Viability of perturbative expansion for neural network field theories

Srimoyee Sen

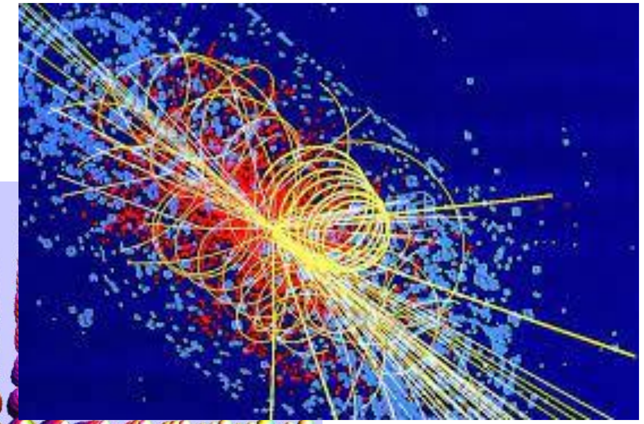
Iowa State University

In collaboration with Varun Vaidya,
U of South Dakota

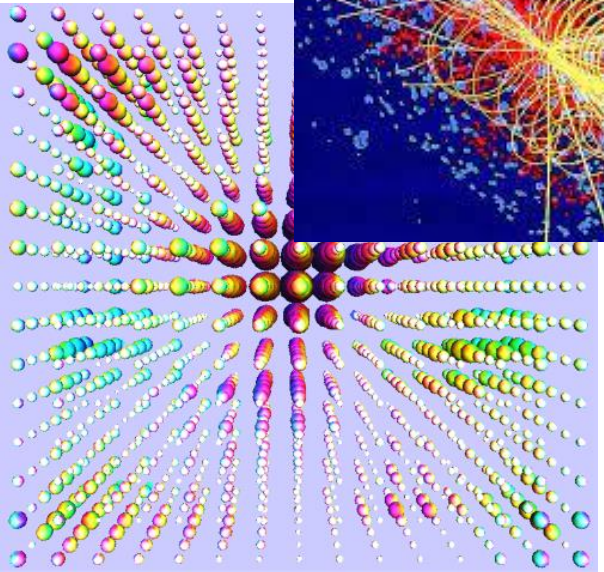
https://arxiv.org/pdf/2508.03810

LITP symposium on Theoretical Physics and AI, University of Michigan

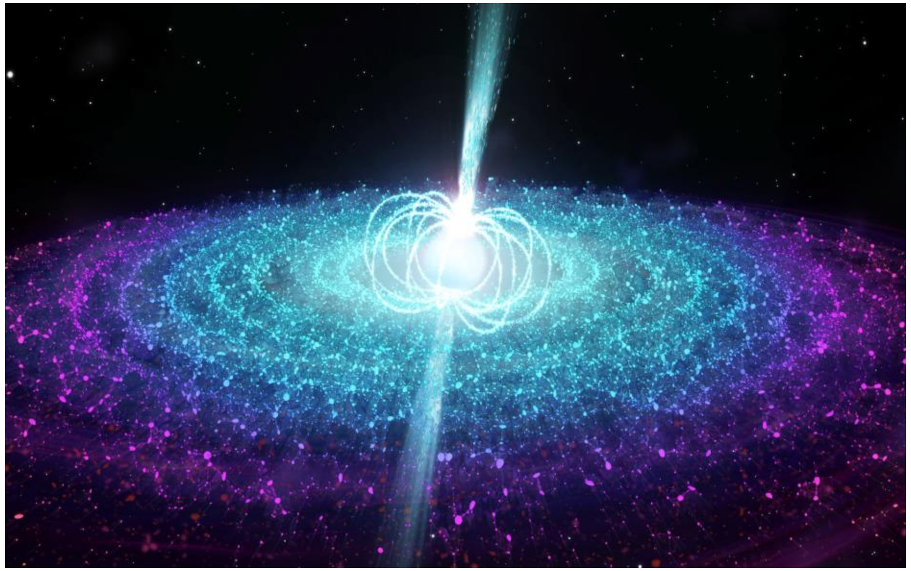
Research at the intersection of different subfields of physics



