Title: A General Approach for Learning Constitutive Relationships (Physical Laws) from Neural Networks

**Authors:** Andrew Temple1†, Matthew Miller2†, Maria Quintana-Hernandez1, Aaron P. Stebner2, Peter C. Collins1\*, Branden B. Kappes1,3\*.

**Affiliations:**

1Department of Materials Science and Engineering, Iowa State University, Ames, IA, 50011, USA.

2Department of Mechanical Engineering, Colorado School of Mines, Golden, CO, 80401, USA.

3KMMD, LLC., Castle Rock, CO, 80108, USA.

\*Correspondence to: Peter Collins (pcollins@iastate.edu), Branden Kappes (bkappes@kmmd.io).

†These authors contributed equally.

**Abstract:** The abstract should be 100-125 words, and organized in this structure: An opening sentence that sets the question that you address and is comprehensible to the general reader, background content specific to this study, results, and a concluding sentence. It should be a single paragraph.

**One Sentence Summary:** Gleaning concise mathematical expressions from the representational power of machine learning models and series expansions.

**Main Text:** In general, this should include a brief (1-2 paragraph) introduction, followed by a statement of the specific scope of the study, followed by results and then interpretations. Please avoid statements of future work or claims of priority, and avoid repeating the conclusions at the end.

**Subheadings** (“Results”, “Discussion”, or more specific subheadings, but not a leading “Introduction”) may be included in Research Articles or Reviews and should be brief, set off by a line break and formatted in bold face. Reports should not have subheadings.

All Figures and Tables should be cited in order, including those in the Supplementary Material (which should be cited as, for example, “Fig. S1”, and “Table S1”). You may include line or page breaks if you would like to place the figures within the text near where they are referenced. Please do not place figures in text boxes.

References should be cited in parentheses with an italic number (*1*). Multiple reference citations are separated by commas *(2, 3)* or if a series, dashes *(4-6)*. References are cited in order by where they first are called out, through the text, captions, and then the supplementary material.

Equations can be included. We do not recommend using the native Word 2007, 2008, 2010, or 2011 equation editor. This can in some cases produce less reliable MathML, the online markup language we use, which may result in display errors. Instead, use the legacy equation editor in Word (Chose Insert > Insert Object > Word Equation) or use Mathtype (recommended). If you enter equations in simple LaTeX, check that they will convert accurately (Word 2007 and higher can convert simple LaTeX equations).

**Introduction.** Various physical phenomenon, those well-described by physical laws (thermodynamics, kinetics, motion) or phenomenological theories (plasticity or [NON METALLURGICAL EXAMPLE]), are captured by constitutive models, models that express cause-effect relationships in an analytical or differential form. Rooted in point estimation and decision theory [REF], machine learning (ML) strategies extend descriptive modeling to include more general correlative relationships. While it is seductive to trust machine learning models to solve the most complex problems, these models, unlike constitutive models, do not differentiate causal relationships from correlative ones and, therefore, are susceptible to the *ad hoc ergo propter hoc* logical fallacy. Should *Y* follow *X* is a necessary but not a sufficient condition to conclude that *X* causes *Y*. Being unable to differentiate causal relationships from correlative ones, machine learning models cannot, in general, develop new physical laws. Because they describe an underlying causal relationship, laws and theories permit us to extrapolate, to hypothesize, theorize, and test hypotheses. Should mastery of our understanding of physical processes suffer, the scientific method in its currently and widely accepted form, becomes harder to apply. Thus, we must: retain mastery of current theory; postulate logically new theories; and leverage mathematical powers to understand complex hyperspaces describing n-variables.

The problem of whether to apply descriptive models to a problem, models such as statistical and machine learning models, or to develop theories that explain physical phenomena seems to become an either-or problem. Much like the houses of the Montagues and Capulets in Shakespeare, the tools of Machine Learning and Phenomenological Relationships seem irreconcilable. Yet the problem lies not with either approach (both alike in dignity), but rather stems from very human limitations. We have great difficulty interpreting problems in higher-order space. It is easy to understand problems that have a single independent variable against which a phenomenon depends, resulting in y-x data plots. It is not much harder to understand problems which have two independent variables, as these can be represented cleanly in a three-dimensional Cartesian space. With some small imagination, we can even visualize and explore problems with three independent variables (say x, y, and time) against which a phenomenon depends, . Beyond this, we generally lack tools to explore an n-dimensional hyperspace.

It was the recognition that these expressions, from the simplest 1D function to the most complex neural network, can be written in multiple mathematically equivalent forms that germinated the idea that two equivalent descriptions of the same phenomenon—one an analytical formulation that captures the governing physics, the other a descriptive model that captures the correlation—may be written in a common form. The next logical step is to postulate that these common forms can be compared, and assessed for their self-similarity, providing a *translation* between one form and another. A Rosetta Stone, if you will, where the vocabularies (the mathematical basis sets that support each function space) of each are translated into a common vocabulary.

The precursors to this thought process began when H.L. Fraser et al. began conducting so-called virtual experiments, which permitted a well-trained neural network to be probed to determine the influence of one variable on a physical property while all of the other potential independent variables were artificially held at their average values [refs]. These virtual experiments were slices through an n-dimensional hyperspace, and it became obvious that the lower dimension slices could be described using simpler functions than the full neural network. This early work was built upon the initial efforts by H.K.D.H. Bhadeshia to solve complex materials science problems using artificial neural networks [refs] and the work of D.J.C. MacKay [refs] to incorporate Bayesian statistics and feedback loops. Following this prior work, with the assumption that a mapping was possible between the basis vectors of the n-variable hyperspace equally represented by the expansion of terms of an artificial neural network and the basis vectors representing the physical processes, Ghamarian and Collins, seeking to establish a constitutive equation for the room temperature yield strength of a titanium alloy given variations in its compositional and microstructural states, applied a hybrid artificial neural network-genetic algorithm method that optimized the unknowns in a postulated physically-based equation, testing the optimized physically-based model against slices through hyperspace function representing the neural network model and the data [refs]. This latter effort was quite inefficient, but resulted in a physically-based equation that was demonstrated, in subsequent work, to be generalized for multiple different variations of processing and with compositional variations [refs].

