`nle-lepton` v4.2: Software to Find Polynomial-like Formulas for Fermion Masses

Kevin Loch*

(Dated: v4.0: 2020/02/19, v4.2: 2020/03/18)

`nle-lepton` is an software program that searches for polynomial-like non-linear equations with three real, positive roots representing the charged lepton masses. A formula of this type might explain why there are three generations of ordinary matter and give insight into the underlying physics of fermion Higgs field Yukawa couplings.

I. INTRODUCTION

Explanations for the structure of the fermion mass spectrum and why there are exactly three generations of ordinary matter remain elusive [1]. In the standard model charged fermion fields acquire mass via Yukawa couplings to the Higgs field after symmetry breaking [2–5]. However, the masses (or the corresponding Yukawa couplings) are free inputs that cannot yet be derived from theory [6]. To address these problems a software program `nle-lepton` [7] has been developed to search for polynomial-like non-linear equations (NLE's) with three positive, real roots. These roots represent the three charged lepton masses, or generically any three user-supplied "solution masses". This type of empirically derived formula might lead to a better understanding of these problems and provide insight into the fermion mass generation process.

Observed charged lepton masses are known to vary (run) with energy scale due to vacuum polarization effects [8]. Specific running masses depend on the chosen renormalization scheme and energy scale. By default `nle-lepton` uses the charged lepton rest masses but can be reconfigured to use other values in the configuration file `nle-lepton.cfg`. However, even the use of rest masses is not necessarily problematic. Some of the factors included in this search such as α_{em} and $\sin^2\theta_W$ also run with energy scale and might lead to an exact or approximate formula using rest masses. If the mass generation process is regulated by the on-shell effective mass, a useful formula might be found from rest masses. In any case, `nle-lepton` is designed to be as flexible as possible and efficiently find formulas that generate the user-supplied masses.

II. SEARCH SPACE AND PROCESSING METHOD

There are an infinite number of polynomial-like non-linear equations with three positive, real roots. Therefore several choices have been be made to limit the search to specific areas of interest.

- The solution masses are paired with a reference mass and other dimensionless factors in a solution mass ratio (smr) to make each term dimensionless. Adding terms with different mass dimensions does not make sense physically. Supported reference masses are v, m_{P}, m_Z, m_W, m_{H^0}, and user-defined m_{user}.
- Exponents are limited to fractional values $\frac{1}{n}$ selected from the range $n = \{-26...-1, 1...26\}$. Proper polynomials have only positive integer exponents. Replicating the charged lepton mass spectrum with a polynomial would require large coefficients on some terms and a large variance between coefficients. Using fractional exponents on each term leads to coefficients that are closer to one and similar in magnitude. This would be expected if coefficient factors outside the radicals were similar or identical between terms. One possible physical explanation for fractional exponents would be if each term represented a length scale derived from a higher dimensional manifold, such as the radius of an n-ball or n-sphere.
- Initial development focused on a "3-term" structure with three mass terms and a constant term of -1. This is the minimum number of terms that can generate three positive, real roots.
- Version 4.1 added a "2-term" mode where two mass terms are mixed to generate a synthetic third mass term.
- Version 4.2 added support for a "1-minus" feature in 2-term mode where one term uses a normal solution mass ratio (smr) while the other term uses $(1 - \text{smr})$. Since $(1 - \text{smr})$ can sometimes be negative there is a choice of

* kevinmloch@gmail.com

polarity (sign) of the synthetically generated term. Version 4.2 uses + polarity so $(1 - \text{smr})$ must be negative for at least one solution mass. Negative polarity could be explored in future versions so that $(1 - \text{smr})$ could always be positive.

- Additional formula structures can be implemented by editing the source code, using the existing modes as a template.

Processing is broken into two main phases. In Phase 1 a minimal formula structure with random exponents is generated and solved for real coefficients. In 3-term mode the minimal formula is

$$C_1 \left(\frac{M}{v}\right)^{\frac{1}{n_1}} - C_2 \left(\frac{M}{v}\right)^{\frac{1}{n_2}} + C_3 \left(\frac{M}{v}\right)^{\frac{1}{n_3}} - 1 = 0, \tag{1}$$

where solving for M gives the three solution masses, C_i are real coefficients, $\frac{1}{n_i}$ are exponents selected from the range $n = \{-26...-1, 1...26\}$, $n_1 < n_2 < n_3$, and the Higgs vacuum expectation value v is temporarily used as the reference mass for each solution mass ratio. The real coefficients C_i are then sent to a factoring engine that tries all possible factor combinations defined in `nle-lepton.cfg`, including substituting v for other reference masses. When the factors are a close enough match an exact formula is constructed from the implied factors and solved for the solution masses and/or other variables in Phase 2. The variables are then compared to their experimental uncertainties and the result is output if they match. This multi-step process (solve for coefficients, factor, solve for masses) allows for billions of coefficient factors and trillions of coefficient, exponent and mass ratio combinations to be tested efficiently.

