The Sign Problem in QCD

Subtleties with the infinite volume limit

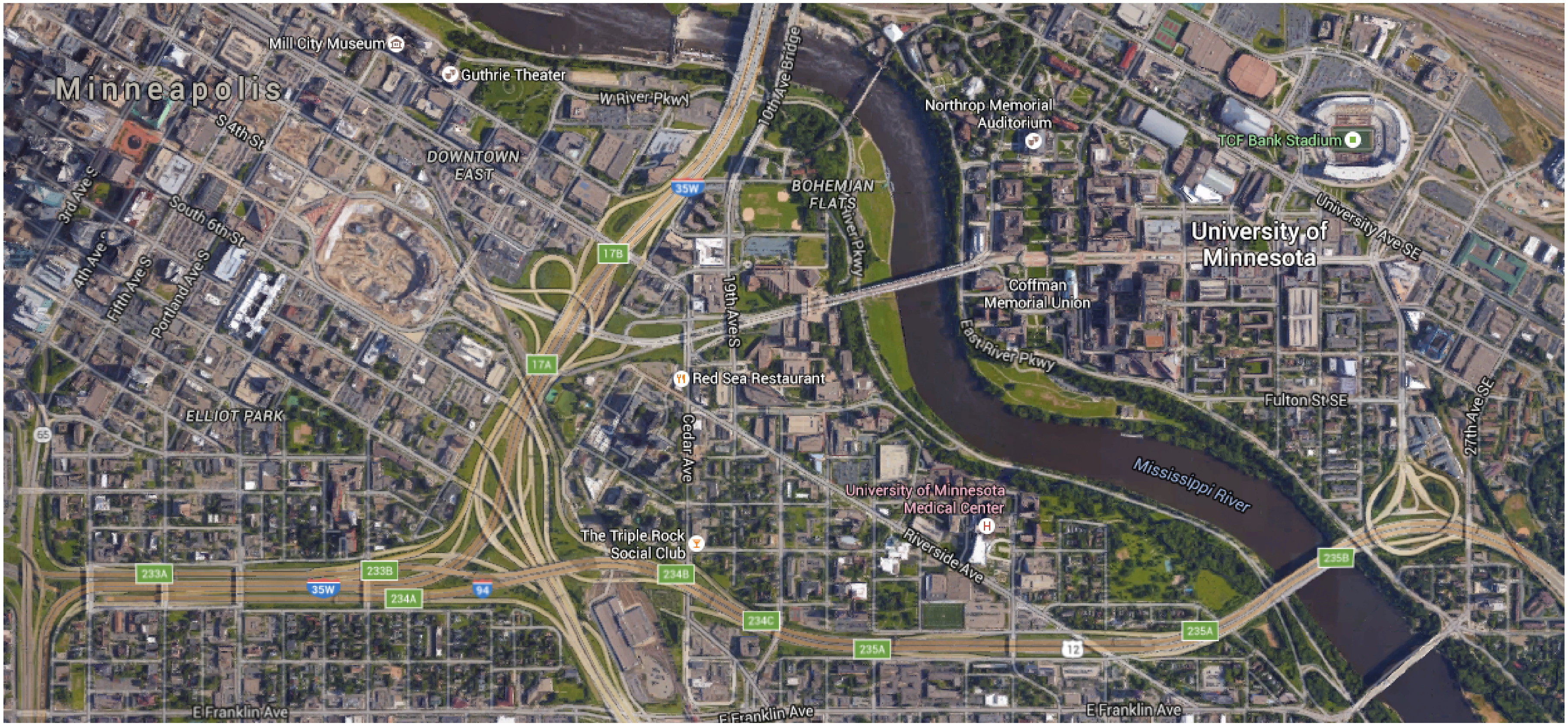
TDC, Yiming Cai◆, Brian McPeak* & Ari Goldbloom-Helzner♯