This previous work strongly suggests that the hypothesis we propose below is valid, that a more fulsome mathematical treatment is warranted, and is the subject of this paper. ~~While we develop the approach based upon a materials science problem, we hope that the general applicability will be apparent, as it is easy to imagine this approach impacting a diverse range of disciplines, including: genetics, public health, biological sciences, earth sciences, information sciences (including signals analysis), physical sciences, applied variants thereof (medicine, environmental activities, engineering), space sciences, and economics.~~ We also include an appendix that contains the expansions of the activation functions and the generating functions of common functions that form the basis of some relevant physics. We recognize that this appendix is far from complete but hope that the logic presented will enable those interested in identifying and developing the additional functions required by specific applications.

**Outline of the rest of the paper.** We first discuss the flexibility of machine learning algorithms, and neural networks more specifically, to identify correlations between an input feature set, *X*, and an output response, *y*. We then discuss the ability of constitutive relationships to mathematically describe the underlying cause of the correlation between *X* and *y*. We then support the functional forms of the neural network and the constitutive relationship on a common basis through a Fourier Series. Finally, we rationalize the equivalency of these functional forms through the uniqueness theorem of Fourier Series. We conclude by presenting a derivation of a mathematical expression for the value at each node in each layer of a neural network as a function of the node values of the previously layer. Taken through the entire architecture of the neural network, this sequential formulation provides a Fourier Series expansion supported on the same basis as the Fourier Series expansion of the candidate constitutive relationship. Because both Series are supported on the same basis, the coefficients are equivalent, providing a generalized numerical solution for any candidate constitutive relationship through its functional equivalence with a trained neural network.

We hypothesize that the physics of an arbitrarily complex process can be extracted by fitting the series expansion of a known, postulated, or potential physical relationship to the coefficients of the series expansion of a trained neural network when those series are supported on the same basis.

Neural networks have been proven capable of representing any arbitrary function [REF], therefore should a correlation exist, a sufficiently complex neural network can capture that correlation. And while not all correlations are causal relationships, all causal relationships are correlations. Therefore, while ML cannot generally expose the form of a constitutive relationship, it can, given sufficient data, accept or reject the null hypothesis: the hypothesis that no relationship exists between an input *X* and an output *y*. Therefore, an ML model that achieves sufficient accuracy indicates that a correlation does exist and that a constitutive relationship may exist.

Functional forms of constitutive relationships have been discovered over centuries to describe the mechanistic relationships between the influence of one-or-more internal or external fields and the change in state of the system. Onsager relations [REF], for example, relate the time-rate-of-change, the flux, of a quantitative variable—heat, mass, charge—to the spatial gradient in a qualitative variable—temperature, concentration, and electric field, respectively. The constant of proportionality between these quantitative and qualitative variables is an intrinsic materials property: the thermal conductivity, diffusivity, or conductivity. Identifying these material properties are one of the fundamental goals of materials science. These properties are the cumulative effect of composition, of processing, and of thermal history; that is, they are the time-and-path dependent effect of thermal, chemical, mechanical, electrical and magnetic stresses on material state.

Under the supposition that some constitutive relationship does exist, then both the neural network and the constitutive form both describe the same functional relationship.

Just as a neural network can represent any function, so too can a Fourier Series [REF]. Therefore, both the neural network and each candidate constitutive relationship can be represented to arbitrary precision by their respective Fourier Series. Since both the neural network (NN) and the constitutive relationship (CR) describe the same functional relationship,

Because the trigonometric basis supporting Fourier Series are orthogonal, two Fourier Series are equivalent if and only if their coefficients are equivalent, and if their coefficients are equivalent, then the Series are identical. And by the uniqueness theorem,[[1]](#footnote-2) the Fourier Series expansion of a function is unique. Therefore, if the Fourier Series of two functional representations are equivalent and unique, then these two functional representations describe the same function.

Fourier Series, being supported on a basis of periodic functions, is itself a periodic function. Therefore, a Fourier Series expansion of an aperiodic function is only valid over the integral domain used to determine the Fourier coefficients.

PETE STOP HERE

This assumes that the function space symmetry is triclinic. This is a strong assumption. A circle is the set of all points . Is there, analogously, a set of functions ? It would seem to follow from the uniqueness theorem for Fourier Series (and therefore, this section must follow the proof/reference/statement of the Fourier uniqueness theorem) that function space is triclinic—e.g., there is only one equation for a circle, otherwise there would be symmetrically equivalent functions that had the same properties as the circle function.

Upon further reflection, symmetry is not a fundamental property of the space, it is a property of a surface/volume in that space that define an object. What is a surface in function space? This is not relevant for this paper, but is an interesting question.

Coefficient generating functions, such as that shown in Equation [[eqn:coefficient generating function example]](#X366edc45ce93eae3d0f666950d2ca016bd7276a), have been tabulated for a number of common mathematical expressions, c.f. Table [[tab:generating functions of common functions]](#Xdf9c3f2924ecfe36eb050c3964517a75d688477). For constitutive relationships composed of a superposition of these terms, the series expansion coefficients become an analytical expression in terms of the physical constants of the constitutive relationships: the Hall-Petch coefficient, ; the Young’s modulus, ; etc. In general, we represent these coefficients as throughtout this work. Any physical constants known a priori reduce the number of in the final coefficient generating function expression.

