# A Rational Approach to B-H Curve Representation

Patrick Diez
Department of Electrical and Computer Engineering
McGill University, Montréal, Canada

December 2015

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of

Master of Engineering

# Abstract

This thesis introduces a method for constructing a rational function that fits the finite set of data points on the B-H plane representing the non-linear response of a ferromagnetic, permeable material. A background of the nature of ferromagnetism is provided, focusing on the mathematical constraints on a function $B(H)$ representing a physically valid magnetic material response curve. Current approaches to B-H curve representation are categorized, reviewed, and evaluated for their physical validity and accuracy of approximation. The method of vector fitting (as seen in the discipline of control systems) is applied to develop the theory of rational curve fitting for B-H data sets; techniques to establish physical validity, and corrections for specific errors that may arise from the curve fitting process, are presented. The resulting rational fitting method is demonstrated to yield valid curves for 192 B-H data sets, with an average root-mean-squared error of 9.49 mT. It is also applied in the finite element method solution of the TEAM 13 problem, yielding results in good agreement with those determined experimentally. Further applications of rational functions to magnetic material modelling, such as the construction of frequency-dependent material response models and hysteresis loops, are proposed for future investigation.

# Résumé

Cette thèse présente une méthode pour construire une fonction rationnelle qui correspond à l'ensemble fini de points de données sur le plan B-H représentant la réponse non-linéaire d'un matériau perméable ferromagnétique. Une analyse de la nature du ferromagnétisme est fournie, en se concentrant sur les contraintes mathématiques sur une fonction $B(H)$ représentant une courbe de réponse physiquement valide des matériaux magnétiques. Les approches actuelles pour représenter les courbes B-H sont classées, examinées et évaluées en fonction de leur validité physique et leur précision d'approximation. La méthode de *vector fitting* (comme on le voit dans la discipline des systèmes de contrôle) est appliquée au développement de la théorie d'ajustement de courbe rationelle pour les ensembles de données B-H. Des techniques pour établir la validité physique, et les corrections d'erreurs spécifiques qui peuvent surgir lors du processus d'ajustement de courbe, sont présentées. On démontre que la méthode d'ajustement de courbe rationnelle qui en résulte produit des courbes valides pour 192 ensembles de données B-H, avec une moyenne erreur quadratique moyenne de 9,49 mT. On applique aussi la méthode dans la solution par méthode des éléments finis du problème TEAM 13, ce qui donne des résultats en bon accord avec ceux déterminés expérimentalement. D'autres applications des fonctions rationnelles dans la modélisation des matériaux magnétiques, telles que la construction de modèles de réponse dépendante de la fréquence, et la construction de boucles d'hystérésis, sont proposées pour des recherches futures.

# Acknowledgements

This thesis was made possible with the support of Infolytica Corporation, which provided its software, MagNet, for use in validating the results of the method presented herein; Derek Dyck, whose knowledge of the domain aided in the procurement of background literature; and Craig Brett, whose work at Infolytica enabled the use of rational B-H curves in MagNet. Special thanks go to Prof. Jon Webb, whose patience and support were invaluable in the completion of this work.

# Contents

**1 Introduction** **6**

**2 Background** **6**

**3 Literature Review** **9**

3.1 Smooth Approaches . . . . . 10

3.2 Piecewise Approaches . . . . . 16

**4 Theory** **19**

4.1 Iterative Linear Least-Squares . . . . . 21

4.2 Basis Functions . . . . . 27

4.3 Non-Linear Least-Squares . . . . . 29

4.4 Curve Validation and Correction . . . . . 30

**5 Implementation Details** **32**

5.1 Linear Systems Solving . . . . . 32

5.2 Polynomial Evaluation and Deflation . . . . . 35

5.2.1 Horner's Scheme . . . . . 35

5.2.2 De Casteljau's Algorithm . . . . . 37

5.3 Polynomial Factorization . . . . . 39

5.4 Partial Fractions Decomposition . . . . . 40

**6 Results** **43**

6.1 Curve Approximation . . . . . 43

6.2 Finite Element Method . . . . . 45

**7 Conclusions** **48**

**8 Appendix** **49**

# 1 Introduction

Since the advent of modern computing, it has been of interest to model the magnetic flux density ($B$) response of magnetically permeable materials subject to an applied magnetic field ($H$). Such models are regularly applied in finite element methods, which in turn discretize and approximate solutions to Maxwell's equations for the purpose of simulating electromagnetic devices. In the case of ferromagnetic materials, the aforementioned response is non-linear: the ratio of magnetic flux density to applied magnetic field – that is, $\mu = B/H$, the *permeability* – varies as a function of the magnitude of the applied field, making modelling non-trivial.

This thesis presents a new method of constructing a continuously differentiable function – notably, a rational function – that maps magnetic field intensity onto magnetic flux density, approximating an input finite data set of B-H points for a given permeable material. Section 2 provides an overview of the physical nature of ferromagnetic materials, and mathematical descriptions of this nature. In section 3, a review of current methods, classified into two categories, is provided. Section 4 presents the method of vector fitting as developed in [1], and its application to the problem of rational function approximation of B-H data sets, with appropriate corrections made to enforce the necessary physical properties of a B-H curve when possible. The next section (5) provides implementation details of the algorithm, notably concerning numerical round-off and stability issues. Finally, the approach is applied to a collection of B-H data sets, as well as the TEAM 13 finite element problem in section 6.

# 2 Background

It is known that the ferromagnetic material's response consists of a series of jumps in magnetic flux density due to the Barkhausen effect [2]. As entire magnetic domains respond to the applied field by sudden changes in their size and orientation, the magnetic flux density in a ferromagnetic material is, strictly speaking, a piecewise constant function of the applied magnetic field. However, this physical nature does not provide a convenient basis for incorporating permeable material models into Maxwell's equations, as the latter ignore the quantized, discrete nature of materials and treat permeability and magnetic flux density as continuously differentiable functions of the magnetic field. To resolve this, it suffices to consider $B$ as a function of $H$ in the limit as the domains of the ferromagnetic material to be modelled become infinitesimal in size; equivalently, in the limit as the Barkhausen jumps become infinitesimal in their magnitude in $B$ and in their spacing in $H$.

Furthermore, ferromagnetic materials exhibit hysteresis: the property that the material's

magnetic flux density response to an applied magnetic field is stateful, and thus $B$ is not simply a single-valued function of $H$. A detailed discussion of hysteresis is beyond the scope of this document, however, it is necessary to identify several common curves and specify what is meant by 'B-H curve' when discussing B-H curve representation.

- *Initial magnetization curve*: The curve starting at the point $B = H = 0$, consisting of the points on the B-H plane passed through by those fields as $H$ is monotonically increased from $H = 0$

- *Main hysteresis loop*: The closed loop on the B-H plane resulting from the (repeated) monotonic decrease in applied magnetic field from a point $(H_{\text{peak}}, B_{\text{peak}})$ to a point $(-H_{\text{peak}}, -B_{\text{peak}})$ followed by the monotonic increase in applied magnetic field back to the point $(H_{\text{peak}}, B_{\text{peak}})$. This loop can be described by two single-valued monotonic functions $B_-(H)$ and $B_+(H)$, where $B_-(H)$ is the branch of the loop resulting from the monotonic decrease of $H$, and $B_+(H)$, the branch resulting from its monotonic increase. The branches are related by the equation $B_-(H) = -B_+(-H) \; \forall \; H$ , and have the additional property that $B_-(H) > B_+(H)$ for all $-H_{\text{peak}} < H < H_{\text{peak}}$ – that is, the two functions only meet at the points $(H_{\text{peak}}, B_{\text{peak}})$ and $(-H_{\text{peak}}, -B_{\text{peak}})$. Furthermore, both branches have a slope $dB/dH$ no less than $\mu_0$, the permeability of free space. A material has many main hysteresis loops: one for each value of $H_{\text{peak}}$.

- *Anhysteretic curve*: The curve resulting from taking the average value of the two branches of a main hysteresis loop, that is: $B_{\text{anhyst}}(H) = \left(B_+(H) + B_-(H)\right)/2$. Unlike main hysteresis loops, anhysteretic curves are single-valued functions of $H$; like main hysteresis loops, a distinct anhysteretic curve exists for each value of $H_{\text{peak}}$.

- *Commutation curve*: The curve consisting of the locus of points $(H_{\text{peak}}, B_{\text{peak}})$ for all main hysteresis loops. Experimentally, points on the commutation curve are determined by applying an oscillating magnetic field (usually sinusoidal) with different amplitudes $H_{\text{peak}}$ and measuring the corresponding maximum magnetic flux densities ($B_{\text{peak}}$).

Main hysteresis loops and anhysteretic curves are inconvenient when attempting to characterize a material, as a single material can only be defined by a (strictly speaking, infinite) set of such curves. In seeking a *simple* non-linear ferromagnetic material model, these are immediately discarded.

In [3], Zirka *et al.* list five "experimentally established magnetization rules" defining aspects of the hysteretic nature of ferromagnetic materials. Significantly, they establish that

hysteresis loops that start at the initial magnetization curve return to the same point on the initial magnetization curve from which they originate. Consequently, this suggests that the initial magnetization curve and the commutation curve are, in fact, one and the same. Irrespective of this, given that many common electromagnetic devices (e.g. motors, transformers) operate with unbiased, alternating magnetic fields, commutation curves are the *de facto* 'B-H curve,' and although many of the processes presented in this document may be used to model the branches of hysteresis loops and anhysteretic curves, it is to be assumed henceforth that the term 'B-H curve' will refer to a commutation curve unless otherwise specified.

Ferromagnetic materials exhibit a Rayleigh region in the vicinity of $B = H = 0$ – that is, a region which obeys the Rayleigh law:

$$B(H) = \mu_{\text{init}} H + \alpha_R \left(\mu_0 H\right)^2 \tag{1}$$

where $\mu_{\text{init}}$ is the *initial permeability*, and $\alpha_R$ the *Rayleigh constant*. Additionally, they approach the line $B(H) \sim \mu_0 \left(M_{\text{sat}} + H\right)$ as $H \to \infty$, where $M_{\text{sat}}$ is the *saturation magnetization*. In [4], Pechstein and Jüttler summarize the properties of a function $f(s)$ mapping applied magnetic field ($H$) to magnetic flux density ($B$) necessary for $f(s)$ to represent a physically valid B-H curve:

(A1) $f$ is continuously differentiable on $\mathbb{R}_{\geq 0}$,

(A2) $f(0) = 0$,

(A3) $f'(s) \geq \mu_0, \ \forall\, s \geq 0$,

(A4) $\lim_{s\to\infty} f'(s) = \mu_0$, where $\mu_0$ denotes the permeability in vacuum.

Hülsmann in [5] presents an equivalent set of properties for a function $H(B)$ and a corresponding function $\nu(B^2) = H(B)/B$, omitting a counterpart for Pechstein and Jüttler's (A4):

1. $H(0) = 0$ (no remanence),
2. $H(B) > 0 \ \forall\, B > 0$ and $\nu(B^2) > 0 \ \forall\, B > 0$ (positivity),
3. $\frac{d}{dB} H > 0$ (monotonicity).

Note that it may appear that Hülsmann's first and third properties imply the second (that is, that a monotonically increasing function passing through the origin is positive for all positive inputs), however this is only the case for a continuous function (a condition which Hülsmann does not stipulate).

For the purposes of this document, the following are necessary conditions for a function $B(H)$ to be considered physically valid:

(i) Continuous differentiability: $B(H)$, and its first derivative $\frac{dB(H)}{dH}$, are defined for all $H \geq 0$. Ideally, $B(H)$ should be smooth (i.e. infinitely differentiable), however requiring this would disqualify most piecewise-defined functions. Continuous differentiability (as opposed to continuity alone) is highly desirable in finite element solvers using the Newton-Raphson method, requiring the evaluation of the derivatives of $B(H)$.

(ii) Zero coercivity: $B(0) = 0$, that is, the function $B(H)$ passes through the origin of the B-H plane. For non-linear permanent magnets, the coercivity is ignored when discussing their B-H curves as it is easily reintroduced after construction of the function $B(H)$ by the evaluation of $B(H - H_c)$, where $H_c$ is the coercivity.

(iii) Monotonicity: $\frac{dB(H)}{dH} \geq \mu_0$ for all $H \geq 0$, that is, $B(H) - \mu_0 H$ is a monotonically increasing function.

(iv) Positivity of magnetization: $B(H) \geq \mu_0 H$ for all $H \geq 0$, that is, the magnetization of the material, $M = \nu_0 B(H) - H$ (where $\nu_0 = \mu_0^{-1}$ is the *reluctivity* of free space), is positive over the entire B-H curve.

(v) Finite, positive saturation magnetization: $\lim_{H\to\infty} (B(H) - \mu_0 H) = \mu_0 M_{\text{sat}}$, where $M_{\text{sat}}$ is a finite, positive value – the saturation magnetization of the material. This implies that $\lim_{H\to\infty} \frac{dB(H)}{dH} = \mu_0$ (condition (A4) of Pechstein and Jüttler), but the latter is insufficient to establish finite $M_{\text{sat}}$, as demonstrated by the counter-example $B(H) = \sqrt{H+1} - 1 + \mu_0 H$, which has a differential permeability ($\frac{dB(H)}{dH}$) of $\mu_0$ as $H \to \infty$, but infinite magnetization.

These five conditions are equally applicable to functions mapping $B$ to $H$.

# 3 Literature Review

Current techniques for representing B-H data sets with continuously differentiable functions can be classified into two categories: those using functions in a small, fixed number of parameters that are smooth over $\mathbb{R}_{\geq 0}$, and those using functions in a (comparatively) large number of parameters that are piecewise-defined over a finite interval of $\mathbb{R}_{\geq 0}$. Examples of functions from the first category can be summarized as 'smooth approaches' and are found in [6–10], whereas examples of functions from the second category – 'piecewise approaches' –

are found in [4, 5, 11, 12], which include means of interpolating as well as approximating the material data set.

## 3.1 Smooth Approaches

An early example of B-H curve representation for use in computational applications is provided by Trutt *et al.* in [6], from 1968, although Fischer *et al.* provide an excellent overview of simple algebraic and transcendental functions capable of representing B-H curves in [7], from 1956. The paper [6] sets out to approximate a material data set by various smooth functions, and prior to doing so establishes the following set of conditions for the method of approximation:

1. If possible, a single function should represent the whole range from the origin to saturation.

2. The approximation should be simple mathematically so that it can be applied by a programmer as a matter of routine and not involve any judgment on the part of the operator.

3. The errors should be as small as possible.

4. The computer time used should be at a minimum.

It is understood from the context in [6] that conditions 2-4 apply to the method used to construct a B-H curve representation from a data set, and not to the evaluation of the resulting approximation. Noteworthy is that, due to its time of writing, Trutt *et al.* claim that "the Gaussian method minimizing the square of the error would be applicable. It has been found, however, that this is very time consuming," implying that least-squares error minimization does not satisfy condition 4. It is arguable that the computational advances of the past four decades have rendered this argument invalid, and that, furthermore, condition 4 can be by-and-large ignored.

Trutt *et al.* proceed to present several different smooth representations, categorized into four categories: power series, hyperbolas, transcendental functions, and Fourier series. Of the power series representations, the first presented is a function of the form

$$B = aH^n \tag{2}$$

where $a$ and $n$ are constants. The second is given by the function

$$aH = bB + (bB)^{2n+1} \tag{3}$$

where $n$ is a positive integer, and $a$ and $b$ are real constants, and it is suggested that $n = 1$, making the function a "cubic parabola." The third, most general power series representation provided in [6] is a polynomial in four terms:

$$H = a_0 + a_1 B + a_n B^n + a_m B^m \tag{4}$$

which the authors say is "very adaptive." Here, $n$ and $m$ are generally integers, making $H$ a polynomial in $B$.