◆ Maryland grad student
* Maryland undergrad, now a Michigan grad student
♯ Montgomery Blair High School student

arXiv:1603.00478

Overview

Overview



Minneapolis
Mill City Museum
Guthrie Theater
W River Pkwy
10th Ave Bridge
Northrop Memorial Auditorium
TCF Bank Stadium
S 4th St
DOWNTOWN EAST
BOHEMIAN FLATS
River Pkwy
3rd Ave S
South 6th St
University of Minnesota
University Ave SE
4th Ave S
Fifth Ave S
Portland Ave S
19th Ave S
Coffman Memorial Union
East River Pkwy
Red Sea Restaurant
Fulton St SE
ELLIOT PARK
Cedar Ave
27th Ave SE
Mississippi River
University of Minnesota Medical Center
The Triple Rock Social Club
Riverside Ave
35W
17B
17A
233A
233B
234A
94
234B
234C
235A
235B
12
65
E Franklin Ave
E Franklin Ave
E Franklin Ave

Overview

- Introduction
  - What is the sign problem and why it matters
  - Why study the θ-term
- The sign problem for QCD with a θ-term
  - A surprise with the infinite volume limit: a puzzle emerges
  - A resolution via the saddle point approximation
  - Implications

The Fermion Sign Problem in Lattice QCD

Houston, we have a problem

The Fermion Sign Problem in Lattice QCD

Houston, we have a problem

- The functional determinant is complex when a chemical potential is included leading to cancelling signs in the functional integral that require accuracy which grows exponentially with the volume; there is no practical way to proceed.
  - Finding a way to evade the fermion sign problem is the most important theoretical issue facing lattice QCD.

The Fermion Sign Problem in Lattice QCD

Houston, we have a problem

- The functional determinant is complex when a chemical potential is included leading to cancelling signs in the functional integral that require accuracy which grows exponentially with the volume; there is no practical way to proceed.
  - Finding a way to evade the fermion sign problem is the most important theoretical issue facing lattice QCD.

- The fermion sign problem is an example of a very general class of sign problems.

Houston, we have a problem

The General Sign Problem

Houston, we have a problem

- The general sign problem almost certainly does not have a solution (i.e. an algorithm with run time polynomial in the volume.)
  - Troyer and Wiese (PRL 2005) showed that there exists a system (a type of spin glass) known to be NP hard that suffers from a sign problem. Thus a generic solution implies that P=NP. It is strongly believed among complexity theorists that P≠NP. However this is unproved.

The General Sign Problem

Houston, we have a problem

- The general sign problem almost certainly does not have a solution (i.e. an algorithm with run time polynomial in the volume.)
  - Troyer and Wiese (PRL 2005) showed that there exists a system (a type of spin glass) known to be NP hard that suffers from a sign problem. Thus a generic solution implies that P=NP. It is strongly believed among complexity theorists that P≠NP. However this is unproved.
  - If you prove either that P≠NP or P=NP, then you win $1,000,000 from the Clay Mathematical institute Thus if you can find a general algorithm for the sign problem, you win a million bucks.

The General Sign Problem

Houston, we have a problem

- The general sign problem almost certainly does not have a solution (i.e. an algorithm with run time polynomial in the volume.)
  - Troyer and Wiese (PRL 2005) showed that there exists a system (a type of spin glass) known to be NP hard that suffers from a sign problem. Thus a generic solution implies that P=NP. It is strongly believed among complexity theorists that P≠NP. However this is unproved.
  - If you prove either that P≠NP or P=NP, then you win $1,000,000 from the Clay Mathematical institute Thus if you can find a general algorithm for the sign problem, you win a million bucks.

All I ask is a 10% finders fee for pointing this out to you!!

Sign Problems in QCD

- The theorem need not exclude viable solutions to the fermion sign problem that depend on particularly features of QCD.

Sign Problems in QCD

- The theorem need not exclude viable solutions to the fermion sign problem that depend on particularly features of QCD.
- It may be useful to get insight and experience from other sign problems that can emerge in QCD.