III. PHASE 1

Each time Phase 1 is run exponents are randomly selected and assigned to the mass terms of the minimal equation 1 in the correct order to generate three real roots. The values of the solution masses are also randomly selected from within their experimental uncertainties unless `phase1_random_samples_enable` is set to `no` in `nle-lepton.cfg`. The function `solveNLEforCoefficients()` then solves equation 1 for coefficients C_i using computational (MC) methods. During this process three copies of equation 1 are used, `sm1_test`, `sm2_test`, `sm3_test`, one for each solution mass,

$$\texttt{sm1\_test} = C_1 \left(\frac{M_{\text{sm1}}}{v}\right)^{\frac{1}{n_1}} - C_2 \left(\frac{M_{\text{sm1}}}{v}\right)^{\frac{1}{n_2}} + C_3 \left(\frac{M_{\text{sm1}}}{v}\right)^{\frac{1}{n_3}} - 1, \tag{2}$$

$$\texttt{sm2\_test} = C_1 \left(\frac{M_{\text{sm2}}}{v}\right)^{\frac{1}{n_1}} - C_2 \left(\frac{M_{\text{sm2}}}{v}\right)^{\frac{1}{n_2}} + C_3 \left(\frac{M_{\text{sm2}}}{v}\right)^{\frac{1}{n_3}} - 1, \tag{3}$$

$$\texttt{sm3\_test} = C_1 \left(\frac{M_{\text{sm3}}}{v}\right)^{\frac{1}{n_1}} - C_2 \left(\frac{M_{\text{sm3}}}{v}\right)^{\frac{1}{n_2}} + C_3 \left(\frac{M_{\text{sm3}}}{v}\right)^{\frac{1}{n_3}} - 1. \tag{4}$$

C_i are set to an initial value of 1.0 and processing successively refines C_i until each copy of the formula converges towards zero and

$$\texttt{precision} = |\texttt{sm1\_test}| + |\texttt{sm2\_test}| + |\texttt{sm3\_test}| < 1.0 \times 10^{-16}, \tag{5}$$

which is more than sufficient to give useful results and safely within the limits of 80-bit extended precision `long double` addition on x86 platforms. For simplicity and speed the Higgs vacuum expectation value v is temporarily used as the reference mass in all three terms. Other reference masses will be substituted in the next step.

IV. FACTORING COEFFICIENTS AND ASSEMBLING FORMULAS FOR PHASE 2

Once the real coefficients C_i are found they are sent to function `cscanner()` which substitutes v for other reference masses (so that all intended mass ratio combinations are tested) and searches for interesting factors for the coefficients. Possible factor combinations are sourced from "static" ingredients with exact or low uncertainty values, and "dynamic" ingredients with significant experimental uncertainty like $\sin^2\theta_W$. Two pre-computed tables of static ingredient combinations are used to accelerate this process, one for inside the radical (`infactors`) and one for outside the radical (`outfactors`). Values for dynamic ingredients are selected at random from within their experimental uncertainties each time phase 1 is run. Filtering of which coefficient factor combinations are considered

a match is controlled by `phase1_filter` and `phase1_int_match_max` in `nle-lepton.cfg`. A match occurs when potential_factor_combination$*C_i$ is within $1.0\times10^{-\text{phase1\_filter}}$ of an integer between 1 and `phase1_int_match_max` inclusive. This factoring process allows for rapid processing of billions of possible factor and reference mass combinations without having to solve equation 1 for each combination.

Up to this point factoring of the coefficients has been done on each term independently of the others. When these terms are combined into complete formulas in phase 2 each unique combination of potential factors for each term needs to be tried and these combinations are assembled in function `verifyMatches()`. This often results in thousands of formulas to be processed from a single run of phase 1. To avoid processing uninteresting formulas, symmetry and complexity scores are assigned to each combination of terms with higher symmetry and lower complexity generally meaning a simpler and more interesting formula. The configuration options `phase2_symmetry_min` and `phase2_complexity_max` are provided to filter the formulas allowed to be passed to phase 2. Restricting these values can greatly speed up the search process and reduce the number of uninteresting results at the expense of possibly missing a more complex but correct formula.