Noteworthy of all of the power series representations of [6] is that they do not provide for a correct representation of the B-H curve in its extrapolation to the limit as $B$ and $H$ tend to infinity. Specifically, it is expected that $dB/dH$ approaches $\mu_0$ as $H$ tends to infinity; likewise, that $dH/dB$ approaches $\nu_0$ as $B$ tends to infinity. Yet for (2),

$$\lim_{H\to\infty} \frac{d}{dH} aH^n = \lim_{H\to\infty} anH^{n-1} = \begin{cases} 0, & \text{if } n < 1 \\ a, & \text{if } n = 1 \\ \infty, & \text{if } n > 1, \end{cases} \tag{5}$$

and thus (2) can only represent a physically valid B-H curve in the limit as $H$ tends to infinity if $a = \mu_0$ and $n = 1$, implying that the material has no magnetization. Similarly, for (3),

$$\lim_{B\to\infty} \frac{d}{dB} \frac{1}{a} \left( bB + (bB)^{2n+1} \right) = \lim_{B\to\infty} \frac{b}{a} \left( 1 + (2n+1)\,(bB)^{2n} \right) = \begin{cases} \frac{b}{a}, & \text{if } n < 0 \\ \frac{2b}{a}, & \text{if } n = 0 \\ \infty, & \text{if } n > 0. \end{cases} . \tag{6}$$

Here, the case for which $n < 0$ seems admissible, specifically, if $2n > -1$ (as this prevents a pole at the origin) and $b = \nu_0 a$. However, under these conditions, (3) fails to represent a physically valid material as:

$$\lim_{B\to\infty} M = \lim_{B\to\infty} (\nu_0 B - H) = \lim_{B\to\infty} -\frac{1}{a} (bB)^{2n+1} = -\infty \tag{7}$$

where $M$ is the magnetization of the material.

That is, assuming $a > 0$ and $b > 0$, the magnetization of the material is infinite and negative in the limit as $B$ tends to infinity. Moreover, any finite power series truncation can be shown to suffer a similar flaw, in that magnetization will requisitely either be zero or infinite in the limit as $B$ or $H$ tends to infinity. These representations *do* have the flexibility to capture material properties from the Rayleigh region (i.e. for small $H$) up to some finite value of $H$ (at which point the material is deemed saturated).

The second category presented by Trutt *et al.* is that consisting of hyperbolas, notably functions of the form

$$B = \frac{H}{a + bH} \tag{8}$$

commonly known as Froelich's equation, as well as functions of the form

$$B = a_0 + a_1 H - \frac{a_2}{H} \tag{9}$$

for which it is noted in [6] that this representation is invalid in the region of $H = 0$.

As with the power series representations, Froelich's equation (8) does not represent a physically valid B-H curve, as

$$\lim_{H\to\infty} M = \lim_{H\to\infty} \left( \frac{\nu_0 H}{a + bH} - H \right) = \frac{\nu_0}{b} - \lim_{H\to\infty} H = -\infty, \tag{10}$$

although this could be corrected by adding $\mu_0 H$ to (8):

$$B = \frac{H}{a + bH} + \mu_0 H, \tag{11}$$

in which case $\lim_{H\to\infty} M = \nu_0 b^{-1}$ and $\lim_{H\to\infty} dB/dH = \mu_0$. Similarly, (9) does allow for a finite, non-zero magnetization in the limit as $H$ tends to infinity, if $a_1 = \mu_0$, where $a_0$ is then the saturation magnetization. These hyperbolic representations therefore have the advantage over power series representations that they remain representative of reality for arbitrarily large $B$ and $H$, but they are incapable of representing B-H curves' Rayleigh regions - for (9) this is evident; for (8) and (11), this is due to the fact that the second derivative has no roots for $H \geq 0$:

$$\frac{d^2}{dH^2} \frac{H}{a + bH} + \mu_0 H = \frac{d^2}{dH^2} \frac{H}{a + bH} = \frac{-2ab}{(a + bH)^3}. \tag{12}$$

Consequently, the representation has no inflection points in the interval $0 \leq H < \infty$ and in order to have finite magnetization it must be concave (down) .

The third category consists of transcendental functions, the first two of which are $H = ae^{bB}$ and $B = a\left(1 - e^{-bH}\right)$, both of which are noted to only be useful "in small intervals of $B$." The function

$$B = e^{H/(a+bH)} \tag{13}$$

is presented for use "over a wide range" in [6], and

$$bH = \sinh\left(\frac{B}{a}\right) \tag{14}$$

is claimed to be a common choice, however like (8), the first does not provide for finite magnetization as $H \to \infty$; the second necessarily has a derivative that tends to infinity as $B$ tends to infinity. Trutt *et al.* provide a good transcendental representation in the form

$$B = a\tan^{-1}(bH) + cH \tag{15}$$

for which the choice $c = \mu_0$ is natural (although not specified in [6]) to make the B-H curve physically valid for all $H$.

The last representation type presented in [6] is that of finite Fourier series. Obviously, this approach can only approximate the B-H curve over a finite interval of values, but has sufficient degrees of freedom to capture details of the curve (such as its Rayleigh region) accurately.

Trutt *et al.* conclude that "a good fit can very seldom be obtained by a single function over the whole useful range of the magnetization characteristic," and proceed to provide a method for segmenting the B-H curve - fundamentally, a piecewise approach.

A well-known representation of $H$ as a function of $B$ is provided by Brauer in [8]:

$$H = \left(k_1 e^{k_2 B^2} + k_3\right) B. \tag{16}$$

Here, a choice of $k_1 > 0,\ k_2 < 0,\ k_3 = \nu_0$ yields a physically valid B-H curve for all values of $B \geq 0$. Brauer, with the intent of solving non-linear magnetic vector potential problems, illustrates the simplicity of the calculation of the reluctivity, $\nu$ and its derivative using this approach:

$$\nu = k_1 e^{k_2 B^2} + k_3 \tag{17}$$

$$\frac{d\nu}{d\left(B^2\right)} = k_1 k_2 e^{k_2 B^2}. \tag{18}$$

This representation is applied to three steel data sets in [8] with negligible error, however none has data points in its Rayleigh region. In fact, from (18), it is concluded that the Brauer model has a non-positive derivative with no roots, and thus cannot capture a material's response in its Rayleigh region.

Hülsmann improves on the Brauer model in [5] with the *Extended Brauer Model*, using the Brauer model as one segment of a piecewise, continuously differentiable function approximating the B-H data set. Although piecewise, this approach consists of a representation in only three segments: an initial, parabolic segment corresponding to the Rayleigh region; a second segment using the Brauer model for the region encompassing the knee (the region of the B-H curve in which the material's differential permeability decreases rapidly with respect to $H$ as it approaches saturation); and a third, linear segment for the saturation region. In [5] the model is presented for $H$ as a function of $B$, notably:

$$H(B) = \begin{cases} \sqrt{\frac{1}{4\alpha^2\nu_{\text{init}}^2} + \frac{B}{\alpha}} - \frac{1}{2\alpha\nu_{\text{init}}} & \text{if } 0 \leq B < B_{\text{exp}} \\ \left(k_1 e^{k_2 B^2} + k_3\right) B & \text{if } B_{\text{exp}} \leq B < B_{\text{lin}} \\ H_{\text{lin}} + \nu_0 \left(B - B_{\text{lin}}\right) & \text{if } B_{\text{lin}} \leq B \end{cases} \tag{19}$$

where $\nu_{\text{init}}$ and $\alpha$ correspond to the linear and quadratic coefficients in Rayleigh's law in the form:

$$B(H) = \nu_{\text{init}}^{-1} H + \alpha H^2; \tag{20}$$

$k_1$, $k_2$, and $k_3$ are the same parameters as those in (16); and $\left(B_{\text{exp}}, H_{\text{exp}}\right)$ and $\left(B_{\text{lin}}, H_{\text{lin}}\right)$ delimit the piecewise segments. Clearly, this model does not implicitly satisfy continuity, much less differentiability at the points $B = B_{\text{exp}}$ and $B = B_{\text{lin}}$. Hülsmann modifies (19) to enforce continuity implicitly by constraining $k_3 = \nu_{d,\text{exp}} - k_1$ and shifting the Brauer segment by $B_{\text{exp}}$ on the $B$ axis, and by $H_{\text{exp}}$ on the $H$ axis:

$$H(B) = \begin{cases} \sqrt{\frac{1}{4\alpha^2\nu_{\text{init}}^2} + \frac{B}{\alpha}} - \frac{1}{2\alpha\nu_{\text{init}}} & \text{if } 0 \leq B < B_{\text{exp}} \\ \left(k_1 e^{k_2\left(B - B_{\text{exp}}\right)^2} + \nu_{d,\text{exp}} - k_1\right)\left(B - B_{\text{exp}}\right) + H_{\text{exp}} & \text{if } B_{\text{exp}} \leq B < B_{\text{lin}} \\ H_{\text{lin}} + \nu_0 \left(B - B_{\text{lin}}\right) & \text{if } B_{\text{lin}} \leq B \end{cases} \tag{21}$$

where $\nu_{d,\text{exp}}$ denotes the differential reluctivity, $\frac{dH}{dB}$ of the parabolic segment at $B = B_{\text{exp}}$. Equation (21) consists of 9 parameters: $\alpha$, $k_1$, $k_2$, $\nu_{d,\text{exp}}$, $\nu_{\text{init}}$, $B_{\text{exp}}$, $H_{\text{exp}}$, $B_{\text{lin}}$, and $H_{\text{lin}}$; however, as Hülsmann notes, $H_{\text{exp}}$ is determined by the evaluation of the first segment at $B = B_{\text{exp}}$, similarly $H_{\text{lin}}$ by the evaluation of the second segment at $B = B_{\text{lin}}$. Two additional constraints establish differentiability at the segment boundaries, leaving 5 degrees of freedom. In [5], $\nu_{\text{init}}$ and the points $\left(B_{\text{exp}}, H_{\text{exp}}\right)$ and $\left(B_{\text{lin}}, H_{\text{lin}}\right)$ are chosen as independent variables, with the remaining parameters defined ($k_2$, implicitly) as functions of these:

$$\begin{aligned}
\alpha &= \alpha\left(\nu_{\text{init}}, B_{\text{exp}}, H_{\text{exp}}\right) \\
&= \frac{\frac{B_{\text{exp}}}{H_{\text{exp}}} - \frac{1}{\nu_{\text{init}}}}{H_{\text{exp}}} \\
\nu_{d,\text{exp}} &= \nu_{d,\text{exp}}\left(\nu_{\text{init}}, B_{\text{exp}}, \alpha\right) \\
&= \frac{1}{2\alpha\sqrt{\frac{1}{4\alpha^2\nu_{\text{init}}^2} + \frac{B_{\text{exp}}}{\alpha}}} \\
k_2 &= k_2\left(\nu_{d,\text{exp}}, B_{\text{exp}}, B_{\text{lin}}, H_{\text{exp}}, H_{\text{lin}}\right) \\
k_1 &= k_1\left(\nu_{d,\text{exp}}, B_{\text{exp}}, B_{\text{lin}}, H_{\text{exp}}, H_{\text{lin}}, k_2\right) \\
&= \frac{\frac{H_{\text{lin}} - H_{\text{exp}}}{B_{\text{lin}} - B_{\text{exp}}} - \nu_{d,\text{exp}}}{e^{k_2\left(B_{\text{lin}} - B_{\text{exp}}\right)^2} - 1}
\end{aligned} \tag{22}$$

where $k_2\left(\nu_{d,\text{exp}}, B_{\text{exp}}, B_{\text{lin}}, H_{\text{exp}}, H_{\text{lin}}\right)$ is the function representing the value of $k_2$ satisfying the equation $\frac{dH}{dB}\big|_{B=B_{\text{lin}}} = \nu_0$, or as derived in [5]:

$$\frac{\frac{H_{\text{lin}} - H_{\text{exp}}}{B_{\text{lin}} - B_{\text{exp}}} - \nu_{d,\text{exp}}}{\nu_0 - \nu_{d,\text{exp}}} = \frac{e^{k_2\left(B_{\text{lin}} - B_{\text{exp}}\right)^2} - 1}{\left(2k_2\left(B_{\text{lin}} - B_{\text{exp}}\right)^2 + 1\right)e^{k_2\left(B_{\text{lin}} - B_{\text{exp}}\right)^2} - 1}. \tag{23}$$

Hülsmann provides algorithms for generating an initial guess for, and solving, the non-linear least-squares problem, that is:

$$\begin{aligned}
&\arg\min_{\mathbf{x}} \left(H(\mathbf{x}, B_1)^2 + \cdots + H(\mathbf{x}, B_n)^2\right) \\
&\mathbf{x} = \left[\nu_{\text{init}}, B_{\text{exp}}, H_{\text{exp}}, B_{\text{lin}}, H_{\text{lin}}\right]
\end{aligned} \tag{24}$$

for $n$ data points, the details of which are found in [5] and are beyond the scope of this review. This approach produces B-H curve representations which are physically valid and continuously

differentiable for all $B \geq 0$. However the magnetization represented by the curve is constant for all $B \geq B_{\text{lin}}$ whereas it is expected that magnetization should only asymptotically approach a constant as $B \to \infty$. Hülsmann provides a test material for the model in [5], but the norm used for measuring the quality of the fit is specific to the Extended Brauer Model. Using the data provided, the RMS error of $H$ as a function of $B$ is 46.73 A / m, and inverting the representation and solving for $H$ for computation of an RMS error of $B$ as a function of $H$ yields 31.52 mT.

Rivas, Martín, and Pereira provide a straightforward approach to second-order rational function B-H curve representation of B-H curves in [9]:

$$M = \frac{a_0 + a_1 H + a_2 H^2}{1 + b_1 H + b_2 H^2}. \tag{25}$$

Rivas *et al.* provide expressions for the coefficients in (25) given the initial magnetic susceptibility ($\chi$), the Rayleigh material constant ($\lambda$), the magnetization of saturation ($M_s$), and the Néel constant ($\alpha$) of a material:

$$\begin{aligned} a_0 &= 0 \\ a_1 &= \chi \\ a_2 &= \frac{\lambda M_s + \chi^2}{M_s + \alpha\chi} \\ b_1 &= \frac{\alpha\lambda + \chi}{M_s - \alpha\chi} \\ b_2 &= \frac{\lambda M_s + \chi^2}{M_s (M_s - \alpha\chi)}. \end{aligned} \tag{26}$$

Unlike previous smooth approaches, that of [9] captures the response of materials in their Rayleigh regions and has physically valid behaviour as $H \to \infty$. Rivas *et al.* proceed to apply second-order rational functions to hysteresis loop modelling, with further improvements to this aspect of the work provided by Pagnola *et al.* in [10]. However, knowledge of the Rayleigh and Néel constants of a material is a requirement for the use of this approach, rendering it unusable when only a B-H data set is available.

## 3.2 Piecewise Approaches

As computing and data storage capacity increased throughout the 20$^{\text{th}}$ century, piecewise representations of B-H curves, requiring the calculation and storage of more parameter values,

became feasible and practical for finite element applications. An early example of such an approach is provided by Forghani *et al.*, which as presented in [11] defines the B-H curve in terms of a piecewise-cubic function of $(B-B_R)^2$ in 14 parameters (where $B_R$ is the remanence of the material): the $B^2$ axis is partitioned into 6 equal-length segments over the domain of the input data set; the value and slope of the piecewise-cubic function at each segments' endpoints comprise the parameters. The use of equal-length segments permits function evaluation to be $O(1)$ in the number of segments.