**Interpretation**

We can use this method for deriving power series for repeatedly composed functions with linear transformations (i.e. ). We represent our original input variable as so that we have coefficients with all other coefficients zero. Then, we update for each composition of the function to obtain for . This can be easily applied to neural networks.  
If a collection of for satisfy both of these constraints, then, the constant coefficients of may be pulled out and added to the sum of . In concise terms we have derived that

In order to make this method computationally feasible, we must truncate the power series at some precision. Say the maximum power we wish on any variable is , then, we simply replace instances of with . Thus, coefficients of order will have no error and our approximation will only have truncation error from dropping terms with order .

**Constraint Analysis**

The constraints form a system called a restricted partition which is a concept within number theory. Fel [REF] has found an explicit solution to the constraints found in the simplified ANN case. We can construct an algorithm that operates in time by iterating through each of the indices from to and checking if they satisfy the constraints. This warrants some future study to see if there are ways to reduce the operation time and find solutions to the system of constraints more quickly.

**Discussion**

Many non-trivial problems in materials science, and in science more broadly, are explained not through a single constitutive relationship, but through a combination of contributing physics.

The goal of this approach is to identify the coefficients of a hypothesized constitutive relationship, coefficients that capture the specific physics of a process (through a least-squares fit between the covector space of a neural network series expansion, , which is a function of the model parameters) and the covector space of the constitutive relationship, , which is a function of the physical constants of the model.

It may be useful to introduce a schematic that not only frames the hypothesis, but also graphically registers key details of the methods section of this paper (Figure [2](#fig:nn-2)).

Schematic of the logic of the approach. The schematic could easily be expanded to show additional details and complexity on both the ANN/ML pathway and the potential physics pathway, including, for example, data selectors, physical dependencies. Similarly, the permutation of multiple architectures and postulated physics could be graphically represented, which would better represent the practical use of this approach.

Schematic of the logic of the approach. The schematic could easily be expanded to show additional details and complexity on both the ANN/ML pathway and the potential physics pathway, including, for example, data selectors, physical dependencies. Similarly, the permutation of multiple architectures and postulated physics could be graphically represented, which would better represent the practical use of this approach.

Having fit the model parameters, , on a vector space spanned by the column vectors of , the coefficients (the covector basis) of the neural network expansion, , capture the functional relationship between the input space and the response space, both affine and non-linear contributions, introduced through those parameters, , and the coefficient generating functions for the activation, e.g. Equations [[eqn:ReLU generating function]](#eqn:ReLU generating function) (Rectified Linear Unit, ReLU) and [[eqn:softmax generating function]](#eqn:softmax generating function) (softmax), respectively (see Appendix).

Naturally, the activation generating functions must match the activation function chosen in the neural network model architecture. Equation [[eqn:ReLU generating function]](#eqn:ReLU generating function) is derived for ReLU activation, the most common hidden layer activation. (Generating functions for other activations are provided in the appendix.) In addition to the hidden layers, activation functions must also be chosen for the output layer. The two most common output activations are linear (Eq. [[eqn:linear generating function]](#eqn:linear generating function)) and softmax (Eq. [[eqn:softmax generating function]](#eqn:softmax generating function)) for regression and classification, respectively.

Both the neural network and the constitutive relationship must depend on the same independent variables. That is, they must be described on the same basis vectors. The fit between the coefficients of the neural network expansion–, the covector space of the neural network’s basis vectors–and the coefficients of the series expansion of the constitutive relationship (its covector space, ) is only possible because both span the same subspace and share a common description of the solution within that subspace, that is, the covector spaces are colinear. That is, suppose that maps between a vector space and its covector space. If and , then if and only if .

One example of the proposed method is performed for yield strength, which is known to depend on intrinsic flow stress, as well as solute and grain size contributions, where the mapping of the polynomial expansions for both the neural network and potential physics :

$$\sigma\_y = \sigma\_f + \sum\_{k=0}^\infty \sum\_{l\_1=0}^k \binom{k}{l\_1, l\_2} \left( a\_k \prod\_{m=1}^2 ({\bf S}\_\text{solute})\_m^{l\_m} + b\_k \prod\_{m=1}^2 ({\bf S}\_\text{Hall-Petch})\_m^{l\_m} \right) \prod\_{m=1}^2{\bf x}\_m^{l\_m}$$

where

|  |  |
| --- | --- |
|  | matrix flow stress |
|  | coefficient generating function for |
|  | coefficient generating function for |
| ${\bf S}\_\text{solute}$ | selection vector/matrix for choosing the solute concentration from the input vector |
| ${\bf S}\_\text{Hall-Petch}$ | selection vector/matrix for choosing the grain size from the input vector |
| ${\bf x} = ([c], d)$ | concentration of a solute and grain size, respectively |

It should be noted here that this common basis is a function of the length (dimension) of the input vector and of the order of the expansion. Therefore, two series will share the same basis vector and the same covector space if and only if they are taken to the same order. In addition, because there is no guarantee that the input vector space directions are orthogonal, there is no guarantee that the cross-term interactions will vanish and, therefore, must be included explicitly. This expansion includes all cross-terms and, through the element-wise exponentiation, also explicitly captures all combinations of powers of all cross-terms.

TBW. This must answer the question: how do we know if we’ve measured the right things–not the number of measurements, but that we have enough information? How do we know that the solution has converged? Example: A model is to be fit to the number of cakes produced by a bakery. If we are given weights of flour and sugar and number of eggs, our model can accurately tell us the *volume* of cakes produced, but not the number. If this bakery makes cupcakes, but the model is trained across a spectrum of bakers, such as purveyors of wedding cakes and catering companies who work with large sheet cakes, then our dimensions (flour, sugar, eggs) are insufficient to fit the number of cakes produced. If, however, we also include number of orders and revenue, some information about the *quanta* of cakes is baked into those two additional dimensions (sorry, I couldn’t help myself). Therefore, a model based only on (flour, sugar, eggs) is dimensionally insufficient, but a model based on (flour, sugar, eggs, order size, revenue) is dimensionally sufficient.