Sign Problems in QCD

- The theorem need not exclude viable solutions to the fermion sign problem that depend on particularly features of QCD.
- It may be useful to get insight and experience from other sign problems that can emerge in QCD.
- A particularly interesting case is QCD and related gauge theories with a θ-term. Theories with θ-terms intrinsically have sign problems

$$\mathcal{L} = \mathcal{L}_{QCD} + \theta \frac{g^2}{32\pi^2} G\tilde{G}$$

In Minkowski space

$$\mathcal{L} = \mathcal{L}_{QCD} + \theta \frac{g^2}{32\pi^2} G\tilde{G}$$

In Minkowski space

- Some interesting features
  - Theta term breaks P and CP.
    - Note that in strong sector CP violation is VERY small (as seen for example in the neutron electric dipole moment.) Why is an interesting question (the strong CP puzzle) which is beyond the scope of this talk.
  - The θ–term in Euclidean problem is imaginary. Thus there is a sign problem.
  - The problem is topological in Eucliden space: the theta term can be written as a total derivative (albeit of a gauge variant quantity) and hence depends on the boundary. Its volume integral is a winding number which is constrained to be an integer, we will denote the winding number as *Q*.

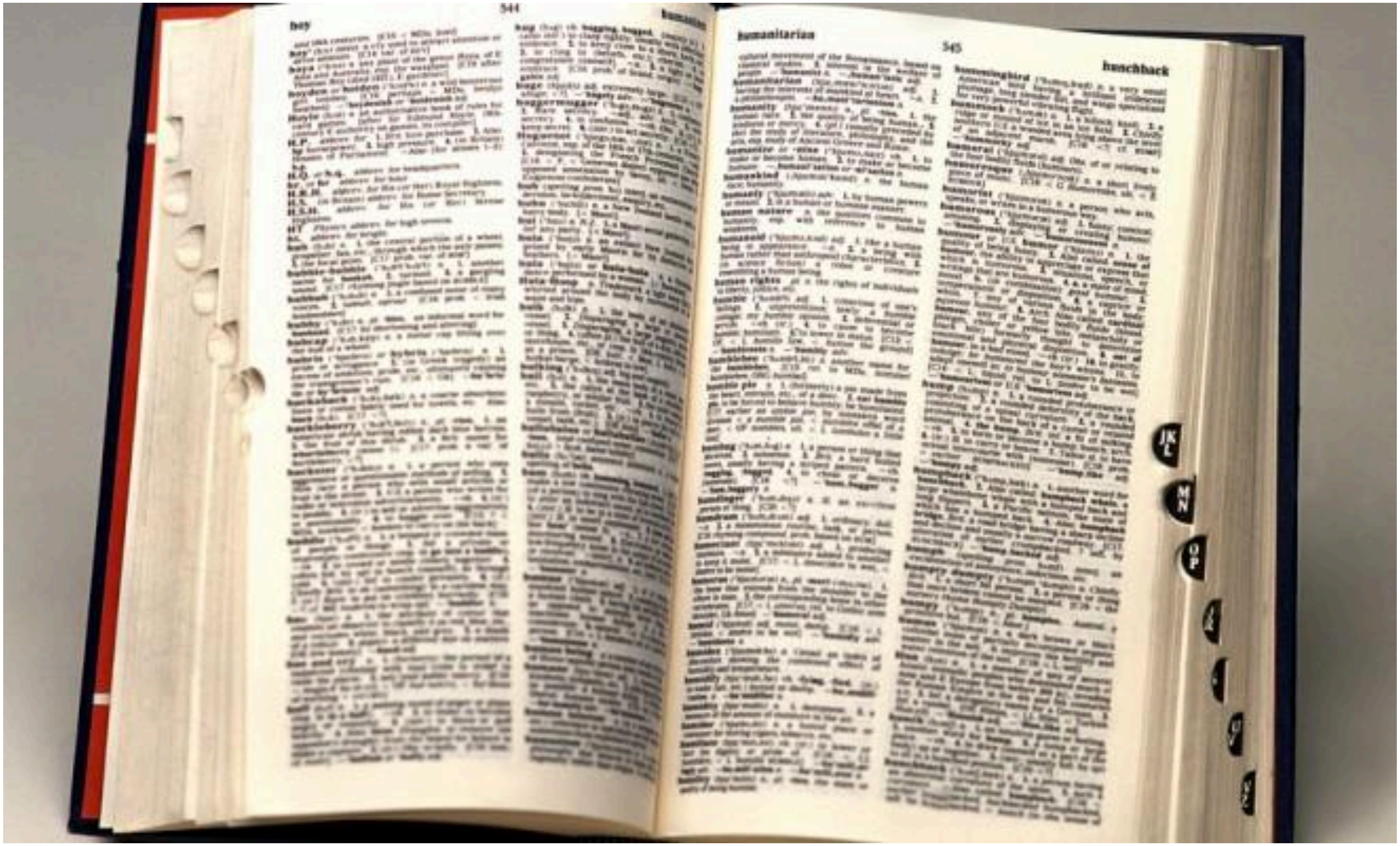
$$\mathcal{L} = \mathcal{L}_{QCD} + \theta \frac{g^2}{32\pi^2} G\tilde{G}$$