Basis functions with local support – one multiplying each parameter – are defined using Hermite polynomials; those for the value at the endpoints, as:

$$f_i\left((B-B_R)^2\right) = \begin{cases} 0 & \text{for } (B-B_R)^2 < (i-1)\Delta B^2 \\ h_{00}\left(\frac{(B-B_R)^2-(i-1)\Delta B^2}{\Delta B^2}\right) & \text{for } (i-1)\Delta B^2 \le (B-B_R)^2 < i\Delta B^2 \\ h_{01}\left(\frac{(B-B_R)^2-i\Delta B^2}{\Delta B^2}\right) & \text{for } i\Delta B^2 \le (B-B_R)^2 < (i+1)\Delta B^2 \\ 0 & \text{for } (i+1)\Delta B^2 \le (B-B_R)^2 \end{cases} \tag{27}$$

and those for slope at the endpoints, as:

$$g_i\left((B-B_R)^2\right) = \begin{cases} 0 & \text{for } (B-B_R)^2 < (i-1)\Delta B^2 \\ h_{10}\left(\frac{(B-B_R)^2-(i-1)\Delta B^2}{\Delta B^2}\right) & \text{for } (i-1)\Delta B^2 \le (B-B_R)^2 < i\Delta B^2 \\ h_{11}\left(\frac{(B-B_R)^2-i\Delta B^2}{\Delta B^2}\right) & \text{for } i\Delta B^2 \le (B-B_R)^2 < (i+1)\Delta B^2 \\ 0 & \text{for } (i+1)\Delta B^2 \le (B-B_R)^2 \end{cases} \tag{28}$$

where $0 \le i \le 6$ is an integer enumerating segment endpoints, $\Delta B^2$ is the length of each segment, and $h_{00}$, $h_{01}$, $h_{10}$, and $h_{11}$ are the Hermite cubics defined over the unit interval:

$$\begin{aligned} h_{00}(x) &= 2x^3 - 3x^2 + 1 \\ h_{01}(x) &= -2x^3 + 3x^2 \\ h_{10}(x) &= x^3 - 2x^2 + x \\ h_{11}(x) &= x^3 - x^2. \end{aligned} \tag{29}$$

By construction, each function $f_i$ has value equal to 1 at precisely one segment endpoint and zero at all other endpoints, and all functions $f_i$ have zero slope at all segment endpoints. Similarly, each function $g_i$ has slope (with respect to $(B-B_R)^2$) equal to 1 at precisely one segment endpoint and zero at all other endpoints, and all functions $g_i$ have zero value at all segment endpoints. Construction of the approximation proceeds by the standard linear

least-squares approach, finding the unique solution to

$$\underset{a_0,\ldots,a_6,b_0,\ldots,b_6}{\arg\min} \sum_j \left( \sum_{i=0}^{6} a_i f_i \left( \left(B_j - B_R\right)^2 \right) + b_i g_i \left( \left(B_j - B_R\right)^2 \right) - H_j \right)^2 \tag{30}$$

for a data set consisting of points $\left(B_j, H_j\right)$, by solution of the corresponding linear system. It is noted in [11] that in practice, this model is applied to the reluctivity of materials, and thus data sets of points $\left(B_j, \nu_j = H_j/B_j\right)$, but the approach is otherwise identical.

The approach presented in [11] does not ensure monotonicity: Forghani *et al.* state that "in the authors' experience, the choice of [piecewise polynomial] functions is very much one of convenience; any lack of monotonicity is due to the widespread use of least-squares fits, which tend to mildly oscillatory approximations." Additionally, the model in [11] has no extension into the saturation of the material beyond the last point of the input data set.

Heise, in [12] uses a similar approach to Forghani *et al.* by modelling reluctivity, $\nu$, as a cubic piecewise function *interpolating* the input data set $(B_i, \nu_i)$. As in [11], Hermite cubics are used as the segments of the function $\nu(B)$, however the work of Heise is based on that of Fritsch and Carlson [13], and thus ensures monotonicity by a similar means to that in [13]. Additionally, Heise defines the asymptotic behaviour of the function to be $\hat{\nu}(z) = \nu_0 - k_1 z^{-1} + k_2 z^{-2}$ for $z_n \leq z$ where $z_n$ is the last value of $B$ in the data set, and $k_1$ and $k_2$ are chosen to establish continuous differentiability at the end of the last cubic segment. Much of this work is focused on incorporation of this material representation into a 2D vector potential formulation of a non-linear magnetostatic finite element problem.

Heise's approach results in material curves representative of physical reality from the Rayleigh region through to saturation, and as it is an interpolative approach, discussion of errors (with repsect to the input data set) is inapplicable. However, the parameters of this representation number on the same order as the points in the input data set, and its extrapolation for the saturation region is solely a function of the value and slope of the last segment of the piecewise interpolant.

Finally, Pechstein and Jüttler present a hybrid of [11] and [12]: in [4] they demonstrate a method of *interproximation* of a B-H data set by $B$-splines – piecewise polynomial functions at whose segment endpoints (called *knots*) $\mathscr{C}^n$ continuity is enforced. As in [11], the number of piecewise segments is less than the number of input data points; as in [12], the number of parameters (i.e. the number of segments) is of the same order as the number of input data points, and monotonicity is ensured by construction. A quadratic minimization problem is solved with constraints placed on the relative error of the piecewise monotonic function

$B(H)$ evaluated at each data point $(H_i, B_i)$; thus the use of the term *interproximation* (a portmanteau of interpolation and approximation) to describe the approach. The piecewise functions resulting from the method of Pechstein and Jüttler share the same properties as other piecewise approaches: a large number of parameters are required to describe them, and the saturation region is defined by a single arctangent segment over the interval starting at the last point and extending to infinity.

# 4 Theory

The goal is to construct a rational function $B(H)$ for a given input data set $\{(H_i, B_i), 1 \leq i \leq n\}$, similar to that presented by Rivas *et al.* in [9], but with a numerator and denominator of higher degree (Rivas *et al.* used a quadratic numerator and denominator). This approach was first introduced by Widger in [14] and expanded upon in [15]; greater detail with regard to its theory and implementation, as well as a new non-linear least-squares improvement, are presented here. As in [14, 15], the objective is to construct rational functions in a relatively small number of parameters that, like the approaches presented in section 3.1, are smooth over the domain $H \geq 0$, and that approximate the input data set with a high degree of accuracy, like the approaches presented in section 3.2. Additionally, as in [15], a correction to the algorithm of [14] that removes non-physical poles and zeros of the rational functions is sought, and a further non-linear minimization of the error introduced.

The rational function $B(H)$ is constructed so as to satisfy the conditions for a physically valid B-H curve introduced in section 2:

(i) Continuous differentiability: $B(H)$, and its first derivative $\frac{dB(H)}{dH}$, are defined $\forall\, H \geq 0$,

(ii) Zero coercivity: $B(0) = 0$,

(iii) Monotonicity: $\frac{dB(H)}{dH} \geq \mu_0 \;\forall\, H \geq 0$,

(iv) Positivity of magnetization: $B(H) \geq \mu_0 H \;\forall\, H \geq 0$,

(v) Finite, positive saturation magnetization: $\lim_{H\to\infty} (B(H) - \mu_0 H) = \mu_0 M_{\text{sat}}$, where $M_{\text{sat}}$ is a finite, positive value.

In order to satisfy (ii) and (v), a rational function $f(H)$ whose numerator and denominator are

of the same degree $d$ is considered, such that:

$$f(H) = \frac{p(H)}{q(H)} = \frac{\sum_{k=1}^{d} p_k H^k}{1 + \sum_{k=1}^{d} q_k H^k} \tag{31}$$
$$B(H) = f(H) + \mu_0 H$$

where, by the fact that the numerator of $f(H)$ has a zero constant term, $B(0) = f(0) = 0$. Assuming $q_d \neq 0$, it follows by construction that:

$$\begin{aligned} \lim_{H\to\infty} (B(H) - \mu_0 H) &= \lim_{H\to\infty} f(H) \\ &= \lim_{H\to\infty} \frac{\sum_{k=0}^{d} p_k H^k}{1 + \sum_{k=1}^{d} q_k H^k} \\ &= \frac{p_d}{q_d} = \mu_0 M_{\text{sat}} \end{aligned} \tag{32}$$

and thus $M_{\text{sat}}$ is finite; for it to be positive, it suffices for $p_d q_d \geq 0$, that is, for $p_d$ and $q_d$ to have the same sign. However, given that $q(0) = 1$, and $q(H)$ cannot have positive real roots if $f(H)$ is to be continuous (condition (i)), it is necessary that $q_d > 0$ (and thus $p_d > 0$).

If, additionally, $p(H)$ and $q(H)$ have no positive real roots, that is, if $p(H) \neq 0 \; \forall \; H > 0$ and $q(H) \neq 0 \; \forall \; H \geq 0$, then $f(H)$ (and thus $B(H)$) is defined for all $H \geq 0$, as is its derivative (and that of $B(H)$), $\left(p'(H)q(H) - p(H)q'(H)\right)/q(H)^2$. Furthermore, this implies that $f(H)$ has no positive real zeros (and thus nor does $B(H)$), as it has the same zeros as $p(H)$. Since $f(H)$ is continuous, and positive in its limit as $H \to \infty$, it follows that $f(H)$ is positive for all $H \geq 0$. Therefore constraining $p(H)$ and $q(H)$ to have no positive real roots establishes conditions (i) and (iv).

Condition (iii), that $\frac{dB(H)}{dH} \geq \mu_0 \; \forall \; H \geq 0$, is established only if $\frac{df(H)}{dH} \geq 0 \; \forall \; H \geq 0$. This is equivalent to the constraint that $p'(H)q(H) - p(H)q'(H)$, the numerator of $\frac{df(H)}{dH}$, has no positive real zeros (as these two functions have the same zeros), and is positive for at least one value of $H \geq 0$. The derivative of $f(H)$ is necessarily positive for some value of $H \geq 0$ if conditions (i), (ii), and (v) are established, given that under these conditions $f(H)$ must continuously vary from a value $f(0) = 0$ to a value $\mu_0 M_{\text{sat}}$ as $H$ tends to infinity.

Thus, the five conditions establishing the physical validity of a B-H curve can be restated in

terms of conditions on the numerator and denominator of $f(H)$, $p(H)$ and $q(H)$ respectively:

(a) Continuous differentiability: $p(H) \neq 0 \; \forall \; H > 0$ and $q(H) \neq 0 \; \forall \; H \geq 0$,

(b) Monotonicity: $p'(H)q(H) - p(H)q'(H) \neq 0 \; \forall \; H \geq 0$,

(c) Positivity of magnetization and finite, positive saturation magnetization: $p_d > 0$ and $q_d > 0$;

noting that condition (ii), zero coercivity, holds by construction of $f(H)$.

With the form of $f(H)$ and $B(H)$ established, the next step is to find the values $\{p_1, \ldots, p_d, q_1, \ldots, q_d\}$ for $f(H)$ that (in some sense) make $B(H)$ a good approximation of the input data set $\{(H_i, B_i)\,,\; 1 \leq i \leq n\}$ at the points $(H_i, B_i)$.

## 4.1 Iterative Linear Least-Squares

As $B(H) = f(H) + \mu_0 H$, it follows that $f(H)$ should be fit to a modified version of the data set $\{(H_i, \mu_0 M_i)\,,\; 1 \leq i \leq n\}$, where $\mu_0 M_i = B_i - \mu_0 H_i$. If the standard squared error function $\varepsilon$ is considered,

$$\varepsilon = \frac{1}{2} \sum_{i=1}^{n} \left( \frac{\sum\limits_{k=1}^{d} p_k H_i^k}{1 + \sum\limits_{k=1}^{d} q_k H_i^k} - \mu_0 M_i \right)^2 , \tag{33}$$

it is apparent that the derivatives of $\varepsilon$ with respect to $p_k$ and $q_k$ are non-linear equations in those variables, and thus a linear least-squares approach is inapplicable. A non-linear least-squares method (e.g. Newton-Raphson) may be considered, however the excessive number of local minima of $\varepsilon$ makes such an approach futile. To illustrate this, $\varepsilon$ may be expressed in rational

form:

$$\begin{aligned}
\varepsilon &= \frac{1}{2}\sum_{i=1}^{n}\left(\frac{\left(\sum_{k=1}^{d} p_k H_i^k\right) - \mu_0 M_i\left(1+\sum_{k=1}^{d} q_k H_i^k\right)}{1+\sum_{k=1}^{d} q_k H_i^k}\right)^2 \\
&= \frac{1}{2}\sum_{i=1}^{n}\left(\frac{\left(\sum_{k=1}^{d} p_k H_i^k\right) - \mu_0 M_i\left(1+\sum_{k=1}^{d} q_k H_i^k\right)}{1+\sum_{k=1}^{d} q_k H_i^k}\left(\frac{\prod_{\substack{1\le j\le n\\ j\neq i}}\left(1+\sum_{j=1}^{d} q_k H_j^k\right)}{\prod_{\substack{1\le j\le n\\ j\neq i}}\left(1+\sum_{j=1}^{d} q_k H_j^k\right)}\right)\right)^2 \\
&= \frac{1}{2}\frac{\sum_{i=1}^{n}\left(\left(\left(\sum_{k=1}^{d} p_k H_i^k\right) - \mu_0 M_i\left(1+\sum_{k=1}^{d} q_k H_i^k\right)\right)\prod_{\substack{1\le j\le n\\ j\neq i}}\left(1+\sum_{j=1}^{d} q_k H_j^k\right)\right)^2}{\prod_{i=1}^{n}\left(1+\sum_{k=1}^{d} q_k H_i^k\right)^2}. \qquad (34)
\end{aligned}$$

From (34) it can be discerned that the numerator and denominator of $\varepsilon$ are multivariate polynomials in $\{p_1,\ldots,p_d,q_1,\ldots,q_d\}$ of degree $2n$. Consequently, the derivatives of $\varepsilon$ with respect to the variables $p_k$ are rational functions whose numerators are of degree $2n-1$, and whose denominators are of degree $2n$; the derivatives with respect to the variables $q_k$ are rational functions whose numerators are of degree $4n-2$, and whose denominators are of degree $4n$. Equating these derivatives with zero and multiplying through by their denominators yields a system of polynomial equations in $2d$ unknowns (the $p_k$ and $q_k$). By homogenizing these polynomials (i.e. by introducing an additional variable multiplying the monomial terms of each polynomial so as to make all monomial terms in a given polynomial of equal degree), Bézout's theorem [16] can be applied to yield an upper bound on the number of solutions to the polynomial system of equations, and thus the number of local minima of $\varepsilon$. This upper bound is the product of the degrees of the polynomials, equal to $(2n-1)^d(4n-2)^d = \left(8n^2-8n+2\right)^d$. For reasonable values of $n$ and $d$, it is apparent that the number of local minima may be extremely large (e.g. for $n=30$ and $d=5$, this upper bound is $\sim 1.6\times 10^{19}$), making the convergence of a non-linear solver (without a good starting point) unlikely, and motivating alternate approaches to the determination of the parameters $\{p_1,\ldots,p_d,q_1,\ldots,q_d\}$.

Turning to the disciplines of circuit analysis and systems theory, it is known that the transfer function of a linear time-invariant system is a rational function of the complex

frequency $j\omega$; equivalently, that the Laplace transform of its impulse response $\mathscr{L}\{h(t)\}(s) = H(s)$ is a rational function in $s$. Numerical determination of the coefficients of this rational function for large systems is often a complicated, ill-conditioned process, whereas evaluation (via experimentation) of their frequency response for a finite set of frequencies $\{j\omega_k\}$ is comparatively simple. Sanathanan and Koerner address the problem of approximating the rational transfer function of a linear dynamic system from a finite set of experimentally determined frequency response values in [1], presenting a method that is equally applicable to the problem of rational function fitting for B-H curve approximation. In [14], Widger adapts the approach of Sanathanan and Koerner to the minimization of the *relative* squared error of a rational function with respect to a B-H data set; here, as in [15], the absolute squared error is used.