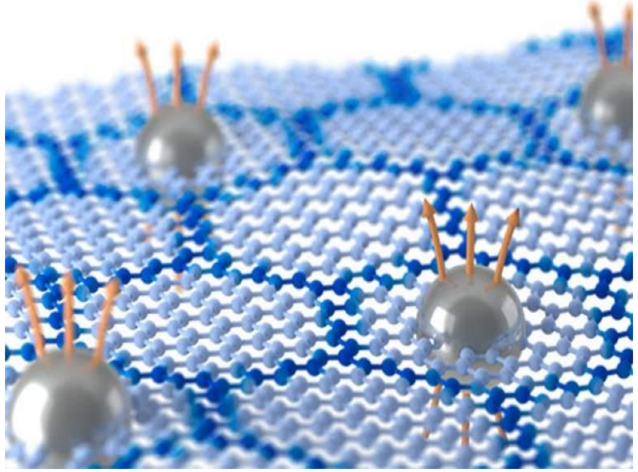
SM and BSM physics



Fundamental physics on discrete space-time



Neutron stars and astrophysics



Quantum materials

Supported by the Department of Energy **Early Career Award** 2021, "Quantum Materials, Lattice Gauge Theory and QCD"

This talk

- On a potential new formulation of quantum field theory known as neural network field theory (NNQFT). (Demirtas, Halverson, Naskar, Maiti, Schwartz, Stoner, Ferko, Zhang)

- Specifically focuses on the 2023 paper "Neural network field theories: non-Gaussianity, actions, and locality,"

----------- Other recent references 2511.16741, 2601.14453, 2512.24420

- Can you use this formulation to simulate QFT with finite resources? and can you extract non-perturbative physics from QFT using NNQFT?

- What would be its advantages and disadvantages in comparison with lattice QFT?

- Which observables would this formulation be suitable for?

This talk

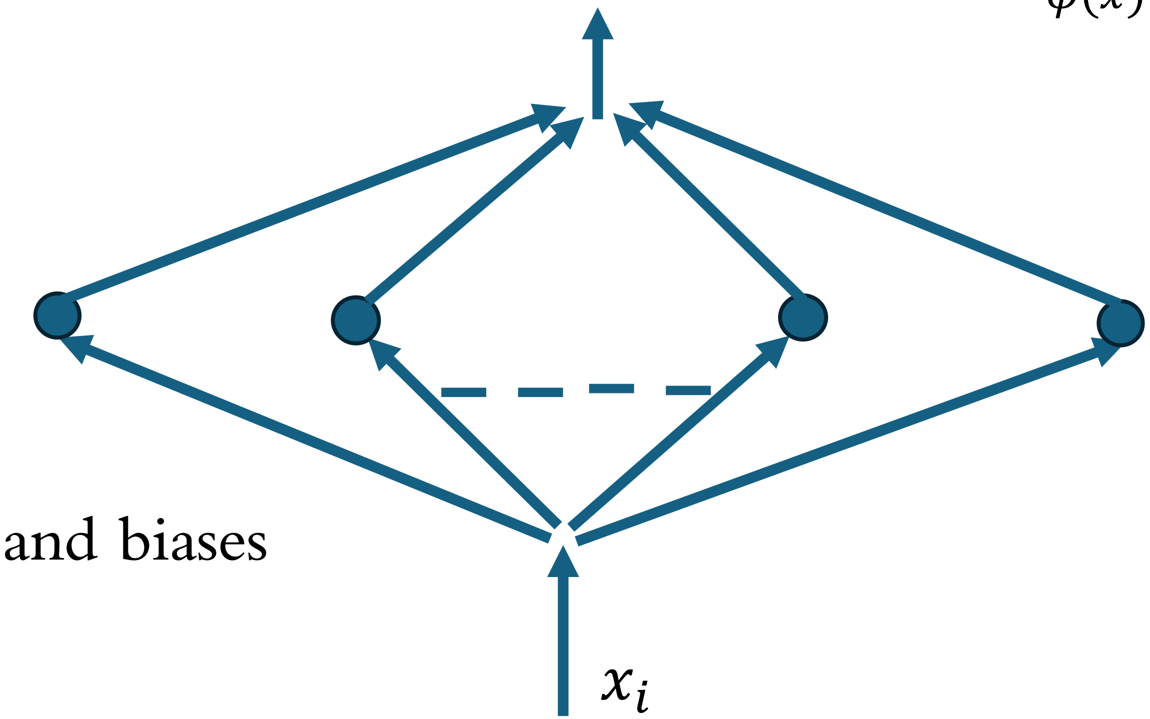
- DHMSS (2023) established how to formulate free scalar field theory and proposed a pathway to introduce interactions, $\lambda\phi^4$.
- Explore perturbative expansion in λ before moving on to non-perturbative physics.
- Compute correlation functions, work out renormalization and estimate the error budget.
- Revelations: The standard process of renormalization does not immediately apply to NNQFT.
- New ideas are needed to make it work.