This is a broader question that may be beyond the scope of this paper. Let’s return to this if, after completing the first pass, we feel that this can be addressed by what we’ve done.

The initial effort was made using the polynomial series exansions to obtain a power series coefficient generating function. However, a finite radius of convergence results in a divergence in the loss function making this approach unstable. The authors are looking to leverage the infinite radius of convergence of the Fourier Series and others to expand the neural network. The current plan is to progress from the neural network to the Fourier Series expansion, which could potentially be converted to a polynomial series making it easier to map directly to the constiutive relationships. Some constiutive relationships, such as 2 and 3-point correlation functions might better fit the Fourier Series anyway.

**Conclusions**

A generalized mathematical framework for proposing a constitutive relationship and fitting the physical constants associated with that constitutive relationship to a data corpus using the generalized fitting framework provided by machine learning/artificial intelligence has been presented. The proposed method maps between an arbitrary constitutive relationship and an artificial neural network model. The resulting covector spaces (coefficients) of the series are colinear. The generating function for the constitutive relationship is in terms of physical constants while that of neural network is determined through the trained model parameters. The method of least-squares is used to fit the physical constants to the trained neural network model parameters through this colinear covector space.

A simplified yield strength model, which includes flow stress, solute concentration, and Hall-Petch strengthening, is provided as an example to demonstrate the more general form for constructing the vector space for these two models.

Neural network series expansions are constructed layer-wise, which allows this framework to handle arbitrarily complex network architectures, including drop out, whitening, and any element-wise activation function. Mathematical descriptions of the rectified linear unit (ReLU) activation commonly used in neural network hidden layers, linear activation for regression networks, and softmax activation for classification networks are derived.

A mathematical framework to create constitutive relationship series expansions is also provided. Select coefficient generating functions for functional forms commonly found in materials physics are presented in the Appendix (Table [[tab:generating functions of common functions]](#Xdf9c3f2924ecfe36eb050c3964517a75d688477)) to serve as a starting point, but any other functions that can be described by a polynomial series may be used.

As a single framework capable of describing arbitrarily complex relationships, this approach is intended to facilitate fits between existing data and any hypothesized constitutive relationships built upon the same vector space as the trained neural network model.

It is critically important to recognize that while we are describing the translation between two reference frames, that of Machine Learning and that of Constitutive Laws of certain physical phenomena, there is no limitation to the number of reference frames (languages) among and between which the relationships can be established...which provides a way for different disciplines, describing the same phenonmena using different expressions, a way to translate their work - EXPAND ... different data, equipment, labs, ...

References and Notes:

1. There is only one reference list including all references in the text, Figure and Table captions, and Supplementary Material. Do not include a second reference list in the Supplementary Material. References only cited in the Supplementary Material are not counted toward length guidelines.
2. Each reference should be on a separate line ending in a period. For a style guide, see [http://www.sciencemag.org/authors/instructions-preparing-initial-manuscript](http://www.sciencemag.org/authors/instructions-preparing-initial-manuscript%20).
3. You should include titles in references and full page ranges. Titles will not be included in the print version of the paper, but will be shown in the online version.
4. Please include the above heading, “References and Notes:”.
5. You can use a numbered list in Word.
6. Each reference should have a separate number.
7. Please do not mix in references with explanatory notes.

**Acknowledgments:** Acknowledgments follow the references and notes but are not numbered. Acknowledgments should be gathered into a paragraph after the final numbered reference. This section should start by acknowledging non-author contributions and then should provide information under the following headings: **Funding:** include complete funding information, including grant numbers; **Author contributions:** a complete list of contributions to the paper [we encourage you to follow the [CRediT](http://docs.casrai.org/CRediT) model]; **Competing interests:** competing interests (including but not limited to patents, financial holdings, professional affiliations, advisory positions, and board memberships) of any of the authors must be listed (all authors must also fill out a separate, internal Conflict of Interest form). Where authors have no competing interests, this should also be declared (e.g., “Authors declare no competing interests.”); and **Data and materials availability:** Include a note explaining any restrictions on materials, such as materials transfer agreements. Note accession numbers to any data relating to the paper and deposited in a public database; include a brief description of the data set or model with the number. If all data are in the paper and supplementary materials include the sentence “All data is available in the main text or the supplementary materials.” All data, code, and materials used in the analysis must be available in some form to any researcher for purposes of reproducing or extending the analysis.

Supplementary Materials:

Materials and Methods

Figures S1-S#

Tables S1-S#

Movies S1-S#

Audio Files S1-S#

External Databases S1-S#

References (*##-##*)

**Materials and Methods**

Fully dense neural network (NN) architectures, such as the one shown in Figure [1](#fig:nn-1), perform a sequence of affine transformations, ${\bold z}\_i \leftarrow \boldsymbol\theta\_i {\bold x}^{(i)}$, followed by element-wise functional operations, $\sigma({\bold z}\_i)$ to introduce non-linearity at each layer; that is, each layer stretches and distorts the underlying space.

