjerringbro, Denmark.

Michael Krätzschmar\textsuperscript{6}
Flensburg University of Applied Sciences, D-24943 Flensburg, Germany.

Abstract.
We discuss strategies for comparing discrete data points to a catalog (reference) curve by means of the Euclidean distance from each point to the curve in a pump's head $H$ vs. flow $Q$ diagram. In particular we find that a method currently in use is inaccurate. We propose several alternatives that are accurate, fast, robust and stable.

Keywords: Geometry, orthogonal distance from point to curve, Schwarz-Christoffel transformation, cubic splines.
PACS: 02.30 Approximation Methods
MSC: 51.1 Geometry. 65.1 Numerical algorithms.

1. The Working Group
This problem was worked on by Jiří Fečkan, Poul Hjorth, Mirza Karamehmedović, Michael Krätzschmar, John Perram, Robert Piché, and Ulrik Ullum.

\textsuperscript{1} Corresponding author: p.g.hjorth@mat.dtu.dk
\textsuperscript{2} M.Karamehmedovic@mat.dtu.dk
\textsuperscript{3} jperram@mip.sdu.dk
\textsuperscript{4} robert.piche@tut.fi
\textsuperscript{5} ullum@ullum.dk
\textsuperscript{6} kraetzschmar@mathematik.fh-flensburg.de
2. The Problem
The company GRUNDFOS compares production samples of a pump to a prototype of the pump. For the prototype, a so called reference curve has been constructed in a \((Q, H)\)-diagram, by taking a large number of measurements and subsequently fitting a polynomial of degree at most 5 to these data points. The shape of the reference curve in the \((Q, H)\)-diagram reflects the classical Bernoulli relation \(H \propto Q^2\) where the pressure head is related to the square of the flux \(Q\). The design point is about half way between \(Q_{\text{min}}\) and \(Q_{\text{max}}\).

The comparison with production samples consists of placing measurement points from samples in the \((Q, H)\)-diagram along with the reference curve, and then for each sample and each measurement point to find the orthogonal distance from the measurement point to the reference curve. See figure 1.

![Figure 1](image1)

**Figure 1.** In the \((Q, H)\)-plane the distance between each single measurement point and the reference curve is to be evaluated.

The details of the method currently used is given in the next section. GRUNDFOS asks for a suggestion for an optimal distance-checking procedure which is sufficiently self contained, e.g., not relying on third-party packages or numerical libraries.

3. The Present Method
The method currently being employed at GRUNDFOS is in more details as follows. To begin with, the catalog curve is approximated by a piecewise linear curve. An algorithm (termed ‘Inverse Schwarz-Christoffel Transformation’, see however the next section) is being used to find the distance of the measurement points to the reference curve (Figure 2).

![Figure 2](image2)

**Figure 2.** The procedure currently used essentially places a grid near the point and the curve and evaluates the position of the point in the coordinates corresponding to the grid.

For each measurement point, a new coordinate system \((x', y')\) is introduced such that the \(y'\)-axis is the line through the point and perpendicular to the piecewise linear segment near the measurement point. The \(x'\)-axis is then the the line through this segment, and the distance from the point to the curve is
read off as the \( y' \)-coordinate of the point. The idea stems from the theory of conformal mappings where it forms a part of the so called Schwarz-Christoffel Transformation. GRUNDFOS accordingly terms their algorithm the "Inverse Schwarz-Christoffel Transformation" (ISC).

There is, however a nonuniqueness problem associated with this method (apart from the fact that it is somewhat computationally cumbersome.) Before suggesting alternative methods, we examine this problem in more detail.