Structure of NNQFT (one layer of neurons)

One layer of neurons and N neurons in that layer, d space-time dimensions

Output: sum of activation functions = field value

$$\phi(x) = \frac{\sqrt{2V_d}}{\sigma_a(2\pi)^d} \sum_{k=1..N} \frac{a_k \cos(b_{kj}x_j + c_k)}{\sqrt{\sum b_{kj}^2 + m^2}}$$



Neurons

Neuron number

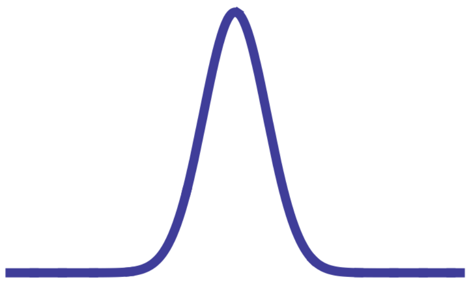
$$\frac{a_k \cos(b_{kj}x_j + c_k)}{\sqrt{\sum b_{kj}^2 + m^2}}$$

a, b, c are weights and biases

Input: space-time coordinate

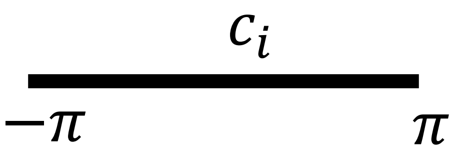
Sampling and computing correlation functions

- Sample a_i, b_{ij}, c_i from the following distribution functions:

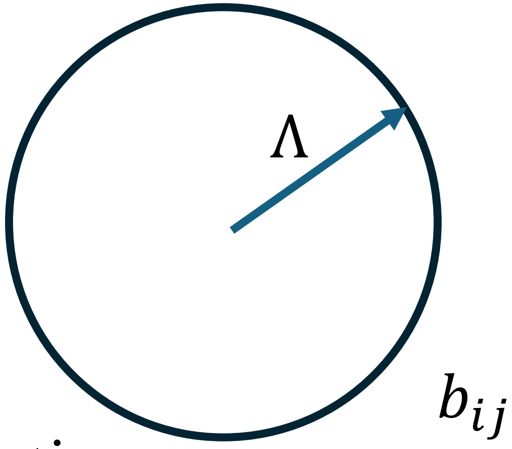


$$P_{a_i} = \frac{\sqrt{N}}{\sqrt{2\pi}\,\sigma_a} e^{-N a_i^2/(2\sigma_a^2)},.$$

$$P_{c_i} = \frac{1-\Theta\big((c_i+\pi)(c_i-\pi)\big)}{2\pi}$$



$$P_{b_{ij}} = \frac{\Theta\left(\Lambda - \sqrt{\sum_i b_{ij}^2}\right)}{V_d}, \quad V_d = \#\,\Lambda^4$$



- Say, we want to compute $< \phi(x)\phi(y) >$ correlation function.

- Sample a, b, c. For each such sample, compute $\phi(x)$ and $\phi(y)$.

- Take their product and then average.

Formally

$$Z_{NN,\text{free}}(J) = \int da_i db_{ij} dc_i \, (P_{a_i} P_{b_{ij}} P_{c_i}) e^{\int d^d x \, J(x)\phi(x)}$$

Interaction:

$$Z_{NN,\text{interaction}}(J) \propto \int da_i db_{ij} dc_i \, (P_{a_i} P_{b_{ij}} P_{c_i}) e^{\int d^d x \, J(x)\phi(x) - \frac{\lambda}{4!}\int \phi(x)^4}$$

P_{abc}

Two-point correlation functions (leading order):

$$< \phi(x)\phi(y) > = \frac{2V_d}{\sigma_a^2(2\pi)^d} \prod_i \int da_i \, db_{ij} dc_i P_{abc} \sum_{k=1..N} \frac{a_k \cos(b_{kj}x_j + c_k)}{\sqrt{\sum b_{kj}^2 + m^2}} \sum_{l=1..N} \frac{a_l \cos(b_{lj}y_j + c_l)}{\sqrt{\sum b_{lj}^2 + m^2}}$$

$k = l$ For a nonzero answer

$$< \phi(x)\phi(y) > = \int_{V_d} \cdot \frac{d^d b}{(2\pi)^d} \frac{1}{b^2 + m^2} e^{ib(x-y)}$$

No error at this point. Exact field theory result.