When considering a set of points $\{(H_i, \mu_0 M_i),\ 1 \leq i \leq n\}$, the approach of Sanathanan and Koerner is to first express the error at each point $(H_i, \mu_0 M_i)$:

$$\varepsilon_i = \frac{p(H_i)}{q(H_i)} - \mu_0 M_i, \tag{35}$$

and then construct a *weighted* error, $\varepsilon_i'$, by multiplying $\varepsilon_i$ by $q(H_i)$:

$$\varepsilon_i' = q(H_i)\,\varepsilon_i = p(H_i) - \mu_0 M_i q(H_i). \tag{36}$$

The sum of these weighted errors, squared, produces a weighted squared error function $\varepsilon'$ which is quadratic in the parameters $\{p_1, \ldots, p_d, q_1, \ldots, q_d\}$:

$$\begin{aligned}
\varepsilon'(p_1, \ldots, p_d, q_1, \ldots, q_d) &= \frac{1}{2}\sum_{i=1}^{n} \varepsilon_i'^2 \\
&= \frac{1}{2}\sum_{i=1}^{n} \left(p(H_i) - \mu_0 M_i q(H_i)\right)^2 \\
&= \frac{1}{2}\sum_{i=1}^{n} \left(\left(\sum_{k=1}^{d} p_k H_i^k\right) - \mu_0 M_i \left(1 + \sum_{k=1}^{d} q_k H_i^k\right)\right)^2.
\end{aligned} \tag{37}$$

Minimization of this error function then proceeds as usual, by equating its derivatives with

respect to the parameters to zero:

$$\begin{aligned}
\forall\ j \in [1,d],\ \frac{d\varepsilon'}{dp_j} &= \sum_{i=1}^{n} \left( \left( \sum_{k=1}^{d} p_k H_i^k \right) - \mu_0 M_i \left( 1 + \sum_{k=1}^{d} q_k H_i^k \right) \right) H_i^j = 0 \\
\forall\ j \in [1,d],\ \frac{d\varepsilon'}{dq_j} &= \sum_{i=1}^{n} \left( \left( \sum_{k=1}^{d} p_k H_i^k \right) - \mu_0 M_i \left( 1 + \sum_{k=1}^{d} q_k H_i^k \right) \right) \left( -\mu_0 M_i H_i^j \right) = 0;
\end{aligned} \tag{38}$$

a $2d \times 2d$ system of linear equations which can be expressed in the form:

$$\mathbf{A}^{\mathrm{T}} \left( \mathbf{A}\vec{x} - \vec{b} \right) = \vec{0}, \tag{39}$$

where the entries of $\mathbf{A}$, $\vec{x}$, and $\vec{b}$ are:

$$\begin{aligned}
\mathbf{A}_{ij} &= \begin{cases} H_i^j, & \text{if } 1 \le j \le d \\ -\mu_0 M_i H_i^{j-d}, & \text{if } d+1 \le j \le 2d \end{cases} \\
\vec{x}_j &= \begin{cases} p_j, & \text{if } 1 \le j \le d \\ q_{j-d}, & \text{if } d+1 \le j \le d \end{cases} \\
\vec{b}_i &= \mu_0 M_i,
\end{aligned} \tag{40}$$

$\mathbf{A}$ being an $n \times 2d$ matrix, and $\vec{x}$ and $\vec{b}$ being $2d \times 1$ and $n \times 1$ column vectors, respectively. Constructing $\varepsilon'$ to linearize the minimization problem was originally proposed by Levy in [17]; Sanathanan and Koerner improved on this approach by noting that the weighted squared error function $\varepsilon'$ can be reweighted by dividing each error $\varepsilon_i'$ by the denominator of $f(H)$, $q(H)$, evaluated at the point $H_i$. This results in an iterative process: initialising $q^{(0)} = 1$, parenthesized superscripts denoting iteration count, the system ${\mathbf{A}^{(m)}}^{\mathrm{T}} \left( \mathbf{A}^{(m)}\vec{x}^{(m)} - \vec{b}^{(m)} \right) = \vec{0}$

is solved repeatedly, where:

$$
\begin{aligned}
\mathbf{A}_{ij}^{(m)} &= \begin{cases} \frac{H_i^j}{q^{(m-1)}(H_i)}, & \text{if } 1 \le j \le d \\ \frac{-\mu_0 M_i H_i^{j-d}}{q^{(m-1)}(H_i)}, & \text{if } d+1 \le j \le 2d \end{cases} \\
\vec{x}_j^{(m)} &= \begin{cases} p_j^{(m)}, & \text{if } 1 \le j \le d \\ q_{j-d}^{(m)}, & \text{if } d+1 \le j \le d \end{cases} \\
\vec{b}_i^{(m)} &= \frac{\mu_0 M_i}{q^{(m-1)}(H_i)} \\
p^{(m)}(H) &= \sum_{j=1}^{d} p_j^{(m)} H^j \\
&= \sum_{j=1}^{d} \vec{x}_j^{(m)} H^j \\
q^{(m)}(H) &= 1 + \sum_{j=1}^{d} q_j^{(m)} H^j \\
&= 1 + \sum_{j=1}^{d} \vec{x}_{j+d}^{(m)} H^j .
\end{aligned}
\tag{41}
$$

According to Sanathanan and Koerner, these "iterations tend to converge rapidly and the coefficients evaluated become effectively those obtained by minimizing the sum of $[\varepsilon_i^2]$ at all the experimental points." However, for a fixed point of this iteration, that is, where $q_1^{(m-1)} = q_1^{(m)}, \ldots, q_d^{(m-1)} = q_d^{(m)}$, it holds that:

$$
\begin{aligned}
\varepsilon'^{(m)} &= \frac{1}{2} \sum_{i=1}^{n} \left( \frac{p^{(m)}(H_i) - \mu_0 M_i q^{(m)}(H_i)}{q^{(m-1)}(H_i)} \right)^2 \\
&= \frac{1}{2} \sum_{i=1}^{n} \left( \frac{p^{(m)}(H_i)}{q^{(m)}(H_i)} - \mu_0 M_i \right)^2 \\
&= \varepsilon^{(m)}.
\end{aligned}
\tag{42}
$$

This does not establish the fixed point of the iteration as a local minimum of the (unweighted) squared error, $\varepsilon$, as the derivatives of $\varepsilon'^{(m)}$ with respect to $p_k$ and $q_k$ (where $q^{(m-1)}(H)$ is considered constant with respect to these variables) are not (necessarily) equal to the same derivatives of $\varepsilon$ (where $q(H)$ is a function of the variables $p_k$ and $q_k$), and thus the latter are not (necessarily) equal to zero for the values $p_k$ and $q_k$ at a fixed point. Nevertheless, Sanathanan

and Koerner's claim can be corroborated by comparison of these derivatives, evaluated at a fixed point of the iteration. The partial derivatives of $\varepsilon'^{(m)}$ with respect to $p_k$ and $q_k$ are:

$$\begin{aligned}\frac{\partial}{\partial p_k}\varepsilon'^{(m)} &= \frac{1}{2}\sum_{i=1}^{n}\left(\frac{p^{(m)}(H_i)-\mu_0 M_i q^{(m)}(H_i)}{q^{(m-1)}(H_i)}\right)\frac{H_i^k}{q^{(m-1)}(H_i)}\\ \frac{\partial}{\partial q_k}\varepsilon'^{(m)} &= \frac{1}{2}\sum_{i=1}^{n}\left(\frac{p^{(m)}(H_i)-\mu_0 M_i q^{(m)}(H_i)}{q^{(m-1)}(H_i)}\right)\frac{-H_i^k}{q^{(m-1)}(H_i)}(\mu_0 M_i)\end{aligned} \tag{43}$$

whereas the partial derivatives of $\varepsilon$ are:

$$\begin{aligned}\frac{\partial}{\partial p_k}\varepsilon &= \frac{1}{2}\sum_{i=1}^{n}\left(\frac{p(H_i)-\mu_0 M_i q(H_i)}{q(H_i)}\right)\frac{H_i^k}{q(H_i)}\\ \frac{\partial}{\partial q_k}\varepsilon &= \frac{1}{2}\sum_{i=1}^{n}\left(\frac{p(H_i)-\mu_0 M_i q(H_i)}{q(H_i)}\right)\frac{-H_i^k}{q(H_i)}\left(\frac{p(H_i)}{q(H_i)}\right)\end{aligned} \tag{44}$$

Comparing the derivatives in (43) and (44), it is apparent that the derivatives of $\varepsilon'^{(m)}$ and $\varepsilon$ with respect to the variables $p_k$ are equal (and zero) at a fixed point; this is to be expected given that both error functions are quadratic (and thus their derivatives, linear) in these coefficients. The same is not, in general, true for the derivatives with respect to the variables $q_k$. However, if $p(H_i)/q(H_i) \approx \mu_0 M_i$ for all $i$ – that is, if the rational function $p(H_i)/q(H_i)$ is a good approximation of the data set at the points $(H_i, \mu_0 M_i)$ – then it follows that the derivatives of $\varepsilon$ with respect to $q_k$ are *approximately* equal to those of $\varepsilon'^{(m)}$ (and thus approximately zero), and that said approximation has an error proportional to that of the rational function approximation of the data set as a whole.

Consequently, the iteration in (41), should it converge, may not converge to a local minimum of $\varepsilon$ – particularly if the squared error at the fixed point is large. In fact, an iterate obtained prior to the fixed point may be of a lower squared error than the fixed point itself. Since the iteration is not assured to converge, the first iterate $M$ for which $\varepsilon^{(M)} \leq \varepsilon^{(m)}$, $\forall\, m < M + \Delta M$ is deemed to be the result of this iterative least-squares fitting method, where $\Delta M$ is a positive integer representing the number of additional iterations to take beyond an iterate $m$ for which $\varepsilon^{(m)} < \varepsilon^{(m+1)}$ to establish that no squared error lower than $\varepsilon^{(m)}$ is achievable; through experimentation, a value of $\Delta M = 20$ is found to be sufficient.

## 4.2 Basis Functions

In [1, 14], the choice to express $p(H)$ and $q(H)$ as a linear combination of powers of $H$ is arbitrary. Although in exact arithmetic, the choice of basis has no effect on the result of the solution to linear systems such as those in (40) and (41), in finite precision arithmetic it may have a notable impact on the numerical stability of the solution process. Many polynomial bases, such as the Legendre polynomials, the Chebyshev polynomials, and the Bernstein polynomials, are intended for use over a fixed interval; often, this interval is $[-1,1]$ or $[0,1]$. For basis polynomials defined over the interval $[a,b]$, a modified data set is constructed, of the form:

$$\left\{ \left( \hat{H}_i = \frac{(b-a)\,H_i}{H_n} + a, \mu_0 M_i \right),\ 1 \leq i \leq n \right\} \tag{45}$$

Similarly, scaled, shifted versions of $p(H)$, $q(H)$, and consequently $f(H)$ are defined as:

$$\begin{aligned} \hat{p}\left(\hat{H}\right) &= p\left(H_n \frac{\hat{H}-a}{b-a}\right) \\ \hat{q}\left(\hat{H}\right) &= q\left(H_n \frac{\hat{H}-a}{b-a}\right) \\ \hat{f}\left(\hat{H}\right) &= f\left(H_n \frac{\hat{H}-a}{b-a}\right) = \frac{\hat{p}\left(\hat{H}\right)}{\hat{q}\left(\hat{H}\right)} \end{aligned} \tag{46}$$

Expressing $\hat{p}(\hat{H})$ and $\hat{q}(\hat{H})$ in terms of linearly independent polynomial basis functions $\beta_k(\hat{H})$ yields:

$$\begin{aligned} \hat{p}\left(\hat{H}\right) &= \sum_{k=0}^{d} \hat{p}_k \beta_k\left(\hat{H}\right) \\ \hat{q}\left(\hat{H}\right) &= \sum_{k=0}^{d} \hat{q}_k \beta_k\left(\hat{H}\right) \end{aligned} \tag{47}$$

With regard to the fitting process, the constraint $\hat{p}(a) = 0$ must be established for $B(H)$ to satisfy the zero coercivity condition, and one of the coefficients of $\hat{q}(\hat{H})$ must be fixed to prevent the system from being under-determined. In the case of an arbitrary basis, $\hat{p}(a) = 0$ implies:

$$\hat{p}\left(a\right) = \sum_{k=0}^{d} \hat{p}_k \beta_k\left(a\right) = 0 \tag{48}$$

and the constraint is established by:

$$\hat{p}_0 = -\frac{1}{\alpha_0(a)} \sum_{k=1}^{d} \hat{p}_k \beta_k(a), \tag{49}$$

where, without loss of generality, $\beta_0(\hat{H})$ is chosen so that $\beta_0(a) \neq 0$ (this choice is always possible due to the fact that a set of linearly independent polynomials cannot all share a common root). Similarly, choosing $\beta_0\left(\hat{H}\right)$ so that $\beta_0(a) \neq 0$ allows the choice of $\hat{q}_0 = 1$ (if $\beta_0(a) > 0$, or $\hat{q}_0 = -1$ if $\beta_0(a) < 0$) to make the linear systems solved during the fitting process non-singular while ensuring that $\hat{q}\left(\hat{H}\right)$ is positive and does not have a root at $\hat{H} = a$. Then, the iteration in (41), for an arbitrary basis $\beta_k(\hat{H})$ (where $\beta_0(a) \neq 0$, and $\hat{H} \in [a,b]$), is:

$$\begin{aligned}
\mathbf{A}_{ij}^{(m)} &= \begin{cases} \frac{\beta_j(\hat{H}_i) - \frac{\beta_j(a)}{\beta_0(a)}}{\hat{q}^{(m-1)}(\hat{H}_i)}, & \text{if } 1 \leq j \leq d \\ \frac{-\mu_0 M_i \beta_{j-d}(\hat{H}_i)}{\hat{q}^{(m-1)}(\hat{H}_i)}, & \text{if } d+1 \leq j \leq 2d \end{cases} \\
\vec{x}_j &= \begin{cases} \hat{p}_j, & \text{if } 1 \leq j \leq d \\ \hat{q}_{j-d}, & \text{if } d+1 \leq j \leq d \end{cases} \\
\vec{b}_i &= \frac{\mu_0 M_i}{\hat{q}^{(m-1)}\left(\hat{H}_i\right)} \\
\hat{p}^{(m)}\left(\hat{H}\right) &= \sum_{j=1}^{d} \hat{p}_j^{(m)} \left(\beta_j\left(\hat{H}\right) - \frac{\beta_j(a)}{\beta_0(a)}\right) \\
&= \sum_{j=1}^{d} \vec{x}_j^{(m)} \left(\beta_j\left(\hat{H}\right) - \frac{\beta_j(a)}{\beta_0(a)}\right) \\
\hat{q}^{(m)}\left(\hat{H}\right) &= \beta_0\left(\hat{H}\right) + \sum_{j=1}^{d} \hat{q}_j^{(m)} \beta_j\left(\hat{H}\right) \\
&= \beta_0\left(\hat{H}\right) + \sum_{j=1}^{d} \vec{x}_{j+d}^{(m)} \beta_j\left(\hat{H}\right);
\end{aligned} \tag{50}$$

from which $p(H)$, $q(H)$, and thus $f(H)$ are easily recovered by inverting the scaling and

shifting in (46):

$$\begin{aligned} p(H) &= \hat{p}\left(\frac{(b-a)H}{H_n}+a\right) \\ q(H) &= \hat{q}\left(\frac{(b-a)H}{H_n}+a\right) \\ f(H) &= \hat{f}\left(\frac{(b-a)H}{H_n}+a\right) = \frac{p(H)}{q(H)}. \end{aligned} \tag{51}$$

## 4.3 Non-Linear Least-Squares

Once the iterative least-squares algorithm yields a fit $\hat{f}\left(\hat{H}\right) = \hat{p}(\hat{H})/\hat{q}(\hat{H})$, it may be improved by applying a Newton-Raphson iteration minimizing the (unweighted) squared error, as experimentation demonstrates that the iterative least-squares algorithm often yields a good starting point for such a process. For this purpose the Bernstein basis,

$$\beta_k\left(\hat{H}\right) = \binom{d}{k}\left(1-\hat{H}\right)^{d-k}\hat{H}^k, \tag{52}$$