V. PHASE 2 AND SAMPLE 3-TERM RESULT

In phase 2 processing is reversed. Formulas are constructed with the factors found in phase 1 and solved for the three solution masses and/or other outputs in function `solveNLEforMasses()` using computational methods similar to phase 1. Before solving each proposed formula, the variables with experimental uncertainty are ranked by relative standard uncertainty and the three with the highest uncertainty are used as outputs (solved for) with the rest used as inputs. This allows for the lowest possible relative uncertainty in the outputs. Only results with all outputs within `phase2_results_window` (default=1.1) of experimental uncertainty are shown unless `phase2_results_always` is set to `yes`.

An example of a result found in 3-term mode is

$$\frac{6\sqrt{2}}{\pi\cos^2\theta_{\mathrm{W}}}\left(\mathrm{R_s}\left(11\right)\frac{3}{\sqrt{2}}\frac{m_W}{M}\right)^{\frac{1}{11}}-\frac{6\sqrt{2}}{\pi\cos^2\theta_{\mathrm{W}}}\left(\mathrm{R_s}\left(10\right)\frac{1}{4}\frac{m_{H^0}}{M}\right)^{\frac{1}{10}}+\frac{1}{2\cos^2\theta_W}\left(\mathrm{R_s}\left(10\right)\frac{4}{3}\frac{M}{m_Z}\right)^{\frac{1}{10}}-1=0, \qquad (6)$$

$$\mathrm{R_s}\left(11\right)=\frac{60}{\pi^6}, \qquad \mathrm{R_s}\left(10\right)=\frac{945}{64\pi^5},$$

where M is solved for the three charged lepton rest masses, $R_{\mathrm{s}}\left(n\right)$ is the the geometric constant for the radius of an n-sphere with surface area S embedded in $n+1$ dimensional euclidean space such that the radius $r=\left(R_{\mathrm{s}}\left(n\right)S\right)^{\frac{1}{n}}$. Note that while two terms appear to have the same exponent of $\frac{1}{10}$, the second term is $-\frac{1}{10}$ in terms of M while the third is $+\frac{1}{10}$ and this difference does allow for three positive, real roots. While it is not clear what significance if any equation 6 could have to real-world physics it is an example of the type of result that can be found with `nle-lepton` in 3-term mode.

VI. 2-TERM MIXED MODES AND SAMPLE RESULT

Starting with version 4.1 2-term mixed mode is supported. In this mode two mass terms are used with a constant term of -1 and the two mass terms are mixed to create a synthetic third (middle) term and this synthetic 3-term formula is then processed similar to 3-term mode. While this was inspired by two terms squared: $(a-b)^2=a^2-2ab+b^2$, the way it is implemented supports other related mixing modes such as full mesh (sum of products) of two copies of a and b: $a^2-4ab+b^2$, or any mixing that can be approximated by $a^2-nab+b^2$, where n is an integer ≥ 1. One interesting property of this mode is that the relative mass spectrum is determined only by the exponents and mixing mode. Since the synthetically generated middle coefficient is not independent of the other two there is no way to solve for a three mass spectrum unless it is inherently supported by the exponents and mixing mode. This provides a powerful check of the applicability of a formula without having to factor or interpret the factoring of the coefficients. This test is represented in the output as `two_term_test` which represents n in $a^2-nab+b^2$. If `two_term_test` is not sufficiently close to an integer ≥ 1 the formula is considered invalid and coefficient factoring is bypassed.

Starting with version 4.2, 2-term mixed mode with $(1-\text{smr})$ is supported. In 2-term mixed mode with `smrfactor_1minus_enable` set to `yes`, processing is similar to 2-term mode except the first term uses $(1-\text{smr})$ inside the radical, where smr = solution mass ratio and smr factors, and the second term uses just smr. Due to the

nature of $(1 - \mathrm{smr})$ specific reference masses and other smr factors need to be explicitly tested each time phase 1 is run instead of being scanned for during coefficient factoring. Configuration file options `smrfactor_*` control these tests.

An example of a formula found in 2-term mode with $(1 - \mathrm{smr})$ is

$$a^2 + 4ab + b^2 - 1 = 0, \tag{7}$$

$$a = \frac{9}{2\pi \cos\theta_W} \left(\frac{\mathrm{R_v}\,(3)}{\sqrt{2}} (1 - \mathrm{smr}) \right)^{\frac{1}{3}}, \qquad b = \frac{4}{3\pi^2 \cos\theta_W} \left(\frac{2\mathrm{R_v}\,(9)}{3\pi} \mathrm{smr} \right)^{\frac{1}{9}},$$

$$\mathrm{R_v}\,(3) = \frac{3}{4\pi}, \qquad \mathrm{smr} = \frac{4M}{\alpha_{\mathrm{em}} v}, \qquad \mathrm{R_v}\,(9) = \frac{945}{32\pi^4},$$

where M is solved for the three charged lepton rest masses, v is the Higgs vacuum expectation value ~ 246.22 GeV/c^2 , θ_W is the weak mixing angle and $R_v\,(n)$ is the geometric constant for the radius of an n-ball with volume V such that $r = (R_{\mathrm{v}}\,(n)\,V)^{\frac{1}{n}}$.