Higher-point correlation functions or two-point at higher order (contractions of multiple pairs involved):

$$< \phi(w_1)\phi(w_2)\phi(w_3)\phi(w_4) > = \cdots . \sum_{k=1..N} \frac{a_k \cos\left(b_{kj} w_{1j} + c_k\right)}{\sqrt{\sum b_{kj}^2 + m^2}} \sum_{l=1..N} \frac{a_l \cos\left(b_{lj} w_{2j} + c_l\right)}{\sqrt{\sum b_{lj}^2 + m^2}}$$

$$\sum_{q=1..N} \frac{a_q \cos\left(b_{qj} w_{3j} + c_q\right)}{\sqrt{\sum b_{qj}^2 + m^2}} \sum_{r=1..N} \frac{a_r \cos\left(b_{rj} w_{4j} + c_r\right)}{\sqrt{\sum b_{rj}^2 + m^2}}$$

Non-field theory results: $k = l = q = r$

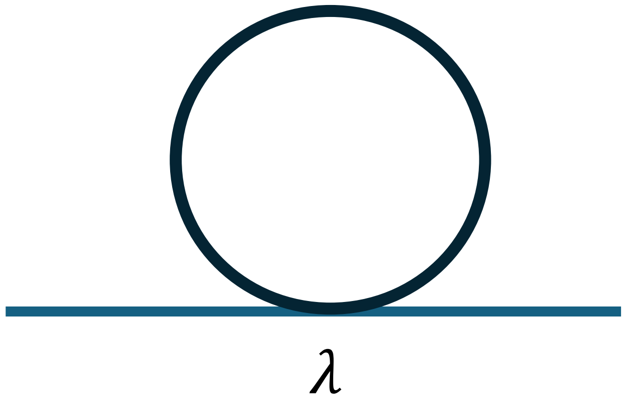
See Kevin Zhang's recent paper on this. 2604.27050

Field theory results arise from pairwise Gaussian integrals: E.g. $k = l \neq q = r$

Leads to an overall factor of $\frac{N(N-1)}{N^2}$

For i pairwise contractions, factor of : $\frac{N(N-1)(N-2)\ldots(N-i+1)}{N^i}$

Two-point at first order and mass renormalization



No N factor.

Three contractions: $\frac{N(N-1)(N-2)}{N^3} \approx 1 - \frac{3}{N}$.

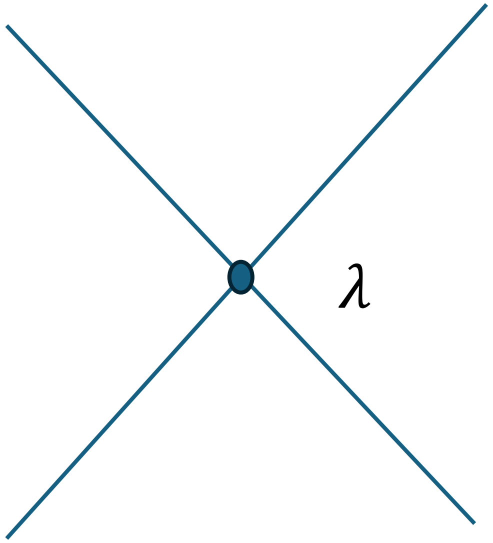
$$G^2(p) = \frac{1}{p^2+m^2} - \frac{\lambda}{2}\left(1-\frac{3}{N}\right)\left(\frac{1}{p^2+m^2}\right)^2 \int \frac{d^d q}{q^2+m^2} + \cdots.$$