4. The Schwarz-Christoffel Transformation

It is useful firstly to provide a brief recapitulation of the Schwarz-Christoffel transformation.\(^7\) Given a finite index set \( I = \{1, \ldots, n\}, n \in \mathbb{N} \), let \( \{x_j\}_{j \in I} \subset \mathbb{R} \) be a set of points on the real axis. Furthermore, introduce the set \( \{\theta_j\}_{j \in I} \subset \mathbb{R} \), as well as the constants \( A, B \in \mathbb{C} \). A Schwarz-Christoffel transformation is now defined by (see [1])

\[
w(z) = A \int_0^z \prod_{j=1}^n (\zeta - x_j)^{\theta_j} d\zeta + B, \quad z \in \mathbb{C}, \quad \Im z \geq 0
\]  

(1)

It can be shown [1] that the mapping in 1 takes the segments

\[
[-\infty, x_1[, \ldots, ]x_1, x_2[, \ldots, ]x_{n-1}, x_n[, ]x_n, \infty[
\]  

(2)

of the real axis into linear segments of a polygonal path in the complex plane such that the 'right turn angles' between the successive segments are specified by \( \theta_j, j \in I \). In the special case where it holds true that

\[
\lim_{z \to -\infty} w(z) = \lim_{z \to \infty} w(z)
\]  

(3)

the upper half-plane is mapped onto a polygon comprising the \( n + 1 \) vertices

\[
w(x_1), \ldots, w(x_n), \lim_{z \to -\infty} w(z)
\]  

(4)

In fact, given any positively oriented polygon with consecutive corners \( w_j \in \mathbb{C} \) and right-turn angles \( \theta_j \in \mathbb{R} \), there exists [1] an injective Schwarz-Christoffel mapping \( w \) from the upper half-plane onto that polygon such that \( w(x_j) = w_j, j \in I \) and \( \lim_{z \to -\infty} w(z) = w_{n+1} \).

Two major comments are in order regarding the use of the ISC mapping for measurement of distances of points from curves. Firstly, a Schwarz-Christoffel transformation generally maps a rectangular coordinate system in the complex plane to a skew coordinate system in that plane. This renders the mapping inadequate for measurements of orthogonal distances of points from curves, as will be illustrated below using a practical example. Secondly, it is clear from the expression 1 that the transformation (and hence also its inverse) is dependent upon the specific choice of the right-turn angles \( \theta_j, j \in I \). Expressed differently, every distinct shape of the given piecewise linear curve (catalog curve) will result, in general, in a distinct ISC mapping. This is expected to render the method sensitive to inaccuracies in catalog curves, and a numerical example given below serves to confirm this point.

A study of the algorithm employed by GRUNDFOS reveals that the company, in fact, is not using the inverse Schwarz-Christoffel transform, as asserted in the problem formulation. It is, however, still of interest to explore the possibility of an application of this method in the present context.

\(^7\) We direct to Reference [1] for a more thorough treatment.
5. Numerical Examples
In this section, numerical examples are provided which illustrate the inadequacy of the ISC transform in the context of determination of the shortest distance of a point to a catalog curve. For completeness, we note that the particular ISC map employed here is specified by the (convenient) choice of parameters reading \((x_1, x_2) = (-1, 0)\).

In the following, a part of a catalog curve is modeled using a piecewise linear approximation, shown in Figure 3a). In the same figure, two points are inscribed such that they have the same shortest distance to the piecewise linear curve. Figure 3b) shows the images of these two points under the ISC mapping. As is evident, these images are points in the complex plane with different shortest distance from the real axis. Thus, the ISC transform is not suitable for direct measurement of shortest distances of points from curves.

We next consider two distinct piecewise linear curves, as shown in Figure 4. The smallest angles between the linear segments of the curves are \(60^\circ\) and \(62^\circ\), respectively. The discrepancy between the curves introduced here may model the effect of inaccuracies in catalog curve representation (piecewise linearisation.) Alternatively, the two piecewise linear curves may simply be assumed to represent two different catalog curves.

The two graphs are shown again in Figures 5a) and 6a), respectively, along with sets of points chosen such that the shortest distances from corresponding points to the respective curves are the same.