In Minkowski space

- Some interesting features
  - Theta term breaks P and CP.
    - Note that in strong sector CP violation is VERY small (as seen for example in the neutron electric dipole moment.) Why is an interesting question (the strong CP puzzle) which is beyond the scope of this talk.
  - The θ–term in Euclidean problem is imaginary. Thus there is a sign problem.
  - The problem is topological in Eucliden space: the theta term can be written as a total derivative (albeit of a gauge variant quantity) and hence depends on the boundary. Its volume integral is a winding number which is constrained to be an integer, we will denote the winding number as *Q*.

$$Q = \int_V \frac{g^2}{32\pi^2} \epsilon^{\alpha\beta\mu\nu} G_{\alpha\beta} G_{\mu\nu}$$

Some definitions



Definitions of key quantities

$$\varepsilon(\theta) = -\lim_{V\to\infty} \frac{1}{V} \log \mathcal{Z}(\theta, V)$$

Where V is the four dimensional volume of the box.

$$\mathcal{Z}(\theta, V) = \int [dA] \, \det[i \not{D}[A] - M] \, \exp(-S_{YM} - i\theta Q)$$

This functional integral can be broken up into a sum over topological sectors.

Definitions of key quantities

$$\varepsilon(\theta) = -\lim_{V\to\infty} \frac{1}{V} \log \mathcal{Z}(\theta, V)$$

Where V is the four dimensional volume of the box.

$$\mathcal{Z}(\theta, V) = \int [dA] \, \det[i\not{D}[A] - M] \, \exp(-S_{YM} - i\theta Q)$$

This functional integral can be broken up into a sum over topological sectors.

$$\mathcal{Z}(\theta, V) = \sum_{Q\in\mathbb{Z}} \mathcal{Z}_Q(V)\, e^{i\theta Q}$$

Oscillates rapidly indicating a sign problem

$$\mathcal{Z}_Q(V) = \frac{1}{\pi}\int_0^\pi \mathcal{Z}(\theta)\, e^{-i\theta Q}\, \mathrm{d}\theta$$

Precise form uses the fact that the theta dependence is even and periodic.

Definitions of some key quantities

$$q \equiv \frac{Q}{V}$$

Defines the topological charge density

$$\tilde{\varepsilon}(q,V) = -\frac{1}{V}\log(\mathcal{Z}_{(qV)}(V))$$

Defines an energy density as a function of topological charge density

$$\tilde{\varepsilon}(q) \equiv \lim_{V\to\infty} \tilde{\varepsilon}(q,V) \;\; \text{with } qV \text{ a positive integer}$$

Definitions of some key quantities

$$q \equiv \frac{Q}{V}$$

Defines the topological charge density

$$\tilde{\varepsilon}(q,V) = -\frac{1}{V}\log(\mathcal{Z}_{(qV)}(V))$$

Defines an energy density as a function of topological charge density

$$\tilde{\varepsilon}(q) \equiv \lim_{V\to\infty} \tilde{\varepsilon}(q,V) \;\; \text{with } qV \text{ a positive integer}$$

Intuitively this limit ought to exist; the quantity is the intensive energy density as a function of the topological charge density which is also intensive. One expects this to be independent of volume.