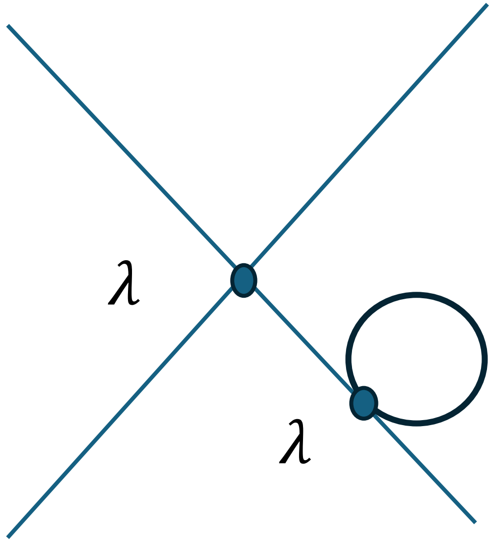
Bubble divergence

$$\delta m^2 \sim \frac{\lambda}{2}\left(1-\frac{3}{N}\right)\int \frac{d^d q}{q^2+m^2} + \cdots.$$

Includes field theory quadratic divergence + NNFT 1/N correction (also divergent)

Four-point correlation function and its renormalization





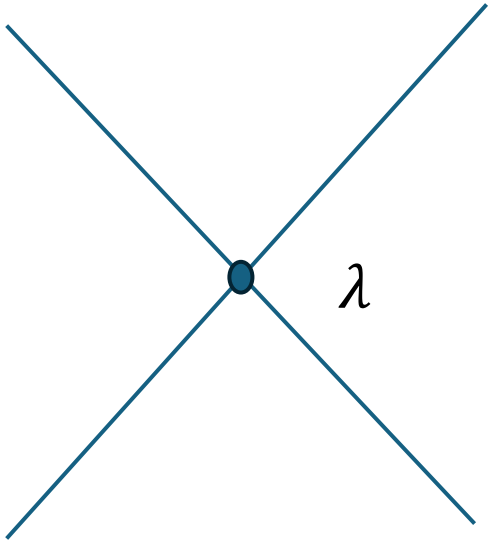
Combinatoric factor

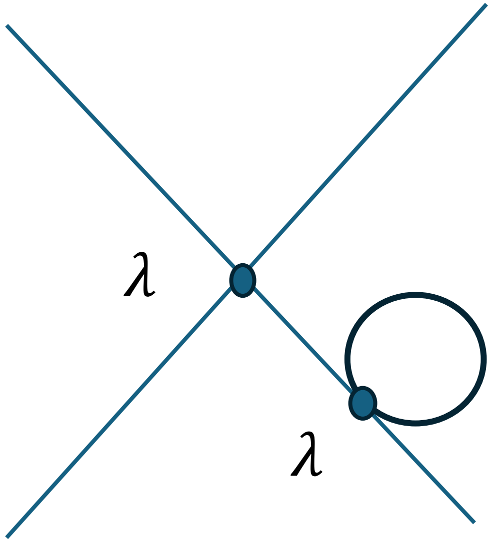
$$\frac{N(N-1)(N-2)(N-3)}{N^4} = 1 - \frac{6}{N} + ..$$

Combinatoric factor:

$$\frac{N(N-1)(N-2)(N-3)(N-4)(N-5)}{N^6} = 1 - \frac{15}{N} + ..$$

Four-point correlation function and its renormalization





The divergence from field theory (red) piece cancels.

But the additional NNFT divergence green piece does not.

Combinatoric factor

$$\frac{N(N-1)(N-2)(N-3)}{N^4} = 1 - \frac{6}{N} + ..$$

Combinatoric factor:

$$\frac{N(N-1)(N-2)(N-3)(N-4)(N-5)}{N^6} = 1 - \frac{15}{N} + ..$$

Four-point correlation function after renormalization

Observable → $\chi_{4,k} = \chi_{4,\tilde{p}} + \chi_{4,\tilde{p}}^2\left(Q(k_1,k_2,k_3,k_4) - Q(\tilde{p}_1,\tilde{p}_2,\tilde{p}_3,\tilde{p}_4)\right)$

Uncancelled divergent piece remains in prediction

$$Q = \left(1 - \frac{3}{N}\right)\left(V(s) + V(t) + V(u)\right) - \frac{\frac{3}{N}\int d^d q.\ \frac{1}{q^2 - \tilde{p}^2 + \tilde{\chi}_2^{-1}}}{(2\pi)^d}\sum_j \frac{1}{k_j^2 - \tilde{p}^2 + \tilde{\chi}_2^{-1}(\tilde{p})} + ..$$