Figures 5b) and 6b) show the images of these points under the ISC transform.\(^8\) An investigation of the presented results reveals that the shortest distances of corresponding image points to the real axis are not the same. Furthermore, we find that the ratio of the distances from the real axis of image points in case a) is not the same as the corresponding ratio of case b), the ratios being 0.484 and 0.486, respectively.

\(^8\) Note that, for convenience, the ordinates in figures 5b) and 6b) start at 0.8 instead of at 0.
Figure 5. The ISC mapping of two points for the first data set.

Figure 6. The ISC mapping of two points for the second data set.

Figure 7. Illustration of change of shortest distance ratio due to simple translation of measurement and reference data.

Hence, a slight change in the catalog curve data evidently results in a different evaluation of the shortest distances and their ratios. This is clearly a disadvantage of the method under investigation.

As a final example, consider Figure 7a), where the original piecewise linear curve (comprising the angle of 60°) is presented together with two pairs of points. One pair of points is obtained using a linear translation of the other one in such manner that both shortest distances to the curve are preserved. Figures 7b) and 7c) show the images of the two point pairs under the ISC transform. It is seen that the shortest distances to the curve, as well as the ratio of these distances, are altered by the simple translation procedure described above.

6. Proposed Alternative Algorithms

In light of the examples of the previous section we conclude that the ISC method used currently has some serious drawbacks; it is both inaccurate and inconsistent. We therefore proceed to suggest four alternatives: Discretization, the Golden Ratio Search, and Rootfinding (Brent’s Method and Sturm Sequence).

The proposed alternative methods are all minimisation algorithms. For a given continuous catalog curve $Q \mapsto (Q, H(Q))$ and a test point $P = (Q_P, H_P)$, the problem amounts to locating the minimum of
the squared distance function

\[ F(Q) = (Q - Q_p)^2 + (H(Q) - H_p)^2 \] (5)

Since \( F \) is continuous it attains its minimum somewhere in its domain \([0, Q_{\text{max}}]\). More than one point on the curve may however realise the minimum, though this does not generically occur in practice. However, the minimum distance value is unique, so the proposed methods do not suffer from the inconsistencies of the ISC method.

6.1. Discretization

This method uses a 'brute force' approach. From a table of \( N \) catalog curve points \( Q_1, \ldots, Q_N \), a table of \( N \) distance values \( F(Q_1), \ldots, F(Q_N) \) is generated; the computer then picks the smallest of these value. This method is simple, but is sufficiently fast, accurate and reliable for this problem.

Suppose the search locates \( F(Q_k) \) as the minimum value, and let \( F(Q^*) \) be the real minimum value of the function \( F \) (assumed twice differentiable). From Taylor’s theorem we then have

\[ |F(Q_k) - F(Q^*)| \leq \max |F''(Q)| \max |Q_i - Q^*|^2 \] (6)

so that error decreases quadratically with the density of curve points.

6.2. Golden Ratio Search

In this standard minimisation algorithm, a succession of minimum-bracketing \( Q \) intervals, starting with the domain endpoints \([a, b] = [0, Q_{\text{max}}]\), is generated as follows. From the two values \( a, b \) compute the intermediate values \( c = a + 0.382(b - a) \) and \( d = a + 0.618(b - a) \). If \( F(c) \leq F(d) \) then continue the search in \([a, d]\), otherwise continue the search in \([c, b]\). The locations of the intermediate points are such that one of the intermediate values can be reused in the next iteration, so once the method gets started only one evaluation of \( F \) per iteration is needed. This algorithm gives quadratic convergence to a local minimum of \( F \) when \( F \) is unimodal – this is typically the case for distance functions considered here.

6.3. Root-Finding

The minimum of \( F \) is achieved either at the end points of the domain \([0, Q_{\text{max}}]\) or at an interior point. In the latter case, we should locate the zeros of the derivative \( F' \). Because the catalog curve function \( H(Q) \) is a polynomial of degree \( n \leq 5 \), the squared distance function \( F \) given by Eqn (5) is a polynomial of degree \( 2n = 10 \), and the derivative is a polynomial of degree \( 2n - 1 \leq 9 \).