(where $\hat{H} = H/H_n$) is convenient, not only because of numerical stability considerations, but also because of the fact that $\beta_0(0) = 1$ and $\beta_k(0) = 0,\ \forall\, k > 0$, meaning that the zero coercivity condition is established by letting $\hat{p}_0 = 0$. The squared error function,

$$\begin{aligned} \varepsilon &= \frac{1}{2}\sum_{i=1}^{n}\left(\frac{\hat{p}\left(\hat{H}_i\right)}{\hat{q}\left(\hat{H}_i\right)} - \mu_0 M_i\right)^2 \\ &= \frac{1}{2}\sum_{i=1}^{n}\left(\frac{\sum\limits_{k=1}^{d}\hat{p}_k\beta_k\left(\hat{H}\right)}{\beta_0\left(\hat{H}\right)+\sum\limits_{k=1}^{d}\hat{q}_k\beta_k\left(\hat{H}\right)} - \mu_0 M_i\right)^2 \end{aligned} \tag{53}$$

is minimized by equating its first derivatives to zero:

$$\begin{aligned} \frac{d\varepsilon}{d\hat{p}_s} &= \sum_{i=1}^{n}\left(\frac{\hat{p}\left(\hat{H}\right)}{\hat{q}\left(\hat{H}\right)} - \mu_0 M_i\right)\frac{\beta_s\left(\hat{H}\right)}{\hat{q}\left(\hat{H}\right)} = 0,\ \forall\, s \in [1,d] \\ \frac{d\varepsilon}{d\hat{q}_s} &= -\sum_{i=1}^{n}\left(\frac{\hat{p}\left(\hat{H}\right)}{\hat{q}\left(\hat{H}\right)} - \mu_0 M_i\right)\frac{\beta_s\left(\hat{H}\right)}{\hat{q}\left(\hat{H}\right)^2} = 0,\ \forall\, s \in [1,d], \end{aligned} \tag{54}$$

resulting in a system of non-linear equations which can be solved with the Newton-Raphson iteration $\vec{x}^{(m+1)} = \vec{x}^{(m)} - \mathbf{J}^{(m)^{-1}} \vec{g}\left(\vec{x}^{(m)}\right)$, where:

$$
\begin{aligned}
\mathbf{J}^{(m)} &= \begin{pmatrix} \mathbf{A}^{(m)} & \mathbf{B}^{(m)} \\ \mathbf{B}^{(m)\mathrm{T}} & \mathbf{C}^{(m)} \end{pmatrix} \\
\mathbf{A}_{st}^{(m)} &= \frac{d^2\varepsilon}{d\hat{p}_s d\hat{p}_t} = \sum_{i=1}^{n} \frac{\beta_s\left(\hat{H}_i\right)\beta_t\left(\hat{H}_i\right)}{\hat{q}^{(m)}\left(\hat{H}_i\right)^2} \\
\mathbf{B}_{st}^{(m)} &= \frac{d^2\varepsilon}{d\hat{p}_s d\hat{q}_t} = \frac{d^2\varepsilon}{d\hat{p}_t d\hat{q}_s} = -\sum_{i=1}^{n} \frac{\beta_s\left(\hat{H}_i\right)\beta_t\left(\hat{H}_i\right)}{\hat{q}^{(m)}\left(\hat{H}_i\right)^2}\left(\frac{2\hat{p}^{(m)}\left(\hat{H}_i\right)}{\hat{q}^{(m)}\left(\hat{H}_i\right)} - \mu_0 M_i\right) \\
\mathbf{C}_{st}^{(m)} &= \frac{d^2\varepsilon}{d\hat{q}_s d\hat{q}_t} = \sum_{i=1}^{n} \frac{\beta_s\left(\hat{H}_i\right)\beta_t\left(\hat{H}_i\right)\hat{p}^{(m)}\left(\hat{H}_i\right)}{\hat{q}^{(m)}\left(\hat{H}_i\right)^3}\left(\frac{3\hat{p}^{(m)}\left(\hat{H}_i\right)}{\hat{q}^{(m)}\left(\hat{H}_i\right)} - 2\mu_0 M_i\right) \\
\vec{x}_s^{(m)} &= \begin{cases} \hat{p}_s^{(m)}, & \text{if } 1 \le s \le d \\ \hat{q}_{s-d}^{(m)}, & \text{if } d+1 \le s \le 2d \end{cases} \\
\vec{g}\left(\vec{x}^{(m)}\right) &= \left[\frac{d\varepsilon}{d\hat{p}_1}, \ldots, \frac{d\varepsilon}{d\hat{p}_d}, \frac{d\varepsilon}{d\hat{q}_1}, \ldots, \frac{d\varepsilon}{d\hat{q}_d}\right]^{\mathrm{T}}\Bigg|_{\hat{p}_1=\vec{x}_1^{(m)},\ldots,\hat{p}_d=\vec{x}_d^{(m)},\hat{q}_1=\vec{x}_{d+1}^{(m)},\ldots,\hat{q}_d=\vec{x}_{2d}^{(m)}} \\
\hat{p}^{(m)}\left(\hat{H}\right) &= \sum_{j=1}^{d} \hat{p}_j^{(m)} \beta_j\left(\hat{H}\right) = \sum_{j=1}^{d} \vec{x}_j^{(m)} \beta_j\left(\hat{H}\right) \\
\hat{q}^{(m)}\left(\hat{H}\right) &= \beta_0\left(\hat{H}\right) + \sum_{j=1}^{d} \hat{q}_j^{(m)} \beta_j\left(\hat{H}\right) = \beta_0\left(\hat{H}\right) + \sum_{j=1}^{d} \vec{x}_{j+d}^{(m)} \beta_j\left(\hat{H}\right).
\end{aligned}
\tag{55}
$$

Convergence of this process is determined by the usual means for a Newton-Raphson iteration (i.e. when $\left|\vec{x}^{(m+1)} - \vec{x}^{(m)}\right|_2 / \left|\vec{x}^{(m+1)}\right|_2$ falls below a given threshold). Should this iteration fail to converge to a local minimum for which $\varepsilon$ is smaller than that achieved by the iterative least-squares process, the result of the iterative least-squares process is used.

## 4.4 Curve Validation and Correction

Neither the iterative least-squares fit nor the non-linear least-squares fit produces a rational function $f(H)$ that is necessarily continuously differentiable or monotonically increasing. Although there is some assurance that the function well-approximates the input data set at the points $(H_i, \mu_0 M_i)$, poles and zeros may occur on the intervals between consecutive data points, as well as on the interval for which $H > H_n$. To establish continuous differentiability of $f(H)$, $p(H)$ and $q(H)$ are factored: if these polynomials are found to have no positive,

real roots, it holds that $f(H)$ is continuously differentiable. Similarly, should the factorization of $p'(H)\,q(H) - p(H)\,q'(H)$ – the numerator of the derivative of $f(H)$ – yield no positive real roots, monotonicity is established for $f(H)$ for $H \geq 0$; additionally, if $p_d > 0$ and $q_d > 0$, then $f(H)$ is monotonically increasing. A rational function for which none of $p(H)$, $q(H)$, and $p'(H)\,q(H) - p(H)\,q'(H)$ has positive real roots, and for which $p_d > 0$ and $q_d > 0$, is considered to *pass* validation; all other rational functions are considered to *fail* validation.

Experimentation demonstrates that rational functions resulting from the fitting process often fail validation due to *pole-zero pairs*: pairs consisting of a positive real root of $p(H)$ and a positive real root of $q(H)$ both occurring in the same interval between two points of the the input data set. That is, if

$$\begin{aligned} &p(H = Z) = 0,\ H_i \leq Z \leq H_{i+1} \\ &q(H = P) = 0,\ H_i \leq P \leq H_{i+1}, \end{aligned} \tag{56}$$

then $P$ and $Z$ constitute a pole-zero pair of $f(H)$ in the interval $H_i \leq H \leq H_{i+1}$. It is shown that removal of this pole-zero pair by division of $p(H)$ by its factor $(H - Z)$ and division of $q(H)$ by its factor $(H - P)$ has only a small impact on the *relative* squared error of the function $f(H)$ with respect to data set $\{(H_i, \mu_0 M_i),\ 1 \leq i \leq n\}$; the relative squared error at the point $(H_i, \mu_0 M_i)$ is defined as:

$$\rho_i^2 = \left(\frac{f(H_i) - \mu_0 M_i}{\mu_0 M_i}\right)^2 = \left(\frac{f(H_i)}{\mu_0 M_i} - 1\right)^2. \tag{57}$$

The change in relative squared error as a consequence of the pole-zero pair removal is then:

$$\begin{aligned} &\left(\frac{f(H_i)}{\mu_0 M_i}(1 + \delta_i) - 1\right)^2 - \left(\frac{f(H_i)}{\mu_0 M_i} - 1\right)^2 = 2\delta_i \frac{f(H_i)}{\mu_0 M_i}\left(\frac{f(H_i)}{\mu_0 M_i}\left(1 + \frac{\delta_i}{2}\right) - 1\right), \\ &\text{where } \delta_i = \frac{Z - P}{H_i - Z}. \end{aligned} \tag{58}$$

and, assuming that $f(H_i) \approx \mu_0 M_i$, the change in relative squared error is proportional to $\delta_i^2$. Given that $\delta_i^2$ is larger at the points $(H_i, \mu_0 M_i)$ and $(H_{i+1}, \mu_0 M_{i+1})$ (the endpoints of the interval containing $Z$ and $P$) than at any other points in the data set, the pole-zero pair removal has the *least* relative impact on the function $f(H)$ when $|Z - P| \ll Z - H_i$ and $|Z - P| \ll H_{i+1} - Z$, that is, when $Z$ and $P$ are close together and centred on the interval $[H_i, H_{i+1}]$. To further mitigate the impact of pole-zero pair removal, the iterative and non-linear least-squares fitting processes may be restarted with the reduced $f(H)$ (or its corresponding $\hat{f}\left(\hat{H}\right)$) used as a

starting point.

In some cases, a zero of $f(H)$ may occur in the interval $0 = H_1 \leq H \leq H_2$, without a nearby pole, causing validation to fail. This is indicative of the fact that the fitting process yielded a rational function $f(H)$ whose initial slope is negative, and can be resolved by restarting the fitting process with an additional constraint: that the slope at $H = H_1 = 0$ (for an arbitrary shift, that the slope at $\hat{H} = \hat{H}_1 = a$) is zero. Since $p(H_1) = p(0) = 0$ (i.e. $\hat{p}\left(\hat{H}_1\right) = \hat{p}(a) = 0$), this is a linear constraint; notably, $\hat{p}'\left(\hat{H}_1\right) = 0$:

$$\begin{aligned}
\left.\frac{d\hat{f}\left(\hat{H}\right)}{d\hat{H}}\right|_{\hat{H}=\hat{H}_1} &= \left.\left(\frac{d}{d\hat{H}}\frac{\hat{p}\left(\hat{H}\right)}{\hat{q}\left(\hat{H}\right)}\right)\right|_{\hat{H}=\hat{H}_1} \\
&= \frac{\hat{p}'\left(\hat{H}_1\right)\hat{q}\left(\hat{H}_1\right) - \hat{p}\left(\hat{H}_1\right)\hat{q}'\left(\hat{H}_1\right)}{\hat{q}\left(\hat{H}_1\right)^2} \\
&= \frac{\hat{p}'\left(\hat{H}_1\right)}{\hat{q}\left(\hat{H}_1\right)} = 0 \\
&\implies \hat{p}'\left(\hat{H}_1\right) = 0.
\end{aligned} \tag{59}$$

Notably, for the monomial basis (for which $\beta_k\left(\hat{H}\right) = \hat{H}^k = H^k$) and the Bernstein basis, this constraint is established by letting $\hat{p}_1 = 0$.

All other validation failures are incorrigible; there is empirical evidence to support the fact that they result from a choice of $d$ (that is, the degree of the numerator and denominator of $f(H)$) that is smaller than what is necessary to accurately represent the input data set by a continuously differentiable, monotonically increasing function. In these cases, as well as in the cases for which the error of the rational approximation (with respect to the data set) is intolerably large, the fitting process is restarted for a rational function whose numerator and denominator are of degree $d+1$.

# 5 Implementation Details

## 5.1 Linear Systems Solving

The iterative least-squares method requires repeated solution of a linear system of the form:

$$\mathbf{A}^{\mathrm{T}}\left(\mathbf{A}\vec{x} - \vec{b}\right) = 0 \tag{60}$$

which may be expressed as:

$$\mathbf{A}^{\mathrm{T}}\mathbf{A}\vec{x} = \mathbf{A}^{\mathrm{T}}\vec{b}. \tag{61}$$

Construction of $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ is computationally expensive, and prone to numerical round-off errors. Consequently, solving (60) is best performed using the QR decomposition of $\mathbf{A}$ [18]:

$$\begin{aligned} \mathbf{A} &= \mathbf{QR} \\ \mathbf{A}^{\mathrm{T}}\mathbf{A} &= \mathbf{R}^{\mathrm{T}}\mathbf{Q}^{\mathrm{T}}\mathbf{QR} = \mathbf{R}^{\mathrm{T}}\mathbf{R} \end{aligned} \tag{62}$$

where $\mathbf{Q}$ is an orthogonal matrix (for which $\mathbf{Q}^{\mathrm{T}}\mathbf{Q} = \mathbf{I}$) and $\mathbf{R}$ is an right-triangular matrix. If $\mathbf{A}$ is an $m$-by-$n$ matrix (where $m > n$), then the bottom $n - m$ rows of $\mathbf{R}$ are zero, and:

$$\mathbf{A} = \mathbf{QR} = \hat{\mathbf{Q}}\hat{\mathbf{R}} \tag{63}$$

where $\hat{\mathbf{Q}}$ is an $m$-by-$n$ matrix consisting of the first $n$ columns of $\mathbf{Q}$, and $\hat{\mathbf{R}}$ is an $n$-by-$n$ square, right-triangular matrix consisting of the first $n$ rows of $\mathbf{R}$. Clearly, since $\mathbf{Q}$ is orthogonal, $\mathbf{R}$ and $\hat{\mathbf{R}}$ have the same rank as $\mathbf{A}^{\mathrm{T}}\mathbf{A}$; if $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ is invertible, so is $\hat{\mathbf{R}}$. Thus, assuming a unique solution exists to (60), it follows that:

$$\begin{aligned} \vec{x} &= \left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\vec{b} \\ &= \left(\hat{\mathbf{R}}^{\mathbf{T}}\hat{\mathbf{R}}\right)^{-1}\hat{\mathbf{R}}^{\mathrm{T}}\hat{\mathbf{Q}}^{\mathrm{T}}\vec{b} \\ &= \hat{\mathbf{R}}^{-1}\hat{\mathbf{R}}^{-\mathrm{T}}\hat{\mathbf{R}}^{\mathrm{T}}\hat{\mathbf{Q}}^{\mathrm{T}}\vec{b} \\ &= \hat{\mathbf{R}}^{-1}\hat{\mathbf{Q}}^{\mathrm{T}}\vec{b} \\ &= \hat{\mathbf{R}}^{-1}\vec{b'} \end{aligned} \tag{64}$$

where $\vec{b'} = \hat{\mathbf{Q}}^{\mathrm{T}}\vec{b}$ is an $n$-element column vector. Instead of constructing $\hat{\mathbf{Q}}$ explicitly, it is noted that:

$$\hat{\mathbf{Q}}^{\mathrm{T}}\mathbf{A}\vec{x} = \hat{\mathbf{R}}\vec{x} = \hat{\mathbf{Q}}^{\mathrm{T}}\vec{b} = \vec{b'} \tag{65}$$

and that orthogonal matrices are closed under multiplication – that is, that the product of two orthogonal matrices is itself an orthogonal matrix. Consequently, $\hat{\mathbf{Q}}$ can be decomposed into a product of orthogonal matrices, multiplied by the $m$-by-$n$ matrix $\hat{\mathbf{I}}$ consisting of the first $n$

columns of the identity matrix:

$$\hat{\mathbf{Q}} = \left( \prod_{k=1}^{K} \mathbf{Q}_k \right) \hat{\mathbf{I}}. \tag{66}$$