Schematic view of a fully dense neural network. Each sequence of affine and non-linear transformations are captured in the function, $f\_i({\bold x}): {\bold x}^{(i+1)} \leftarrow \sigma(\boldsymbol\theta\_i {\bold x}^{(i)})$

Schematic view of a fully dense neural network. Each sequence of affine and non-linear transformations are captured in the function, $f\_i({\bold x}): {\bold x}^{(i+1)} \leftarrow \sigma(\boldsymbol\theta\_i {\bold x}^{(i)})$

A typical neural network, as given in (insert neural network reference), is nothing more than an arbitrary function generator, but at present, the network weights can not map back to analytic forms that capture and describe the underlying physics. There are, however, many such mappings through polynomial series expansions (possible reference to other polynomial series expansion applications?). We hypothesize that the physics of a process can be extracted by fitting the polynomial expansions of known physical relationships to the polynomial coefficients of a polynomial series expansion.  
The first step is to obtain an activation generating function for the chosen activation function in the neural network. There exists many potential options for choosing an activation function such as softplus, softmax, ReLU (rectified linear units), and logistic sigmoid. The details of the derivations for the generating functions of these common activation functions can be found in the Appendix. The generating function for the ReLU is dependent on both the input variables and the network weights. Although ReLU (rectified linear units) have become a more common activation function, its discontinuity at requires an infinite series to fully capture the behavior at this transition. Note: add more pros and cons about each of the common activation functions if necessary

**Iterative Determination of ANN Series Expansion Coefficients**

Any series expansion description of a deep neural network, which is necessarily multilayered, requires a explicit generating function for the coefficients of each layer; a generating function that is a function only of the coefficients of the previous layer and the coefficient generating function of the series expansion of the current layer’s activation function.

A derivation for the polynomial expansion coefficient generating function for vector-valued function (layer 1 to layer 2) is presented below, which is an extension of a scalar expansion, provided in [[appendix]](#appendix).

**Polynomial Expansion of a Vector Layer**

We can represent the recursive structure of an ANN as a series of transformations on a particular power series. Using this method, we are able to write an algorithm that will compute the coefficients of a power series determined by some ANN up to an arbitrary order of approximation.  
Here, we give a derivation of the generalized method applicable to any feed-forward ANN. We present a special case of this derivation where each layer of the ANN has a single neuron in Section [[sec:polynomial series scalar]](#sec:polynomial series scalar).

For the sake of illustration, we present the next derivation in terms of a simplified ANN with one neuron in each layer. For the treatment of a unsimplified ANN, see Appendix.

**Notation**

Suppose that is a vector, is a vector, is a matrix, and is an analytic function. Since is analytic, it can be represented as

Suppose that we can represent each entry of as a power series of the entries of a vector , where the coefficients are .

Finally, suppose that we have the relation

where is being applied element-wise to a vector.  
Similar to the scalar case, this represents a multi-neuron layer of an ANN with input , weights , output , and activation . The value of can be interpreted as the input to the entire ANN.

**Objective**

Similarly, just as each entry of can be represented as a power series, the same can be done for , where the coefficients are . Rewriting Equation ([[eqn:vector layer relation]](#eqn:vector layer relation)) in terms of Equations ([[eqn:vector layer activation]](#eqn:vector layer activation)) and ([[eqn:vector x series]](#eqn:vector x series)), we obtain

where is a collection of non-negative integers that are indices such that

**Coefficient Extraction**

To find the coefficients from Equation ([[eqn:vector y expansion]](#eqn:vector y expansion)), we must find terms satisfying index constraints and power constraints. Similar to a scalar layer expansion, the power constraints can be simplified as given in the Appendix. If a collection of and satisfy all of these constraints, then the value

may be pulled out to the sum of . In concise terms we have derived

The only issue is that of finding solutions to the index and power constraints.

Suppose that is a vector, is a vector, is a matrix, and is an analytic function. Since is analytic, it can be represented as

Suppose that we can represent each entry of as a power series of the entries of a vector .

Finally, suppose that we have the relation

where is being applied element-wise to a vector.  
Similar to the scalar case, this represents a multi-neuron layer of an ANN with input , weights , output , and activation . The value of can be interpreted as the input to the entire ANN.

**Objective**

We wish to represent each entry of as a power series of the entries of . That is,

Rewriting Equation ([[eqn:vector layer relation]](#eqn:vector layer relation)) in terms of Equations ([[eqn:vector layer activation]](#eqn:vector layer activation)) and ([[eqn:vector x series]](#eqn:vector x series)), we obtain

where is a collection of non-negative integers that are indices such that

**Coefficient Extraction**

To find the coefficients from Equation ([[eqn:vector y expansion]](#eqn:vector y expansion)), we must find terms satisfying index constraints

and power constraints

Similar to the scalar case, notice that the power constraints can be simplified since any solution must also satisfy

resulting in

If a collection of and satisfy all of these constraints, then the value

may be pulled out to the sum of . In concise terms we have derived

At this point, it is infeasible to attempt to write Equation ([[eqn:scalar y expansion]](#eqn:scalar y expansion)) directly as a power series. Instead, we will extract the coefficients for by observing constraints on the coefficients.  
First, as a result of the multinomial theorem, we have a constraint on the inner sum in Equation ([[eqn:scalar y expansion]](#eqn:scalar y expansion)), which is constraint on the index of the sum so we shall call this an index constraint.  
Second, the coefficient is associated with the term so we must form an equality between and the power on in Equation ([[eqn:scalar y expansion]](#eqn:scalar y expansion)). To satisfy this equality, a constant is required. This is a constraint on the power of the scalar so we shall call this a power constraint. This constraint implies . Therefore, both the index constraint and power constraint can be reduced to a finite series instead of an infinite series.

The only issue is that of finding solutions to the index and power constraints.

**Polynomial Expansion of a Vector Layer**

This derivation will be a generalization of the derivation given in Section [[sec:polynomial series scalar]](#sec:polynomial series scalar) where each layer of the ANN may have an arbitrary number of neurons. We show a similar result to before where we find coefficients for a power series after linear and nonlinear transformations.

**Notation**

Suppose that is a vector, is a vector, is a matrix, and is an analytic function. Since is analytic, it can be represented as

Suppose that we can represent each entry of as a power series of the entries of a vector .

Finally, suppose that we have the relation

where is being applied element-wise to a vector.  
Similar to the scalar case, this represents a multi-neuron layer of an ANN with input , weights , output , and activation . The value of can be intrepreted as the input to the entire ANN.