Sign Problem is Exponentially Sensitive in Volume

For simplicity
focus on $\varepsilon(\pi)$

Sign Problem is Exponentially Sensitive in Volume

$$\varepsilon(\pi) - \varepsilon(0) = -\lim_{V\to\infty} \frac{1}{V} \log\left(\frac{A(V) - B(V)}{A(V)}\right)$$

For simplicity focus on $\varepsilon(\pi)$

$$\text{with} \;\; A(V) \equiv \frac{\mathcal{Z}_{Q=0}(V)}{\mathcal{Z}(\theta = 0, V)}$$

Factor of 2 in *B* since $\mathcal{Z}_Q = \mathcal{Z}_{-Q}$

$$\text{and} \;\; B(V) \equiv -\frac{2\sum_{Q=1}^{\infty}(-1)^Q \mathcal{Z}_Q(V)}{\mathcal{Z}(\theta = 0, V)}$$

Sign Problem is Exponentially Sensitive in Volume

$$\varepsilon(\pi) - \varepsilon(0) = -\lim_{V\to\infty} \frac{1}{V} \log\left(\frac{A(V) - B(V)}{A(V)}\right)$$

For simplicity focus on ε(π)

$$\text{with } \; A(V) \equiv \frac{\mathcal{Z}_{Q=0}(V)}{\mathcal{Z}(\theta = 0, V)}$$

$$\text{and } \; B(V) \equiv -\frac{2\sum_{Q=1}^{\infty}(-1)^Q \mathcal{Z}_Q(V)}{\mathcal{Z}(\theta = 0, V)}$$

Factor of 2 in *B* since $\mathcal{Z}_Q = \mathcal{Z}_{-Q}$

In deriving this we have used the well known—and easy to see on physical grounds—fact that

$$\lim_{V\to\infty} -\frac{1}{V} \log\left(Z_{Q=0}(V)\right) = \epsilon(0)$$

Sign Problem is Exponentially Sensitive in Volume

$$\varepsilon(\pi) - \varepsilon(0) = -\lim_{V\to\infty} \frac{1}{V} \log\left(\frac{A(V) - B(V)}{A(V)}\right)$$

For simplicity focus on $\varepsilon(\pi)$

$$\text{with} \;\; A(V) \equiv \frac{\mathcal{Z}_{Q=0}(V)}{\mathcal{Z}(\theta=0,V)}$$

Factor of 2 in *B* since $\mathcal{Z}_Q = \mathcal{Z}_{-Q}$

$$\text{and} \;\; B(V) \equiv -\frac{2\sum_{Q=1}^{\infty}(-1)^Q \mathcal{Z}_Q(V)}{\mathcal{Z}(\theta=0,V)}$$

In deriving this we have used the well known—and easy to see on physical grounds—fact that

$$\lim_{V\to\infty} -\frac{1}{V}\log\left(Z_{Q=0}(V)\right) = \epsilon(0)$$

Thus to capture the θ dependence in ε, *B* must cancel *A* to one part in $\exp(V(\varepsilon(\pi) - (0)))$. Thus one needs both *A* and *B* at this accuracy.

Kinky Behavior

$\varepsilon(\theta) = \varepsilon(\theta + 2\pi)$ Q is an integer:the generating function must be periodic in θ $\mathcal{Z}(\theta, V) = \sum_{Q\in\mathbb{Z}} \mathcal{Z}_Q(V)\, e^{i\theta Q}$

$\varepsilon(\theta) = \varepsilon(-\theta)$ since QCD is CP invariant except for θ term

Combined with periodicity in θ, this means **the point $\theta = \pi$ is CP invariant.** If $\varepsilon(\theta)$ has a **kink** at $\theta = \pi$, the system **spontaneously breaks CP.**

Kinky Behavior

$\varepsilon(\theta) = \varepsilon(\theta + 2\pi)$ Q is an integer:the generating function must be periodic in θ $\mathcal{Z}(\theta, V) = \sum_{Q \in \mathbb{Z}} \mathcal{Z}_Q(V)\, e^{i\theta Q}$