Then $\hat{\mathbf{R}}$ can be constructed from $\mathbf{A}$ by premultiplying $\mathbf{A}$ by a series of orthogonal matrices, as can $\mathbf{Q}^{\mathrm{T}}\vec{b}$:

$$\begin{aligned} \hat{\mathbf{R}} = \hat{\mathbf{Q}}^{\mathrm{T}}\mathbf{A} &= \left( \hat{\mathbf{I}}^{\mathrm{T}} \prod_{k=K}^{1} \mathbf{Q}_k^{\mathrm{T}} \right) \mathbf{A} \\ \hat{\mathbf{Q}}^{\mathrm{T}}\vec{b} &= \left( \hat{\mathbf{I}}^{\mathrm{T}} \prod_{k=K}^{1} \mathbf{Q}_k^{\mathrm{T}} \right) \vec{b}. \end{aligned} \tag{67}$$

Specifically, the matrices $\mathbf{Q}_k$ can be chosen to represent Givens rotations - a row operation multiplying the rows $i$ and $j$ of the matrix they premultiply, taking the form:

$$\hat{\mathbf{Q}}_{k,st} = \begin{cases} 1, & \text{if } s = t \neq i \text{ and } s = t \neq j \\ \cos(\theta), & \text{if } s = t = i \text{ or } s = t = j \\ -\sin(\theta), & \text{if } s = i \text{ and } t = j \\ \sin(\theta), & \text{if } s = j \text{ and } t = i \\ 0, & \text{otherwise.} \end{cases} \tag{68}$$

Then, considering the partial right-triangularization of $\mathbf{A}$ after premultiplication by $k'$ matrices $\mathbf{Q}_k$,

$$\mathbf{R}_{k'} = \left( \prod_{k=k'}^{1} \mathbf{Q}_k^{\mathrm{T}} \right) \mathbf{A}, \tag{69}$$

it is evident that $i$, $j$, and $\theta$ in (68) can be chosen for $\mathbf{Q}_{k'+1}$ so as to introduce a new zero in the matrix $\mathbf{R}_{k'+1}$ that does not exist in the matrix $\mathbf{R}_{k'}$, without introducing any new non-zeros. The first $m-1$ Givens rotations are chosen so that the first column of $\mathbf{R}_{m-1}$ has zeros in all but its first entry; the next $m-2$ Givens rotations are chosen so that the second column of $\mathbf{R}_{2m-3}$ has zeros in all but its first two entries, and so on. Each Givens rotation can be applied to $\mathbf{A}$ and $\vec{b}$ in-place and concurrently; a total of $K = mn - \left(n^2+n\right)/2$ rotations are required to reduce the system $\mathbf{A}\vec{x} \approx \vec{b}$ to its right-triangular form $\hat{\mathbf{R}}\vec{x} = \vec{b}'$.

In terms of computational complexity, the QR decomposition algorithm described (when $\hat{\mathbf{Q}}$ is not constructed) requires $2mn^2 - \left(2n^3 + 6n^2 - 2n\right)/3$ floating-point multiplications and

$3\left(mn-\left(n^2+n\right)/2\right)$ trigonometric function evaluations to convert the matrix $\mathbf{A}$ to its right-triangular form $\hat{\mathbf{R}}$, and an additional $mn-\left(n^2+n\right)/2$ multiplications to perform the operation $\hat{\mathbf{Q}}^{\mathrm{T}}\vec{b}$. Determining the solution $\vec{x}$ is then a simple matter of performing back substitution on the matrix $\hat{\mathbf{R}}$ and right-hand side vector $\vec{b}'$, requiring $n^2-n$ multiplications and $n$ divisions.

This same algorithm is applicable to square systems for which $m=n$, and thus is also applied for the solution of the linear systems resulting from the Newton-Rapshon iteration in section 4.3.

## 5.2 Polynomial Evaluation and Deflation

Experimentation demonstrates that the monomial basis ($\beta_k(x)=x^k$) does not cause significant numerical round-off errors when applied to the iterative least-squares fitting method for rational functions of degree 9 or less, but this is not the case for the Newton-Raphson method – thus the use of the Bernstein basis. Each of these bases is evaluated by numerically stable means: the monomial basis using Horner's scheme [19], and the Bernstein basis using De Casteljau's algorithm [20]. Additionally, Horner's scheme has a second application: deflation of a polynomial, that is, synthetic division of a polynomial (expressed in the monomial basis) $p(x)$ by one of its linear factors $(x-z)$, where $z$ is a root of the polynomial.

### 5.2.1 Horner's Scheme

Given a polynomial expressed in the monomial basis,

$$p(x)=\sum_{k=0}^{d}a_k x^k \tag{70}$$

it can be rewritten as:

$$p(x)=a_0+x\left(a_1+x\left(a_2+\cdots+x\left(a_{d-1}+x\left(a_d\right)\right)\ldots\right).\right) \tag{71}$$

From this expression, Horner's scheme for evaluating the polynomial at a specific value of $x=x_0$ follows directly: beginning with $b_d=a_d$, the values $b_{d-1},\ldots,b_0$ are computed

iteratively as:

$$\begin{aligned} b_d &= a_d \\ b_{d-1} &= a_{d-1} + b_d x_0 \\ &\vdots \\ b_{k-1} &= a_{k-1} + b_k x_0 \\ &\vdots \\ b_0 &= a_0 + b_1 x_0 \end{aligned} \tag{72}$$

where, for the purpose of evaluation of $p(x)$, the values $b_k$ need not be stored after the computation of $b_{k-1}$. Put differently, $p(x_0)$ can be expressed in terms of $a_k$ and $b_k$:

$$\begin{aligned} p(x_0) &= a_0 + x\,(a_1 + x_0\,(a_2 + \cdots + x_0\,(a_{d-1} + x_0\,(b_d))\ldots)) \\ p(x_0) &= a_0 + x\,(a_1 + x_0\,(a_2 + \cdots + x_0\,(a_{d-2} + x_0\,(b_{d-1}))\ldots)) \\ &\vdots \\ p(x_0) &= a_0 + x_0\,(a_1 + \cdots + x_0\,(a_{d-k+1} + x_0\,(b_{d-k}))\ldots) \\ &\vdots \\ p(x_0) &= a_0 + x_0\,(b_1) \\ p(x_0) &= b_0. \end{aligned} \tag{73}$$

As stated in [19], evaluation of polynomials by Horner's scheme provides better numerical stability than the naïve approach, and is of the same computational complexity. Additionally, Horner's scheme can also be used to construct the polynomial $p(x)/(x-z)$, where $z$ is a root of the polynomial $p(x)$. Dividing the degree-$d$ polynomial $p(x)$ by $(x-z)$ yields an expression for $p(x)$ in terms of the $(d-1)$-degree quotient, $q(x)$, and a constant (zero) remainder $r$:

$$p(x) = r + q(x)\,(x-z)\,. \tag{74}$$

Using the $b_k$ in (72), it is shown that $q(x)$ can be expressed in terms of $b_1, \ldots, b_d$, when $p(x)$ is evaluated using Horner's scheme at the point $x_0 = z$:

$$\begin{aligned} p(x) &= a_0 + a_1 x + \cdots + a_{d-1} x^{d-1} + a_d x^d \\ &= (b_0 - b_1 z) + (b_1 - b_2 z)\, x + \cdots + (b_{d-1} - b_d z)\, x^{d-1} + b_d x^d \\ &= b_0 + \left( b_1 x + \cdots + b_{d-1} x^{d-1} + b_d x^d \right) - \left( b_1 + b_2 x + \cdots + b_d x^{d-1} \right) z \\ &= b_0 + \left( b_1 + \cdots + b_{d-1} x^{d-2} + b_d x^{d-1} \right) x - \left( b_1 + b_2 x + \cdots + b_d x^{d-1} \right) z \\ &= b_0 + \left( b_1 + \cdots + b_{d-1} x^{d-2} + b_d x^{d-1} \right) (x - z) \\ &= r + q(x)\,(x - z) \end{aligned} \tag{75}$$

where $a_k = b_k - b_{k+1} z$, and $p(z) = b_0 = r = 0$. Since the division of $p(x)$ by $(x - z)$ is unique, it follows that $q(x) = \sum_{k=1}^{d} b_k x^{k-1}$, and thus evaluation of $p(x)$ at one of its roots $z$ using Horner's scheme yields the coefficients of $q(x)$ – the deflation of $p(x)$ by its root, $z$. This application of Horner's scheme is used for the deflation of the numerator and denominator of rational functions in the process of pole-zero pair removal.

### 5.2.2 De Casteljau's Algorithm

Evaluation of Bernstein polynomials by De Casteljau's algorithm is derived directly from the recursive definition of the Bernstein basis polynomials. The base case for this recursive definition is:

$$\beta_{0,0}(x) = 1 \tag{76}$$

where $\beta_{d,k}$ denotes the $k^{\text{th}}$ polynomial in the Bernstein basis of degree $d$. Then, the Bernstein polynomials of degree $d$ are defined in terms of those of degree $d - 1$:

$$\beta_{d,k}(x) = \begin{cases} (1 - x)\, \beta_{d-1,0}(x), & \text{if } k = 0 \\ (1 - x)\, \beta_{d-1,k}(x) + x \beta_{d-1,k-1}(x), & \text{if } 0 < k < d \\ x \beta_{d-1,d-1}(x), & \text{if } k = d. \end{cases} \tag{77}$$

Using this definition, a polynomial of degree $d$ expressed as a linear combination of the Bernstein polynomials of degree $d$ may be expressed in terms of the Bernstein polynomials of

degree $d-1$:

$$\begin{aligned}
p(x) &= \sum_{k=0}^{d} a_k \beta_{d,k}(x) \\
&= a_0\beta_{d,0}(x) + a_1\beta_{d,1}(x) + \cdots + a_{d-1}\beta_{d,d-1}(x) + a_d\beta_{d,d}(x) \\
&= a_0(1-x)\beta_{d-1,0}(x) + a_1\left((1-x)\beta_{d-1,1}(x) + x\beta_{d-1,0}(x)\right) + \dots \\
&\quad + a_{d-1}\left((1-x)\beta_{d-1,d-1}(x) + x\beta_{d-1,d-2}(x)\right) + a_d x\beta_{d-1,d-1} \\
&= \left((1-x)a_0 + xa_1\right)\beta_{d-1,0}(x) + \left((1-x)a_1 + xa_2\right)\beta_{d-1,1}(x) + \dots \\
&\quad + \left((1-x)a_{d-2} + xa_{d-1}\right)\beta_{d-1,d-2}(x) + \left((1-x)a_{d-1} + xa_d\right)\beta_{d-1,d-1}(x) \\
&= \sum_{k=0}^{d-1} \left((1-x)a_k + xa_{k+1}\right)\beta_{d-1,k}(x).
\end{aligned} \tag{78}$$

This results in De Casteljau's algorithm, as introduced in [20], where the coefficients of a degree $d-1$ polynomial, expressed as a linear combination of the Bernstein basis polynomials of degree $d-1$, can be computed (for a given value of $x$) from the coefficients of a degree $d$ polynomial (expressed as a linear combination of the Bernstein basis polynomials of degree $d$). Concretely, where $p(x) = \sum_{k=0}^{d} a_k^{(0)} \beta_{d,k}(x)$, evaluation of $p(x)$ at $x = x_0$ consists of iteratively computing the coefficients $a_k^{(i)}$:

$$\begin{aligned}
a_0^{(1)} &= (1-x_0)a_0^{(0)} + x_0 a_1^{(0)},\ \dots,\ a_{d-1}^{(1)} = (1-x_0)a_{d-1}^{(0)} + x_0 a_d^{(0)} \\
&\vdots \\
a_0^{(i)} &= (1-x_0)a_0^{(i-1)} + x_0 a_1^{(i-1)},\ \dots,\ a_{d-i}^{(i)} = (1-x_0)a_{d-i}^{(i-1)} + x_0 a_{d-i+1}^{(i-1)} \\
&\vdots \\
a_0^{(d-1)} &= (1-x_0)a_0^{(d-2)} + x_0 a_1^{(d-2)},\ a_1^{(d-1)} = (1-x_0)a_1^{(d-2)} + x_0 a_2^{(d-2)} \\
a_0^{(d)} &= (1-x_0)a_0^{(d-1)} + x_0 a_1^{(d-1)} \\
&= p(x_0).
\end{aligned} \tag{79}$$

De Casteljau's algorithm, although providing numerical stability, comes at an increased computational cost when compared with Horner's scheme - it requires $d^2 + d$ multiplications and $O(d)$ space to evaluate the polynomial $p(x)$ at a point $x = x_0$ when expressed as a linear combination of Bernstein basis polynomials. However, given that the polynomials involved in the rational fitting process are of a small degree (i.e. $d \leq 9$), this cost is not prohibitive.

## 5.3 Polynomial Factorization

For the purpose of validation, it is essential that the factorization of a polynomial $p(x)$ is performed accurately; to this end, Aberth's method [21] is employed. This method is an iterative approach whose convergence depends on a reasonable initial choice of approximations to the roots of $p(x)$; the quality of this approximation, as well as the convergence of Aberth's method, are improved by constructing a polynomial $p^{\circ}(x)$ whose roots are equal in complex argument, but scaled in magnitude relative to those of $p(x)$ so as to be assured to exist within an annulus on the complex plane.

For a polynomial $p(x) = \sum_{k=0}^{d} a_k x^k$ expressed in the monomial basis, Fujiwara provides an upper bound $R$ on the magnitude of its roots in [22]:

$$R = 2\max\left(\left|\frac{a_{d-1}}{a_d}\right|, \left|\frac{a_{d-2}}{a_d}\right|^{\frac{1}{2}}, \ldots, \left|\frac{a_1}{a_d}\right|^{\frac{1}{d-1}}, \left|\frac{a_0}{2a_d}\right|^{\frac{1}{d}}\right). \tag{80}$$

This same bound can be used to establish a lower bound on the magnitude of the roots of $p(x)$, by considering the reciprocal polynomial $p^*(x)$ of $p(x)$:

$$\begin{aligned} p^*(x) &= x^d p\left(x^{-1}\right) \\ &= x^d \sum_{k=0}^{d} a_k x^{-k} \\ &= \sum_{k=0}^{d} a_k x^{d-k} \\ &= \sum_{k=0}^{d} a_{d-k} x^d, \end{aligned} \tag{81}$$

that is, the polynomial constructed by reversing the order of the coefficients of $p(x)$. It is evident that if $z \neq 0$ is a root of $p(x)$, then $z^{-1}$ is a root of $p^*(x)$, since $p^*\left(z^{-1}\right) = z^d p(z) = 0$. Therefore, a minimum bound $r$ can be placed on the magnitude of the roots of $p(x)$ based on the work of Fujiwara, assuming $p(0) \neq 0$ – if $p(0) = 0$, then $a_0 = 0$, and this root can be treated *a priori* by dividing $p(x)$ by $x$:

$$\begin{aligned} \frac{1}{r} &= 2\max\left(\left|\frac{a_1}{a_0}\right|, \left|\frac{a_2}{a_0}\right|^{\frac{1}{2}}, \ldots, \left|\frac{a_{d-1}}{a_0}\right|^{\frac{1}{d-1}}, \left|\frac{a_d}{2a_0}\right|^{\frac{1}{d}}\right) \\ r &= \frac{1}{2}\min\left(\left|\frac{a_0}{a_1}\right|, \left|\frac{a_0}{a_2}\right|^{\frac{1}{2}}, \ldots, \left|\frac{a_0}{a_{d-1}}\right|^{\frac{1}{d-1}}, \left|\frac{2a_0}{a_d}\right|^{\frac{1}{d}}\right) \end{aligned} \tag{82}$$

With the magnitude of the roots of $p(x)$ bound above and below by $R$ and $r$, respectively, the polynomial $p^\circ(x)$ is constructed so that its roots exist on the annulus on the complex plane whose area is divided equally in two by the circle of unit magnitude; empirically, this is found to yield good results when employing Aberth's method:

$$\begin{aligned} p^\circ(x) &= \sum_{k=0}^{d} a_k^\circ x^k \\ a_k^\circ &= \left(\frac{R^2+r^2}{2}\right)^{\frac{k}{2}} a_k \end{aligned} \tag{83}$$

where, if $z^\circ$ is a root of $p^\circ(x)$, then $z = z^\circ\sqrt{\left(R^2+r^2\right)/2}$ is a root of $p(x)$; a root $z^\circ$ of $p^\circ(x)$ lies on the annulus $r^\circ = \sqrt{2r^2/\left(R^2+r^2\right)} \leq |z^\circ| \leq \sqrt{2R^2/\left(R^2+r^2\right)} = R^\circ$.