**Objective**

We wish to represent each entry of as a power series of the entries of . That is,

Rewriting Equation ([[eqn:vector layer relation]](#eqn:vector layer relation)) in terms of Equations ([[eqn:vector layer activation]](#eqn:vector layer activation)) and ([[eqn:vector x series]](#eqn:vector x series)), we obtain

where is a collection of non-negative integers that are indices such that

**Coefficient Extraction**

To find the coefficients from Equation ([[eqn:vector y expansion]](#eqn:vector y expansion)), we must find terms satisfying index constraints

and power constraints

Similar to the scalar case, notice that the power constraints can be simplified since any solution must also satisfy

resulting in

If a collection of and satisfy all of these constraints, then the value

may be pulled out to the sum of . In concise terms we have derived

The only issue is that of finding solutions to the index and power constraints.

**Softmax Activation**

For example, consider the Softmax activation function which is given in Equation (???)

$$\sigma(\bold z)\_i = \frac{e^{z\_i}}{\sum\limits\_{j=1}^k e^{z\_j}}$$

The goal is to obtain a activation generating function for the chosen activation function in the neural network.

The partion function, $\bold Z$, is defined below as

$$\bold Z = \sum\limits\_{j=1}^k e^{z\_j}$$

The series representation for the exponential is given by the following

Therefore

$$\sigma(\bold z)\_i = \sum\limits\_{k=0}^{\infty} \alpha\_{k}z\_{i}^{k}$$

Where $\alpha\_k = \frac{1}{{\bold Z}k!}$ is the generating function for the Softmax

**ReLU Activation**

The ReLU (rectified linear units) function is another commonly used activation function and is given below in Equation ([[eqn:ReLU]](#eqn:ReLU)):

The generating function for the ReLU is given in Equation ([[eqn: ReLU generating function]](#eqn: ReLU generating function)) and it is dependent on both the input variables and the network weights.

$$\alpha\_n = \begin{cases} 1 & \bold z > 0 \\ 0 & \text{otherwise} \end{cases} \label{eqn: ReLU generating function}$$

Similarly, the linear generating function is given in Equation ([[eqn: linear generating function]](#eqn: linear generating function)) but it is only dependent upon $\bold n$.

$$\alpha\_n = \begin{cases} 1 & \bold n = 1 \\ 0 & \text{otherwise} \end{cases} \label{eqn: linear generating function}$$

Although ReLU (rectified linear units) have become a more common activation function, its discontinuity at requires an infinite series to fully capture the behavior at this transition.

From Gradshteyn and Ryzhik - 0.314 Power series raised to powers.

$$\left(\sum\limits\_{k=0}^{\infty}a\_{k}x^{k}\right)^{n} = \sum\limits\_{k=0}^{\infty}c\_{k}x^{k} \\ \label{eqn:power series raised to powers}$$

where

And therefore it can be seen that

From Gradshteyn and Ryzhik - 1.1 Power of Binomials, 1.111 Power Series

And similarly

Once again the substitution for a power series raised to a power is applied.

After some rearranging of term, the result is given below.

And therefore the following is obtained.

Where

From Gradshteyn and Ryzhik - 0.313 Division of power series.

where

And therefore

where is the generating function for the softplus

At this point, it is infeasible to attempt to write Equation ([[eqn:scalar y expansion]](#eqn:scalar y expansion)) directly as a power series of . Instead, we will extract the coefficients for by observing constraints on the coefficients.  
First, as a result of the multinomial theorem, we have a constraint on the inner sum in Equation ([[eqn:scalar y expansion]](#eqn:scalar y expansion)) that is

where each must be a non-negative integer. This is a constraint on the index of the sum so we shall call this an index constraint.  
Second, the coefficient is associated with the term so we must form an equality between and the power on in Equation ([[eqn:scalar y expansion]](#eqn:scalar y expansion)). To satisfy this equality, we note that

where is some constant. This implies that

This is a constraint on the power of the scalar so we shall call this a power constraint.  
Notice that the power constraint implies . Therefore, both the index constraint and power constraint can be reduced to a finite series instead of an infinite series.

If a collection of for satisfy both of these constraints, then, the constant coefficients of which are

may be pulled out and added to the sum of . In concise terms we have derived that

Thus, we have derived a tractable method of computing the coefficients of a power series after a linear and nonlinear transformation. We have included some worked examples of these coefficients in Appendix [6.6](#sec:worked scalar coeffs).

ReLU activation,

is discontinuous in the first derivative at . Therefore, the coefficient generating function of this activation must either be either a function of the input data, ${\bf z}$ or a small modification must be made to the softplus,

In the limit as approaches infinity, this converges to the ReLU. Practically, though, can be assigned a large value and the coefficient generating function no longer depends on the input data, see Equation [[eqn:modified softplus generating function]](#X930ec476a78273d6dfba12a791f73f83148c5ce) in the appendix. However, because of the high computational cost of expressing the coefficients using the modified softplus, and the relative low cost of forward evaluation of the trained neural network in order to apply Equation [[eqn:ReLU generating function]](#eqn:ReLU generating function), expressing the coefficient generating function of ReLU in terms of the input data, as in Equation [[eqn:ReLU generating function]](#eqn:ReLU generating function), rather than the training-data-agnostic approach in Equation [[eqn:modified softplus generating function]](#X930ec476a78273d6dfba12a791f73f83148c5ce) would seem more practical.