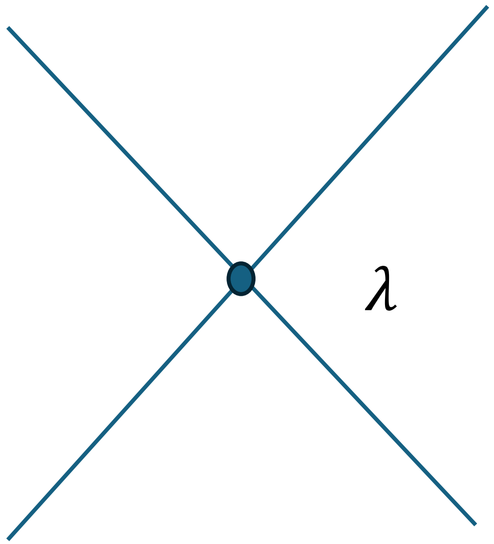
$$V(s) = \frac{1}{2}\int \frac{d^d q}{(2\pi)^d}\cdot\frac{1}{(q^2 + m^2)((q+s)^2 + m^2)}$$

A way out

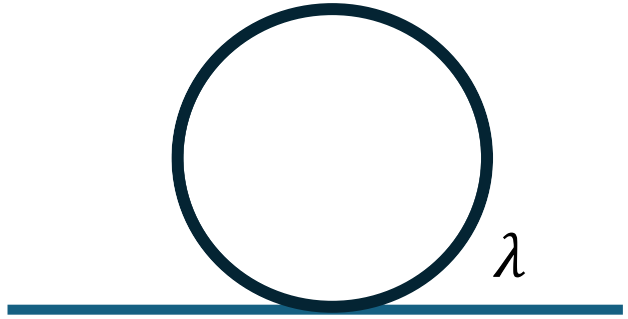
Modify the interaction to:

$$\frac{\lambda}{4!}\int d^d x \sum_{k_1\neq k_2\neq k_3\neq k_4} \phi_{k_1}(x)\,\phi_{k_2}(x)\,\phi_{k_3}(x)\,\phi_{k_4}(x)$$

Retains:



But eliminates completely

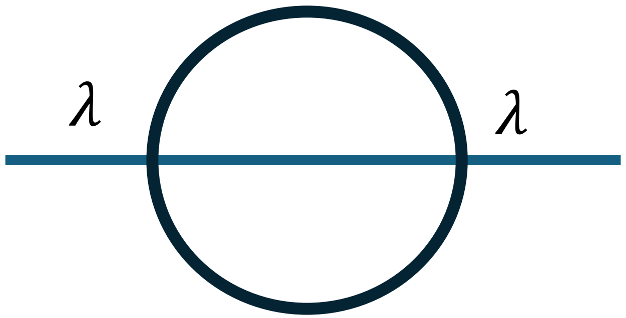


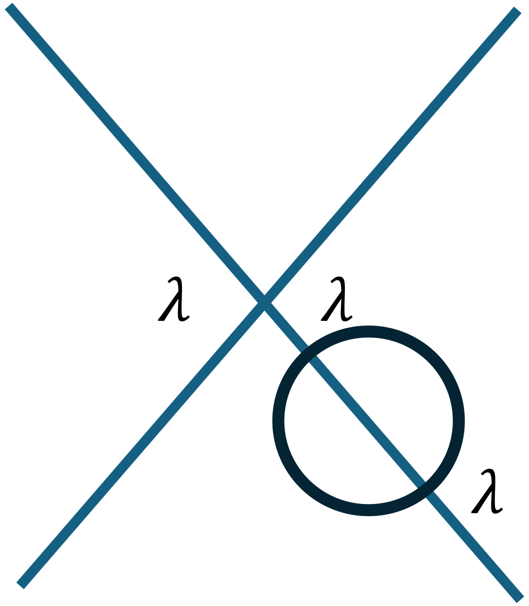
Fixes the problem up to order λ^2 !!