The derivation of Aberth's method is beyond the scope of this document, and is detailed in [21]. The method consists of choosing initial approximations to the roots of $p^\circ(x)$, $z_1^{(0)}, \ldots, z_d^{(0)}$, and then iteratively constructing better approximations to said roots:

$$\forall\, k \in [1,d],\; z_k^{(n+1)} = z_k^{(n)} - \frac{\dfrac{p^\circ\left(z_k^{(n)}\right)}{p^{\circ\prime}\left(z_k^{(n)}\right)}}{1 - \dfrac{p^\circ\left(z_k^{(n)}\right)}{p^{\circ\prime}\left(z_k^{(n)}\right)} \displaystyle\sum_{j\neq k} \frac{1}{z_k^{(n)} - z_j^{(n)}}}. \tag{84}$$

Iteration ends when $\left(\sum_{k=1}^{d}\left|z_k^{(n+1)} - z_k^{(n)}\right|\right)/d$, the average change in the position of the roots on the complex plane, drops below a specified threshold. As suggested by Aberth in [21], the initial approximations to the roots should be chosen so as to not lie on a line on the complex plane. Given the earlier bounds established on the roots of $p^\circ(x)$, initial approximations are chosen to have a relatively even distribution in the annulus:

$$z_k^{(0)} = \left(r^\circ + \left(R^\circ - r^\circ\right)\frac{k-1}{d-1}\right) e^{\left(\frac{2k-1}{2d}\pi j\right)} \tag{85}$$

– points lying on a rotated, scaled segment of the Archimedean spiral in the complex plane.

## 5.4 Partial Fractions Decomposition

Both in the monomial basis, and in the Bernstein basis, the linear coefficients resulting from the rational fitting process to a B-H data set are found, empirically, to be of vastly varying magnitude. Additionally, evaluations of the integral and derivative (as are required

for loss calculations and Newton iterations, respectively, in finite element applications) are computationally expensive when the rational function is expressed in these forms. To rectify these issues, it is convenient to perform the partial fractions decomposition of the proper rational function $\mu(H) - \mu_0 = f(H)/H = p(H)/(Hq(H)) = r(H)/q(H)$ expressed in the monomial basis, where $r(H) = p(H)/H$ is easily computed given that $p(0) = 0$.

The first step in performing the partial fractions decomposition of $r(H)/q(H)$ is to factor the polynomial $q(H)$ – as it is a polynomial with real coefficients, it is expected to have (including multiplicity) $L$ linear factors of the form $(H - a_k)$, and $Q$ quadratic factors of the form $(H - b_k)^2 + c_k^2 = (H - (b_k + jc_k))(H - (b_k - jc_k))$, where $L + 2Q = d$, the degree of $q(H)$. Based on empirical evidence, the multiplicity of the roots of $q(H)$ can be assumed to always be 1; no cases (aside from contrived data sets) for which this assumption is invalid have been found. Then $r(H)/q(H)$ is expressed in its partial fractions form:

$$\begin{aligned}
\frac{r(H)}{q(H)} &= \sum_{k=1}^{L} \frac{d_k}{H - a_k} + \sum_{k=1}^{Q} \frac{e_k H + g_k}{(H - b_k)^2 + c_k^2} \\
&= \frac{\sum_{k=1}^{L} d_k \frac{q(H)}{H - a_k} + \sum_{k=1}^{Q} (e_k H + g_k) \frac{q(H)}{(H - b_k)^2 + c_k^2}}{q(H)} \\
r(H) &= \sum_{k=1}^{L} d_k \frac{q(H)}{H - a_k} + \sum_{k=1}^{Q} (e_k H + g_k) \frac{q(H)}{(H - b_k)^2 + c_k^2}.
\end{aligned} \tag{86}$$

Determination of the values $d_k$, $e_k$, and $g_k$ then consists of solving a $d$-by-$d$ system of equations. First, the coefficients of the deflations of $q(H)$ by the factors $(H - a_k)$ and $(H - b_k)^2 + c_k^2$ are computed in the monomial basis using Horner's scheme:

$$\begin{aligned}
\forall\, k \in [1, L],\ \lambda_k(H) &= \frac{q(H)}{H - a_k} = \sum_{i=0}^{d-1} \lambda_{k,i} H^i \\
\forall\, k \in [1, Q],\ \kappa_k(H) &= \frac{q(H)}{(H - b_k)^2 + c_k^2} = \sum_{i=0}^{d-2} \kappa_{k,i} H^i.
\end{aligned} \tag{87}$$

Then, the coefficients of $r(H)$ are expressed in terms of these coefficients:

$$\begin{aligned} r(H) &= \sum_{i=0}^{d-1} r_i H^i \\ &= \sum_{k=1}^{L} d_k \lambda_k(H) + \sum_{k=1}^{Q} (e_k H + g_k)\, \kappa_k(H) \\ &= \sum_{k=1}^{L} d_k \sum_{i=0}^{d-1} \lambda_{k,i} H^i + \sum_{k=1}^{Q} \sum_{i=0}^{d-2} e_k \kappa_{k,i} H^{i+1} + g_k \kappa_{k,i} H^i \\ &= \sum_{i=0}^{d-1} \left( \sum_{k=1}^{L} \left( d_k \lambda_{k,i} \right) + \sum_{k=1}^{Q} \left( e_k \kappa_{k,i-1} + g_k \kappa_{k,i} \right) \right) H^i \end{aligned} \tag{88}$$

where, for brevity of notation, $\kappa_{k,-1} = \kappa_{k,d-1} = 0$. This results in the $d$-by-$d$ linear system

$$\begin{pmatrix} \lambda_{1,0} & \cdots & \lambda_{L,0} & 0 & \cdots & 0 & \kappa_{1,0} & \cdots & \kappa_{Q,0} \\ \lambda_{1,1} & \cdots & \lambda_{L,1} & \kappa_{1,0} & \cdots & \kappa_{Q,0} & \kappa_{1,1} & \cdots & \kappa_{Q,1} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \lambda_{1,d-2} & \cdots & \lambda_{L,d-2} & \kappa_{1,d-3} & \cdots & \kappa_{Q,d-3} & \kappa_{1,d-2} & \cdots & \kappa_{Q,d-2} \\ \lambda_{1,d-1} & \cdots & \lambda_{L,d-1} & \kappa_{1,d-2} & \cdots & \kappa_{Q,d-2} & 0 & \cdots & 0 \end{pmatrix} \begin{bmatrix} d_1 \\ \vdots \\ d_L \\ e_1 \\ \vdots \\ e_Q \\ g_1 \\ \vdots \\ g_Q \end{bmatrix} = \begin{bmatrix} r_0 \\ \vdots \\ r_{d-1} \end{bmatrix} \tag{89}$$

whose solution yields the partial fractions decomposition of $r(H)/q(H)$. Then $\mu(H)$ takes the form:

$$\mu(H) = \mu_0 + \frac{r(H)}{q(H)} = \sum_{k=1}^{L} \frac{d_k}{H - a_k} + \sum_{k=1}^{Q} \frac{e_k \left( H + \frac{g_k}{e_k} \right)}{(H - b_k)^2 + c_k^2} \tag{90}$$

and it is noted that, conveniently,

$$\lim_{H \to \infty} f(H) = \mu_0 M_{\text{sat}} = \sum_{k=1}^{L} d_k + \sum_{k=1}^{Q} e_k. \tag{91}$$

The derivative of $\mu(H)$ with respect to $H$ is easily computed from its partial fractions

decomposition,

$$\frac{d\mu(H)}{dH} = \sum_{k=1}^{L} \frac{-d_k}{(H-a_k)^2} + \sum_{k=1}^{Q} \left( \frac{e_k}{(H-b_k)^2 + c_k^2} - \frac{2\,(e_k H + g_k)\,(H-b_k)}{\left((H-b_k)^2 + c_k^2\right)} \right) \tag{92}$$

and when $\mu(H)$ and its derivative are evaluated at the same point (as may occur during Newton-Raphson iterations in a finite element method), the floating-point operations required are:

- $3L+6Q-1$ additions/subtractions;
- $L+5Q$ multiplications;
- $2L+3Q$ divisions.

# 6 Results

## 6.1 Curve Approximation

To demonstrate the efficacy of the presented rational B-H curve approximation algorithm, it is applied to the Cobalt Steel - Hiperco 50, Stainless Steel - 416, and Castings - Cast Iron B-H data sets available from [23]. As is evidenced by the linear-log plot in Figure 1, the algorithm produces a good fit for all three data sets: the root-mean-squared (RMS) errors for the rational approximations to the data sets are 1.58 mT, 1.69 mT, and 3.75 mT, respectively. Additionally, the curves provide extrapolations beyond the last points in the data sets with reasonable saturation magnetizations: 2.275 T for the cobalt steel, 1.693 T for the stainless steel, and 1.711 T for the cast iron (the last points in these data sets have $\mu_0 M$ values of 2.264 T, 1.589 T, and 1.47 T, respectively). The rational function approximating the cobalt steel data set is of degree 5, that approximating the stainless steel data set is of degree 7, and that approximating the cast iron data set is of degree 4. These rational functions (expressing permeability as a function of magnetic field strength) are found in the appendix. Construction and validation of all three functions required 2.28 seconds on a single-core of an Intel Core 2-generation Xeon processor running at 2.50 GHz.

The algorithm is also applied to 197 B-H data sets from Infolytica Corporation's finite element software, MagNet [24], of which 5 fail to yield a pole-free, monotonically increasing rational function of degree 9 or lower, resulting in 192 successful approximations: a 97%

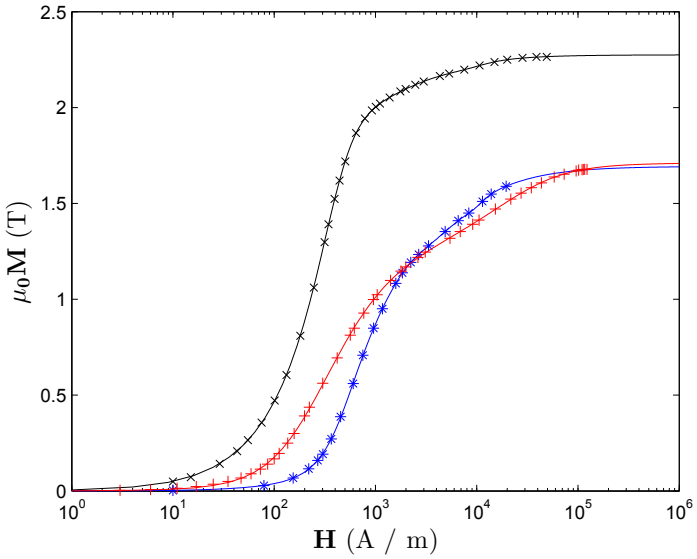


Figure 1: Magnetization (multiplied by $\mu_0$) versus magnetic field intensity for three sample materials: Cobalt Steel - Hiperco 50 data set ('x' markers), Stainless Steel - 416 data set ('*' markers), Castings - Cast Iron data set ('+' markers), and their respective rational approximations (smooth lines). Horizontal axis is logarithmic; vertical axis is linear.

success rate. All 5 data sets that fail to produce valid rational approximations consist of fewer than 18 data points; the 11 data sets whose rational approximation has an RMS error greater than 50 mT consist of fewer than 23 data points. Of the 181 remaining data sets yielding valid rational approximations, the average size of a data set is 69 data points. Based on this empirical evidence, it is concluded that the primary cause of failure of the algorithm is insufficiently large data sets; no other property common to the data sets yielding poor or invalid rational approximations is identified. The average RMS error of the 192 data sets' approximating rational functions is 9.49 mT. Construction and validation of all 192 functions required 195 seconds on a single-core of an Intel Core 2-generation Xeon processor running at 2.50 GHz; an additional 533 milliseconds were spent constructing functions for the 5 data sets that failed validation.

Visual inspection indicates that few oscillations (i.e. spurious changes in concavity) are present in the rational functions resulting from the fitting process: a desirable characteristic that is often absent in piecewise-polynomial representations. Pole-zero pair correction is necessary for 18 data sets to be represented by a physically valid rational function. An additional 25 are represented by rational functions more accurately due to pole-zero pair correction, with an average decrease in RMS error of 3.93 mT.

| Degree | No. of Data Sets |
|---|---|
| 3 | 30 |
| 4 | 38 |
| 5 | 21 |
| 6 | 44 |
| 7 | 21 |
| 8 | 17 |
| 9 | 21 |

Table 1: Degrees of 192 rational approximations of data sets from Infolytica Corporation's MagNet [24]

Table 1 provides a summary of the degree of the rational functions resulting from the approximation of MagNet's data sets. A histogram illustrating the distribution of RMS errors of the 181 rational approximations whose RMS error is less than 50 mT is found in Figure 2.

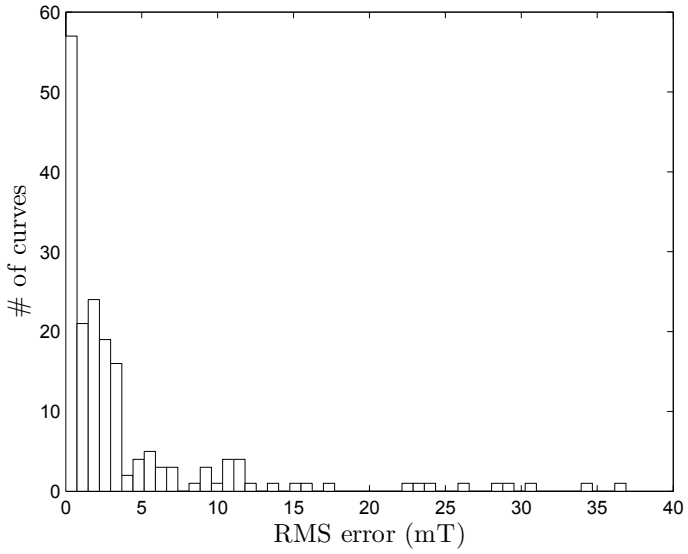


Figure 2: Histogram of RMS errors of 181 B-H rational curves; bars are 0.8 mT wide.

## 6.2 Finite Element Method

To establish that the rational approximations are suitable for finite element applications, the TEAM 13 problem [25] is solved. The TEAM 13 problem consists of a 1000 AT coil surrounded by two C-shaped 3.2 mm-thick steel channels, symmetrically offset with respect to the x-axis (as seen in Figure 3(b)), magnetically coupled by a 3.2 mm-thick steel plate centered in the coil and between the channels. Full details of the TEAM 13 are be found

in [25].