If ${\bf x} = ([c], d)$ are the concentration of a solute and grain size, respectively, then the constitutive relationship expansion would be,

$$\begin{aligned} \sigma\_y &= \sigma\_f + C [c]^{2/3} + \frac{k}{\sqrt{d}} \nonumber \\ &= \sigma\_f + \underbrace{\sum\_{k=0}^{\infty} a\_k [c]^k}\_{\text{solute}} + \underbrace{\sum\_{k=0}^{\infty} b\_k d^k}\_{\text{Hall-Petch}} \nonumber \\ &= \sigma\_f + \sum\_{k=0}^{\infty} a\_k \ensuremath{{k}^{\circ +}} \sum\_{k=0}^{\infty} b\_k \ensuremath{{k}^{\circ \label}}{eqn:simplified strengthening}\end{aligned}$$

where

|  |  |
| --- | --- |
|  | Matrix flow stress. |
|  | Coefficient generating function for . |
|  | Coefficient generating function for . |
| ${\bf S}\_\text{solute}$ | Selection vector/matrix for choosing the solute concentration from the input vector. |
| ${\bf S}\_\text{Hall-Petch}$ | Selection vector/matrix for choosing the grain size from the input vector. |

(Coefficient generating functions can be found in Table [[tab:generating functions of common functions]](#Xdf9c3f2924ecfe36eb050c3964517a75d688477).) Combining Equations [[eqn:simplified strengthening]](#eqn:simplified strengthening) and [[eqn:hadamard exponent vector]](#eqn:hadamard exponent vector),

$$\begin{gathered} \sigma\_y = \sigma\_f + \sum\_{k=0}^\infty a\_k \left[ \sum\_{l\_1=0}^k \binom{k}{l\_1, l\_2} \prod\_{m=1}^2 (({\bf S}\_\text{solute})\_m {\bf x}\_m)^{l\_m} \right] \\ + \sum\_{k=0}^\infty b\_k \left[ \sum\_{l\_1=0}^k \binom{k}{l\_1, l\_2} \prod\_{m=1}^2 (({\bf S}\_\text{Hall-Petch})\_m {\bf x}\_m)^{l\_m} \right],\ l\_2 = k - l\_1\end{gathered}$$

which further simplifies to

$$\begin{gathered} \sigma\_y = \sigma\_f + \sum\_{k=0}^\infty \sum\_{l\_1=0}^k a\_k \binom{k}{l\_1, l\_2} \prod\_{m=1}^2 ({\bf S}\_\text{solute})\_m^{l\_m} \prod\_{m=1}^2{\bf x}\_m^{l\_m} \\ + \sum\_{k=0}^\infty \sum\_{l\_1=0}^k b\_k \binom{k}{l\_1, l\_2} \prod\_{m=1}^2 ({\bf S}\_\text{Hall-Petch})\_m^{l\_m} \prod\_{m=1}^2{\bf x}\_m^{l\_m}\end{gathered}$$

such that the term $\prod\_{m=1}^2{\bf x}\_m^{l\_m}$ serves as the common basis set over which the summation occurs, so that now, having a common basis, this simplifies to,

$$\sigma\_y = \sigma\_f + \sum\_{k=0}^\infty \sum\_{l\_1=0}^k \binom{k}{l\_1, l\_2} \left( a\_k \prod\_{m=1}^2 ({\bf S}\_\text{solute})\_m^{l\_m} + b\_k \prod\_{m=1}^2 ({\bf S}\_\text{Hall-Petch})\_m^{l\_m} \right) \prod\_{m=1}^2{\bf x}\_m^{l\_m}$$

Equation [[eqn:ANN power series coefficient generating function]](#Xd4fa64d6580317957bf7361ba0566ce8adadd06) shows that the polynomial series expansion for the first layer of a neural network,

similarly relies on the element-wise exponential, $(\bullet)\he{n}$, as does the expansion of all layers. Unlike scalar exponentiation, element-wise exponentiation does not distribute, as seen in Equation [[eqn:nondistributive hadamard]](#eqn:nondistributive hadamard), and because element-wise exponentiation does not distribute, this equation explicitly captures all possible (second, ; third, ; fourth, ; etc.) cross-interactions of each term in $(\boldsymbol{\theta}{\bf x})$ at all polynomial orders. This is equivalent, then, to expanding over the basis set that includes all cross-interactions in the input vector space for both the constitutive relation and neural network polynomial expansions.

Data preprocessing is an important step in training a neural network to avoid implicit bias. Commonly, data is whitened, also known as scaling or standardization, ${\bf z}\_s^{(k)}: {\bf z}\_s^{(k)} = \frac{{\bf z}^{(k)} - \overline{{\bf z}^{(k)}}}{\sigma}$, where $\overline{{\bf z}^{(k)}}$ is the arithmetic mean and the standard deviation of data in layer, . However, a model trained on such scaled data would no longer share the vector space of the constitutive relationship. To ensure that both the neural network and constitutive relation expansions share a common vector space, and thus a common covector space, this whitening procedure must be integrated into the construction of the neural network architecture.

Procedurally, the neural network expansion proceeds as described in Section [3](#methods). The input to each layer, which is the output from the previous layer, is subjected first to an affine transformation, then to an activation function. The activation function is completely described though its polynomial expansion and the corresponding coefficient generating function. Using this same structure, then, data whitening can be applied to any layer, including the input layer, as an identity tranform ($\boldsymbol{\theta} = {\bf I}\_d$, where is the dimension of the source layer) followed by the whitening expansion whose coefficient generating function is simply,

$$\alpha\_k({\bf z}) = \begin{cases} -\overline{\bf z}/\sigma & \mbox{if } k = 0, \\ 1/\sigma & \mbox{if } k = 1, \\ 0 & \mbox{otherwise} \end{cases}$$

where

|  |  |
| --- | --- |
| $\overline{\bf z}$ | The mean of the data into the standarization layer. |
|  | The standard deviation of the data into the standardization layer. |

By introducing such a whitening layer, data standardization can be included at any point in the neural network architecture.