Note the polarity (sign) of the synthetic middle term is + instead of -. This is because $(1 - \mathrm{smr})$ can be negative for one or more solution masses and allows for three real roots with positive polarity. The choice of middle term polarity was made after initial tests found more interesting results with positive polarity. The source code can easily be modified for negative polarity and to expect $(1 - \mathrm{smr})$ to always be positive.

VII. INSTALLATION, CONFIGURATION, AND OPERATION

To install: unpack the source archive, change to the `src` directory and type `make`. Manually copy the binary executable `nle-lepton` to the desired executable directory.

An external user-editable configuration file (default: `nle-lepton.cfg`) is used to set operating modes, reference values, and coefficient factor parameters. Example configurations are provided for 3-term, 2-term and 2-term with 1-smr. Each option has a comment describing its function and valid values so they will not be exhaustively repeated here. Once example of coefficient factor parameters is `outfactor_pi_exp_up_max` and `outfactor_pi_exp_down_max`. These control what factors of π are tested outside the radical of each term. Setting them to `outfactor_pi_exp_up_max` = 3, and `outfactor_pi_exp_down_max` = 2, would enable testing of π^{-3}, $\pi^{-\frac{3}{2}}$, π^{-2}, π^{-1}, $\pi^{-\frac{1}{2}}$, 1, π, $\pi^{\frac{1}{2}}$, π^2, π^3, $\pi^{\frac{3}{2}}$. Obviously increasing the ranges of any parameter increases the time it takes to scan coefficient factors and the number of potential phase 2 formulas to be solved. To disable a factor from being tested set `_up_max` = 0, and `_down_max` = 1 for that ingredient.

Copy one of the examples of `nle-lepton.cfg` to a working config file and customize it to your search. You can specify the location of the configuration file at run-time with the `-c` command line option. Help is available with the `-h` command line option, and an external random seed can optionally be set with the `-s` option. The file `README.txt` in the source tree provides a summary of command line operation and output formatting.

VIII. CONCLUSION

A semi-automated approach to finding empirical mass formulas has advantages and disadvantages. Trillions of possible formulas can be tried and evaluated automatically to some extent but human analysis is ultimately required and a large number of results can make that process inefficient. More challenging is selecting the correct search parameters and formula structure to scan with. It is possible that NLE's are the wrong approach to this problem or that the right search ingredients are not included yet. The author fully expects that further development will be required before a useful formula offering clues to the underlying physics of fermion mass generation is found, if ever. `nle-lepton` was designed to be a platform adaptable to different investigative approaches through configuration parameters, including user-defined input options, and changes to the source code to support additional formula structures and features. `nle-lepton` is released under a standard BSD 3-clause open source license for maximum flexibility by other developers.

[1] M. Ibe, A. Kusenko, T. T. Yanagida, "Why three generations?", IPMU16-0014, 10.1016/j.physletb.2016.05.025, arXiv:1602.03003 [hep-ph]

[2] S. Weinberg, Phys. Rev. Lett. 19, 1264 (1967)
[3] A. Salam, Elementary Particle Theory, ed. N. Svartholm (Almquist and Wiksells, Stockholm, 1969), p. 367
[4] S.L. Glashow, J. Iliopoulos, and L. Maiani, Phys. Rev. D 2, 1285 (1970).
[5] H. E. Logan, "TASI 2013 lectures on Higgs physics within and beyond the Standard Model", arXiv:1406.1786 [hep-ph]
[6] H. Fritzsch, Z. Xing, "Mass and Flavor Mixing Schemes of Quarks and Leptons", Prog.Part.Nucl.Phys.45:1-81,2000, 10.1016/S0146-6410(00)00102-2, arXiv:hep-ph/9912358
[7] https://github.com/kevinloch/nle-lepton/releases
[8] Z. Xing, H. Zhang, S. Zhou, "Updated Values of Running Quark and Lepton Masses", Phys.Rev.D77:113016,2008, 10.1103/PhysRevD.77.113016, arXiv:0712.1419 [hep-ph]