Removes uncancelled divergences up to this order

However

Similar divergences at higher order in λ, e.g. the sunset diagram.
Quadratically divergent



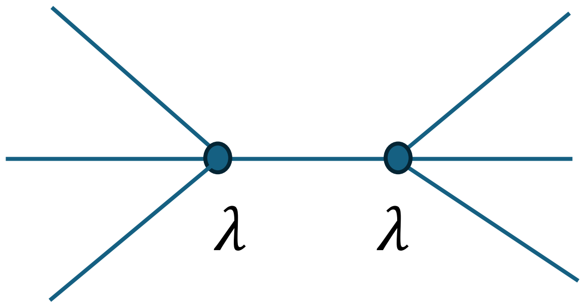


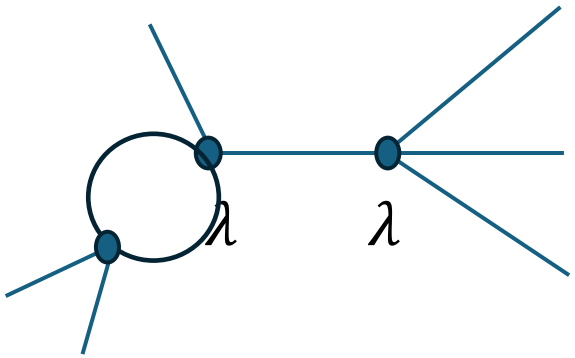
Structure of 1/N corrections different, divergences won't cancel again

Thus, the fix does not quite work beyond order λ^2

Higher point correlation function

Even at order λ^2 higher point correlation functions remain divergent with the modification even though four point is finite.





Remains divergent

And the fix does not quite work at order λ^2 simultaneously for all correlation functions.

Summary

- The free NNQFT formulation is a QFT formulation in continuum space-time.

- If extended to interacting theories, it will be a non-perturbative formulation of QFT in the continuum that can be simulated on a computer.

- Obstructions to the original proposal remain.

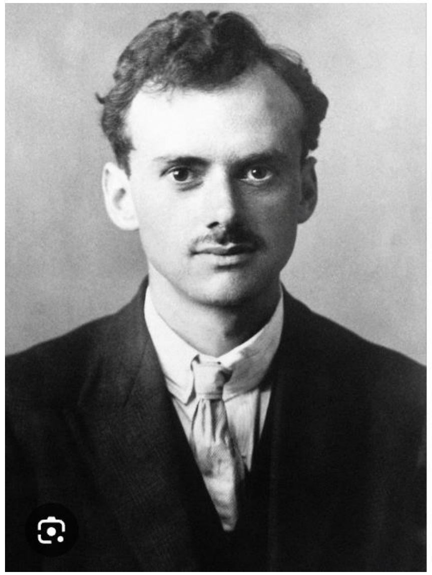
- The usual method of perturbative renormalization does not seem to work. Leaves one with uncancelled divergences in predictions.

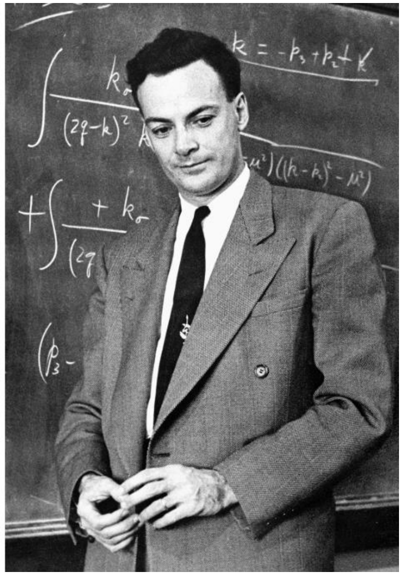
- Are there other ways of making this formulation work?

Summary

- If we are able to make it work perturbatively, we will have to understand non-perturbative renormalization.
- We will need to identify for which kinds of observables NNFT can offer advantages over lattice QFT.
- Examples could be, definition in continuum space-time which will preserve exact rotational invariance and translation invariance.
- No need for a Brillouin zone, no fermion doubling problem?

Outlook





Quote from Feynman about Dirac's book that got him to think about QED:

I couldn't understand the book very well because I really wasn't up to it. But there in the last paragraph at the end of the book it said, "Some new ideas are here needed." And so there I was. Some new ideas were needed? OK! So I started to think of new ideas.

That's where we are with regards to NNFT: "Some new ideas are here needed."

"Take the world from another point of view"- Yorkshire television program