**Activation function****s: the logistic sigmoid**

is a special case of the generating function for the Euler polynomial coefficients,

where, for ,

The Euler polynomials at ,

where is the Bernoulli number. Since Bernoulli numbers of odd index, with the exception of , are zero, for . Therefore, the summand and limits of Equation ([[eqn:sigmoid Euler expansion]](#eqn:sigmoid Euler expansion)) change to

The series representation of

such that,

and therefore,

$$\begin{aligned} \sigma(x) & = & \frac{1}{2} - \sum\_{n=1}^\infty 2 \frac{(-1)^n}{\pi^{2n}} \left( \sum\_{k=0}^\infty \frac{1}{(2k+1)^{2n}} \right) x^{2n-1} \\ & = & \frac{1}{2} - \sum\_{n=1}^\infty 2 \frac{(-1)^n}{\pi^{2n}} \left( 4^{-n} \left( 4^n - 1 \right) \zeta(2n) \right) x^{2n-1} \nonumber \\ & = & \frac{1}{2} - \sum\_{n=1}^\infty \underbrace{2 \left( \frac{-1}{4\pi^2} \right)^n \left( 4^n - 1 \right) \zeta(2n)}\_{a\_n} x^{2n-1} \nonumber \\ & = & \sum\_{n=0}^\infty a\_n x^n,\ a\_n = \left\{ \begin{array}{l l} 1/2 & n = 0 \\ -2 \left( \frac{-1}{4\pi^2} \right)^{(n+1)/2} \left( 4^{(n+1)/2} - 1 \right)\zeta(n+1) & n\ \text{odd} \\ 0 & n\ \text{even} \end{array}\right. \label{eqn:sigmoid zeta expansion}\end{aligned}$$

**Coefficient Generating Functions for Common Functions**

Central to this approach is the connection ability to represent a constitutive relationship and a neural network on the same vector/covector space. This is done through a polynomial series expansion of both the neural network, covered in the text, and the constitutive relationship. Select generating functions are provided here.

c | c c c c c c k & & & & & &   
0 & & – & & & &   
1 & & – & & & &   
2 & & – & & & &   
⋮&  
n & & $\left\{\begin{array}{c l} C & \text{if}\ k = n \\ 0 & \text{otherwise} \end{array}\right.$ & & & &   
Constraint & & & & & &

[tab:generating functions of common functions]

**Coefficients of the Scalar Polynomial Series**

A result of the coefficient derivation of polynomial series is that we may construct formulae for the coefficients of an arbitrary function by simply solving for the constraints in Equation ([[eqn:scalar power series result]](#eqn:scalar power series result)). We provide a few instances of worked coefficients here.

We find . so . Therefore,

We find . so and . Further, . Therefore,

We find . For this coefficient, there are multiple solutions to the constraints.

* and either
  + and or
  + and

Therefore,

A close up of a map

Description automatically generated**Fig. 1.** Comparison of the exponential function (black) to power series approximations taken to 6th order. The inset indicates schematically the linear independence of the polynomial basis functions. These are used to approach , which exists as a point in the function space spanned by the basis functions of the series.

**Fig. 2.** You can place graphics in-line above each caption. Please do not use text boxes to arrange figures. All images should be JPEG, PNG, TIFF, or similar standard format. High resolution (preferably editable PDF or Adobe Illustrator format) figures will be requested following review.

**Table 1.** Start this caption with a short description of your table. Format tables using the Word Table commands and structures. Do not create tables using spaces or tab characters.

(Please delete before submission) Supplementary materials should be included in a separate supplementary materials file. A template for this file can be found at: <http://www.sciencemag.org/sites/default/files/Science_Supplementary_Materials_Word_template.docx>.

***Science* Manuscript Template Instructions**

**General Instructions on using this template and submitting a manuscript to *Science*:** Thank you for preparing a manuscript for submission to *Science*. Using this template, or following the guidelines below, will help us in processing your paper. Our goal is to be able to identify each section of your manuscript so that we can accurately record the title, authors, abstract, etc. and to enrich it by including reference links and an accurate layout.

Please use the actual template, which starts on page 2. When you are ready to submit, please delete the text on this cover page.

You can submit your paper at <https://cts.sciencemag.org>.

Additional information for authors is available at [http://www.sciencemag.org/authors/science-information-authors](http://www.sciencemag.org/site/feature/contribinfo/index.xhtml).

If you are using LaTeX, please convert your paper into a Word .docx file if possible. If this is not possible, please use our LaTeX template and upload a PDF version of your paper. Some conversion approaches are available here: <http://www.tug.org/utilities/texconv/textopc.html>.

So that we can identify the parts of your paper, and even if you do not use our template, please begin each section with the specific key words listed below, some of which are followed by a colon. Please do not use paragraph breaks in the title, author list, or abstract. The author list, corresponding author email(s), and affiliation(s) should be checked carefully because they will be published as listed in the manuscript.

**Title:** No more than 96 characters, lacking jargon and abbreviations where possible.

**Authors:**

**Affiliations:**

**One Sentence Summary:** No more than 40 characters.

**Abstract:** 125 words or less.

**Main Text:**

**References and Notes:** (Followed by a numbered list); only a single reference list should be provided for the main text and supplemental information.

**Acknowledgments:** Split into sections as described below.

**List of Supplementary Materials:** Include a list, noting which references are only cited in the SM.

**Fig. #.** (Begin each figure caption with a label, “**Fig. 1.**”, for example, as a new paragraph.)

**Table #.** (Begin each table caption with a label “**Table 1.**”, for example, as a new paragraph.)

**Supplementary Materials:** Comprising Materials and Methods, figures, and tables; should be in a separate file.

Please use the .docx format (all versions after Word 2007). If you chose not to use this template, please include page numbers in your submitted file. We also encourage use of line numbers.

More specific formatting instructions are provided in the template which follows.

1. Vagharshakyan, A. On the uniqueness problem for Fourier series. *Real Anal. Exch.* **29**, 939–946 (2004). [↑](#footnote-ref-2)