There are many methods available to find a single function zero. A popular choice is Brent’s method, also known as the van Wijngaarden-Deker-Brent method \([7]\). This algorithm combines root bracketing, interval bisection, and inverse quadratic interpolation, and does not require function derivatives. The algorithm is described in a number of textbooks, including \([5]\) and \([7]\).

A general root-finding algorithm may find a solution that corresponds to a local maximum or to a local minimum that is not a global minimum. To ensure against this, one could factor the computed zero out of the polynomial \( F' \) (“deflation”) and restart.

Alternatively, there are methods available to reliably find all the real zeros of a real polynomial. For example, reference \([2]\) gives an easily programmable algorithm that uses Sturm’s sequences \([6]\) to bracket the roots, and subsequently uses bisection to pinpoint the zeros.

7. Reference Curves and Tolerances

The problem brought to us by the company is a straightforward geometry question, but the underlying problem is more involved: how to evaluate how well a sample meets a tolerance, and how this tolerance is determined. We had too little background information to be able to do any detailed work on this question, but we have some general suggestions that could be the basis of further study.
7.1. Reference Curve Fitting

The reference curve is constructed based on measurements of a number of standard pumps. Orthogonal distance is used as a curve-fitting criterion. The public domain code ODRPACK[4], is available to compute polynomial reference curves based on orthogonal distance regression. This code could be useful in automating the reference curve construction.

Orthogonal distance regression is also known as error-in-variables modelling in statistical literature. Standard statistical regression tools are available to help determine the appropriate degree of the polynomial and to provide confidence regions above and below the reference curve. These confidence regions might provide a better way of evaluating sample pumps than the geometrically motivated tolerance curves currently used (see below).

7.2. Splines

Instead of fitting a single polynomial, one could construct a reference curve that is a piecewise cubic polynomial with continuous first and second derivatives, that is, a cubic spline. This would generally produce a better fit to measurement data while at the same time limiting the degree of the polynomial. (High degree polynomials are more likely to have undesirable oscillations and make it harder to compute the zeros of the derivative function $F'$ in the Root-Finding method.)

When using cubic splines the interval $[0, Q_{\text{max}}]$ is divided into $m$ (not necessarily equally spaced) subintervals $[Q_i, Q_{i+1}]$. The cubic spline restricted to any subinterval is a polynomial of degree 3. All together, these $m$ polynomials of degree 3 form a function that is a piecewise polynomial and furthermore twice continuously differentiable. Because the eye cannot distinguish between a twice continuously differentiable and an infinitely continuously differentiable function there is no aesthetic reason not to prefer a cubic spline instead of a polynomial as a catalog curve. ODRPACK could be used for automatic fitting of cubic splines to measurement data based on orthogonal distance regression.

Using cubic splines instead of a single polynomial curves would change the Root-Finding Method for finding the distance between a given point and the catalog curve in the following way. Instead of looking for all the zeros of a polynomial of degree $2n - 1$ in the interval $[0, Q_{\text{max}}]$, we find for each subinterval $[Q_i, Q_{i+1}]$ all the zeros of a 5th order polynomial belonging to this interval. Let $F$ denote the set of all the values of $F$ at these zeros together with the values of $F$ at node points $Q_1, \ldots, Q_m$. Then $F$ is finite with no more than $6m + 1$ elements and the continuous minimization problem can be again be replaced by the discrete minimization problem. If the cubic spline is a convex or concave function (which will typically be the case for a catalog curve) the number of points in $F$ will be close to $m$ instead of being close to $6m + 1$.