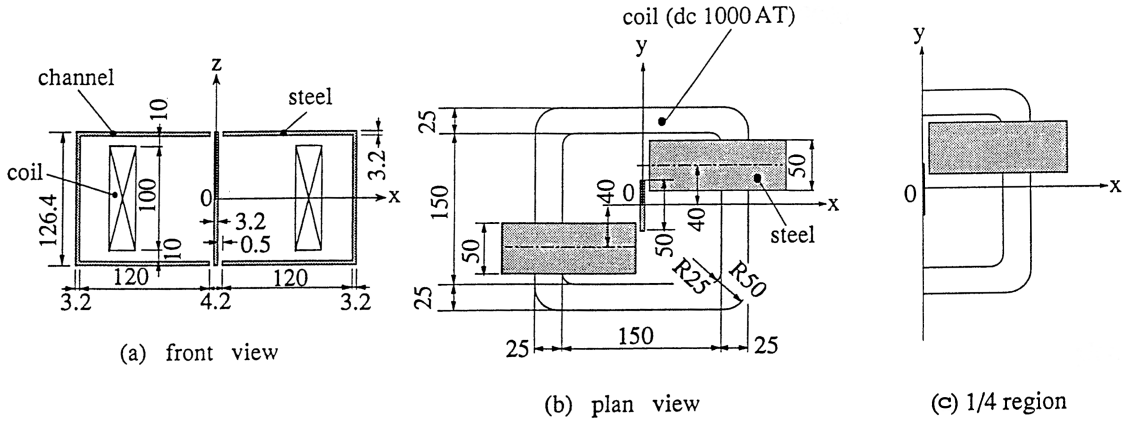


Figure 3: Diagrams of the TEAM 13 problem solid geometry from [25]. Dimensions are in mm.

The data set fit for the non-linear material in this problem consists of 42 points, and is the same as is used in [15]: "the 27 points provided in [25] for $0 \leq B \leq 1.8$, an additional 13 points, evenly distributed in $B$, in the range $1.825 \leq B \leq 2.175$ (in which $B$ is a quadratic function of $H$, specified in [25]), and two points at $B = 2.25$ T and 2.375 T (in the interval in which $B$ is a linear function of $H$, again, specified in [25]);" this data set is found in Table 2 in the appendix. The resulting fit has an RMS error of 7.24 mT and consists of a degree-7 rational function. In its partial fractions decomposed form, it is:

$$\mu(H) = \mu_0 + \frac{0.689308}{H + 6842.83} + \frac{1.13862\,(H + 178.661)}{(H - 26.4768)^2 + 443.224^2} + \frac{0.213382\,(H - 249.172)}{(H - 221.138)^2 + 90.4063^2} + \frac{0.124607\,(H - 5885.15)}{(H - 16985.0)^2 + 20790.4^2}. \tag{93}$$

The function $f(H) = (\mu(H) - \mu_0)H$ (where $\mu(H)$ is as defined in (93)) is plotted with the data points used to construct it in Figure 4. The finite element problem is constructed using the 1/4 region (as specified in [25], and seen in Figure 3(c)) with odd-periodic and magnetic field normal boundaries. An air region three times larger than the bounding box of the solid geometry encloses the steel plates and (copper) coils, all of which is meshed with tetrahedral elements. The finite element method solver employed uses second-order polynomial basis functions for the magnetic scalar potential in each tetrahedral element [24].

The TEAM 13 problem statement includes the value of the average flux density (in the steel) measured at 25 different points using a sensing coil. To compare these with the field solution resulting from the finite element method used, the average flux densities normal to

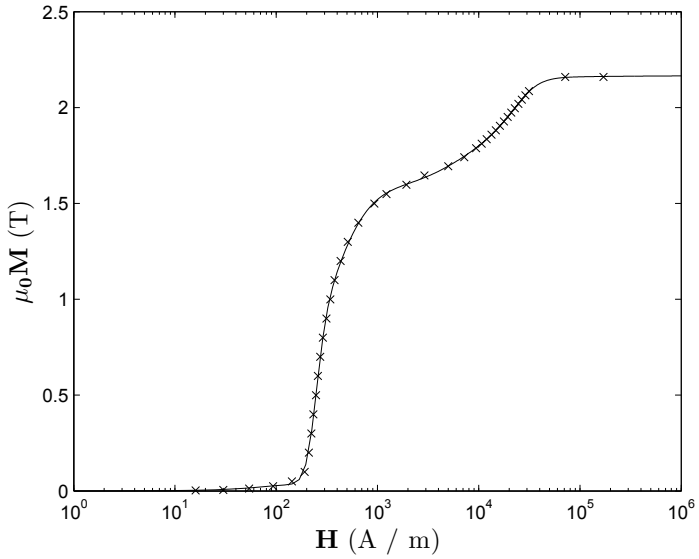


Figure 4: Magnetization (multiplied by $\mu_0$) versus magnetic field intensity for TEAM 13 steel ('x' markers) and its rational approximation (smooth line). Horizontal axis is logarithmic; vertical axis is linear.

the sensing coil cross-sections are computed by numeric integration of the B field solution; these flux densities, along with the measurements provided in [25], are displayed in Figure 5.

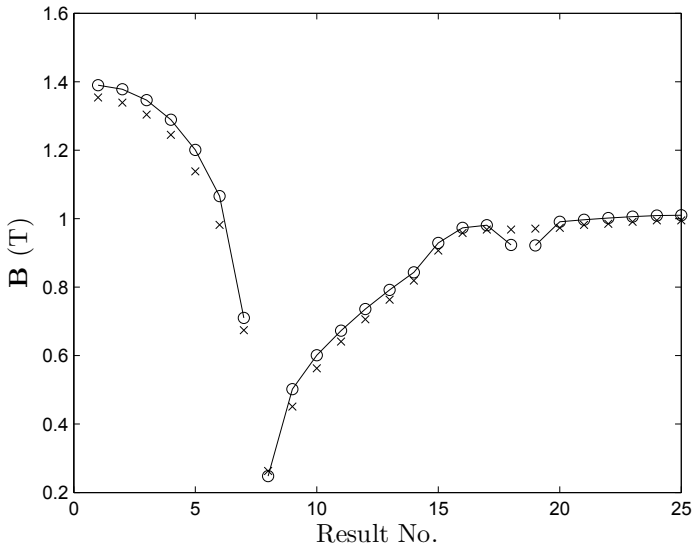


Figure 5: Comparison of finite-element results achieved with rational B-H curve representation ('o' markers) and measured results from TEAM 13 ('x' markers)

# 7 Conclusions

Current approaches to the construction of continuously differentiable B-H curves do not provide as elegant and accurate an approximation as the rational functions presented here: smooth approaches (such as those presented by Trutt *et al.* [6]) do not approximate data sets for modern materials with sufficient accuracy, whereas piecewise approaches have definitions in a large number of parameters, and extrapolation regions defined solely based on value and slope of the penultimate piecewise segment evaluated at the last data point. Rational functions provide a smooth (i.e. infinitely differentiable), accurate approximation to input B-H data sets defined over all values of $H \geq 0$, and are capable of representing the vast majority of B-H data sets. Shortcomings of the rational fitting algorithm, notably the possible existence of pole-zero pairs in intervals spanning consecutive data points, are overcome with corrections to its resulting rational functions whose impact on the quality of the approximation is limited. Evaluation of the rational functions require a small number of floating-point operations, on the same order as polynomials of similar degree. When applied to finite element problems, rational functions are demonstrated to produce results in agreement with those determined by experimentation.

The method presented here addresses the issue of approximating a data set in the B-H plane by a univariate rational function $B(H)$; it may be possible to extend this approach to data sets where $H$ and one or more other variables, such as frequency or temperature, to produce multivariate rational functions (e.g. $B(H, f, T)$) approximating these data sets. Such an approximation has particular value in the context of modern design problems, where frequency- and temperature- dependence of the results are non-negligible. Extension of rational B-H curve representation to rational B-H-$f$ *surface* or B-H-$f$-$T$ *hypersurface* representation is non-trivial, as pole-zero correction does not extend to higher dimensions, necessitating further research. Additionally, hysteresis loops and anhysteretic curves are likely candidates for rational function approximation. It is plausible, given the shape and properties of these curves, that rational fitting may also be applied to them, and the resulting functions used in hysteresis models (such as the Jiles-Atherton [26]). In short, the versatility and expressiveness of rational functions is only partially explored in this thesis; it is expected that they may be applicable to the smooth representation of many other data sets pertaining to non-linear magnetic materials.

# 8 Appendix

$$\mu(H) = \mu_0 + \frac{1.83926}{H + 421.224} + \frac{0.422798\,(H + 341.845)}{(H - 197.632)^2 + 394.636^2} + \frac{0.0128382\,(H - 11836.4)}{(H - 7019.43)^2 + 6693.82^2} \text{ H / m}$$

Rational permeability of Cobalt Steel - Hiperco 50; data set obtained from [23]

$$\mu(H) = \mu_0 + \frac{1.35710}{H + 1684.66} + \frac{0.391934\,(H - 337.895)}{(H - 246.599)^2 + 455.720^2} + \frac{-0.0627054\,(H - 2736.34)}{(H - 1703.87)^2 + 1727.36^2} + \frac{0.00664084\,(H - 12191.3)}{(H - 10031.3)^2 + 3294.42^2} \text{ H / m}$$

Rational permeability of Stainless Steel - 416; data set obtained from [23]

$$\mu(H) = \mu_0 + \frac{0.424708\,(H + 32849.8)}{(H + 19171.1)^2 + 18818.7^2} + \frac{1.28663\,(H + 35.7335)}{(H + 140.617)^2 + 206.224^2} \text{ H / m}$$

Rational permeability of Castings - Cast Iron; data set obtained from [23]

| H (A / m) | B (T) | H (A / m) | B (T) | H (A / m) | B (T) |
|---|---|---|---|---|---|
| 0 | 0 | 313 | 0.9 | 11998 | 1.85 |
| 16 | 0.0025 | 342 | 1 | 13350 | 1.875 |
| 30 | 0.005 | 377 | 1.1 | 14752 | 1.9 |
| 54 | 0.0125 | 433 | 1.2 | 16209 | 1.925 |
| 93 | 0.025 | 509 | 1.3 | 17729 | 1.95 |
| 143 | 0.05 | 648 | 1.4 | 19321 | 1.975 |
| 191 | 0.1 | 933 | 1.5 | 20996 | 2 |
| 210 | 0.2 | 1228 | 1.55 | 22768 | 2.025 |
| 222 | 0.3 | 1934 | 1.6 | 24657 | 2.05 |
| 233 | 0.4 | 2913 | 1.65 | 26690 | 2.075 |
| 247 | 0.5 | 4993 | 1.7 | 28904 | 2.1 |
| 258 | 0.6 | 7189 | 1.75 | 31360 | 2.125 |
| 272 | 0.7 | 9423 | 1.8 | 71620 | 2.25 |
| 289 | 0.8 | 10691 | 1.825 | 171092 | 2.375 |

Table 2: TEAM13 B-H data set for rational approximation.

# References


[1] C. K. Sanathanan and J. Koerner, "Transfer Function Synthesis as a Ratio of Two Complex Polynomials," *IEEE Trans. Autom. Control,* vol. 8, no. 1, pp. 56-58, Jan. 1963.

[2] D. C. Jiles, "Dynamics of Domain Magnetization and the Barkhausen Effect," *Czechoslovak J. Phys.*, vol. 50, no. 8, pp. 893-988, 2000.

[3] Sergey E. Zirka, Yury I. Moroz, Philip Marketos, and Anthony J. Moses, "Congruency-Based Hysteresis Models for Transient Simulation," *IEEE Trans. Magn.,* vol. 40, no. 2, pp. 390-399, Mar. 2004.

[4] C. Pechstein and B. Jüttler, "Monotonicity-preserving interproximation of B-H-curves," *J. Comput. and Appl. Math.,* vol. 196, no. 1, pp. 45-57, Nov. 2006.

[5] T. Hülsmann, *Nonlinear Material Curve Modeling and Sensitivity Analysis for MQS-Problems,* M.S. thesis, Faculty Inf., Media, Elect. Eng., Bergische Univ. Wuppertal, Wuppertal, Germany, 2012.

[6] F. C. Trutt, E. A. Erdélyi, and R. E. Hopkins, "Representation of the Magnetization Characteristic of DC Machines for Computer Use," *IEEE Trans. Power App. Syst.,* vol. 87, no. 3, pp. 665-669, Mar. 1968.

[7] J. Fischer and H. Moser, "Die Nachbildung von Magnetisierungskurven durch einfache algebraische oder transzendente Funktionen," *Archiv. für Elecktrotecnik*, XLII. Band, Heft 5, pp. 286-299, 1956.

[8] J. R. Brauer, "Simple Equations for the Magnetization and Reluctivity Curves of Steel," *IEEE Trans. Magn.,* vol. 11, no. 1, p. 81, Jan. 1975.

[9] J. Rivas, J. M. Zamarro, E. Martín, and C. Pereira, "Simple Approximation for Magnetization Curves and Hysteresis Loops," *IEEE Trans. Magn.,* vol. 17, no. 4, pp. 1498-150, Jul. 1981.

[10] M. R. Pagnola, F. D. Saccone, A. Ozols, and H. Sirkin, "Improvement to the second-order rational functions approximation for hysteresis cycles of magnetic materials," *Int. J. Comput. Math. Elect. Electron. Eng.,* vol. 28, no. 6, pp. 1579-1589, Sep. 2009.

[11] B. Forghani, E. M. Freeman, D. A. Lowther, and P. P. Silvester, "Interactive Modeling of Magnetization Curves," *IEEE Trans. Magn.,* vol. 18, no. 6, pp. 1070-1072, Nov. 1982.

[12] B. Heise, "Analysis of a Fully Discrete Finite Element Method Problem for a Nonlinear Magnetic Field Problem," *SIAM J. Numer. Anal.,* vol. 31, no. 3, pp. 745-759, Jun. 1994.

[13] F. N. Fritsch and R. E. Carlson. "Monotone Piecewise Cubic Interpolation," *SIAM J. Numer. Anal.,* vol. 17, no. 2, pp. 238-246, Apr. 1980.

[14] G. F. T. Widger, "Representation of magnetisation curves over extensive range by rational-fraction approximations," *Proc. Inst. Elect. Eng.,* vol. 116, no. 1, pp. 156-160, Jan. 1969.

[15] P. Diez and J. P. Webb, "A Rational Approach to B-H Curve Representation," to appear in *IEEE Trans. Magn.,* vol. 52, no. 3, Mar. 2016.

[16] É. Bézout, *Théorie générale des équations algébriques,* Paris, France: Ph.-D. Pierres, 1779.

[17] E. C. Levy, "Complex curve fitting," *IRE Trans. on Autom. Control,* vol. AC-4, pp. 37-44, May 1959.

[18] W. M. Gentleman, "Error Analysis of QR Decompositions by Givens Transformations," *Linear Algebra and its Applications*, vol. 10, no.3, pp. 189-197, Jun. 1975.

[19] C. Sidney Burrus, J. W. Fox, G. A. Sitton, and S. Treitel, "Horner's Method for Evaluating and Deflating Polynomials," Rice University. Houston, Texas, United States, Nov. 2003.

[20] P. De Casteljau, "Courbes et surfaces à poles," *Andr e Citro en Automobiles SA*, Paris, 1963.

[21] O. Aberth, "Iteration Methods for Finding all Zeros of a Polynomial Simultaneously," *Math. Comp.,* vol. 27, no. 122, Apr. 1973.

[22] M. Fujiwara, "Über die obere Schranke des absoluten Betrages der Wurzeln einer algebraishen Gleichung," Tôhoku Math. J., vol. 10, pp. 167-171, Jul. 1916.

[23] *Free BH Curves – MagWeb* [Online]. Available: `http://www.magweb.us/free-bh-curves/`, accessed Dec. 18, 2014.

[24] *MagNet 7.5.1*. Montréal, Canada: Infolytica Corporation, 2015.

[25] *TEAM Problem 13: 3-D Non-Linear Magnetostatic Model* [Online]. Available: `http://www.compumag.org/jsite/images/stories/TEAM/problem13.pdf`, accessed Dec. 22, 2014.

[26] D. C. Jiles and D. L. Atherton, “Theory of Ferromagnetic Hysteresis,” *J. Magn. Magn. Mater.,* vol. 61, no. 1-2, pp. 48-60, Sep. 1986.