7.3. Tolerance curves

GRUNDFOS defines tolerance curves by adding $\pm 7\%$ to the $H$ values and $\pm 9\%$ to the $Q$ values, thereby generating 4 curves. There are two displaced set of $H$ values plotted against $Q$, and two sets of $Q$ values displaced from the catalog curve. See figure 8.

![Figure 8. The 'tolerance curves' obtained by horizontal and vertical displacement of the catalog curve.](image-url)
Since the tolerance of samples is measured in terms of perpendicular distance of a test point to the catalog curve, it is far from clear that the curves drawn in the above manner are consistent with that measure. To resolve this, we construct curves that are obtained from the catalog curve by displacing the curve in directions normal to the tangent at each point.

The amount displaced we chose to be equal in both directions and of magnitude the r.m.s. error

$$\frac{1}{100} \sqrt{(9Q)^2 + (7H(Q))^2}$$  \hspace{1cm} (7)

The normal direction vector is given by

$$\mathbf{n} = \frac{(-H'(Q), 1)}{\sqrt{1 + H'(Q)^2}}$$  \hspace{1cm} (8)

Thus, for the point \((Q, H(Q))\) on the curve, the corresponding curve on the upper and lower tolerance curve have the parametric expressions:

\[(Q, H(Q)) \pm \frac{1}{100} \sqrt{(9Q)^2 + (7H(Q))^2} \frac{(-H'(Q), 1)}{\sqrt{1 + H'(Q)^2}}\]  \hspace{1cm} (9)

The resulting curve (for a typical shape of the catalog curve) is shown superimposed on the GRUNDFOS curves in figure 9.

![Figure 9](image)

**Figure 9.** The tolerance curves obtained by displacement of the catalog curve along the direction in which tolerance is measured. Note that the curve is different from the envelope of the curves of figure 8.

Notice that the upper curve generated in this way (which is consistent with the calculation of the sample distance values) lies above the horizontal-vertical shifted curves. Similarly, the lower curve lies below the horizontal-vertical shifted curves. The differences are quite significant near the duty point, indicating that the tolerance curves used up to now have been too stringent.

**8. Conclusion**

In this report we have first given reasons GRUNDFOS should change the procedure for comparison of sample measurements with the catalog curve; indeed the procedure currently used may give misrepresenting and misleading results. We have then proposed four separate algorithms to solve the problem.

Of these four algorithms, we recommend *Discretization* for its simplicity, speed and robustness. The error is easily controllable through the number of points in the discretization.

---

[1] E. B. Saff and A. D. Snider, 1993 *Fundamentals of Complex Analysis for Mathematics, Science and Engineering*, 2nd edition, Prentice Hall

[2] See, e.g., [http://www.acm.org/pubs/tog/GraphicsGems/gems/Sturm/main.c](http://www.acm.org/pubs/tog/GraphicsGems/gems/Sturm/main.c)

[3] I.J. Anderson, M.G. Cox, A.B. Forbes, J.C. Mason and D.A. Turner 1998 *An Efficient and Robust Algorithm for Solving the Foot Point Problem*, in Mathematical Methods for Curves and Surfaces II, Dahlen et al (eds.) Vanderbilt Press, ISBN 0-8265-1315-8.
[4] Paul T. Boggs and Janet R. Donaldson, *ALGORITHM 676, ODRPACK: Software for Weighted Orthogonal Distance Regression*, ACM Transactions on Mathematical Software.

[5] Brent, R.P. 1973 Ch.3-4 in *Algorithms for Minimization Without Derivatives*, Englewood Cliffs, NJ: Prentice-Hall

[6] Forsythe, G.E.; Malcolm, M.A.; and Moler, C.B. 1977 7.2 in *Computer Methods for Mathematical Computations*. Englewood Cliffs, NJ: Prentice-Hall

[7] Press, W.H.; Flannery, B.P.; Teukolsky, S.A.; and Vetterling, W.T. 1992 *Numerical Recipes in FORTRAN: The Art of Scientific Computing*, 2nd ed. Cambridge, England: Cambridge University Press