$\varepsilon(\theta) = \varepsilon(-\theta)$ since QCD is CP invariant except for θ term

Combined with periodicity in θ, this means **the point $\theta = \pi$ is CP invariant.** If $\varepsilon(\theta)$ has a **kink** at $\theta = \pi$, the system **spontaneously breaks CP.**

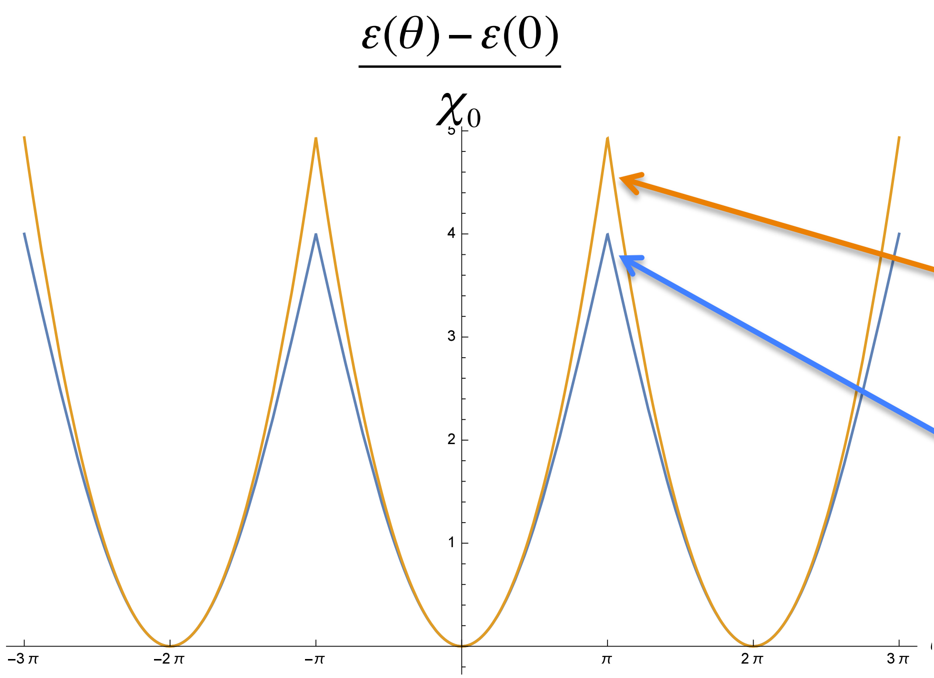
This possibility is not crazy. It happens for QCD in known limiting cases:

Kinky Behavior

$\varepsilon(\theta) = \varepsilon(\theta + 2\pi)$ Q is an integer:the generating function must be periodic in θ $\mathcal{Z}(\theta, V) = \sum_{Q\in\mathbb{Z}} \mathcal{Z}_Q(V)\, e^{i\theta Q}$

$\varepsilon(\theta) = \varepsilon(-\theta)$ since QCD is CP invariant except for θ term

Combined with periodicity in θ, this means **the point $\theta = \pi$ is CP invariant.** If $\varepsilon(\theta)$ has a **kink** at $\theta = \pi$, the system **spontaneously breaks CP.**



This possibility is not crazy. It happens for QCD in known limiting cases:

Infinitely many colors

Two flavors of infinitesimally light quarks

The energy was given in units of χ_0

$$\chi_0 \equiv \varepsilon''(0) = \lim_{V\to\infty}\left(\frac{\frac{1}{V}\sum_{Q\in Z} Q^2 Z_Q(V)}{\sum_{Q\in Z} Z_Q(V)}\right)$$

The topological susceptibility at zero

Sets the scale for variations of $\varepsilon(\theta)$

The energy was given in units of χ_0

$$\chi_0 \equiv \varepsilon''(0) = \lim_{V\to\infty}\left(\frac{\frac{1}{V}\sum_{Q\in Z} Q^2 Z_Q(V)}{\sum_{Q\in Z} Z_Q(V)}\right)$$

The topological susceptibility at zero

Sets the scale for variations of $\varepsilon(\theta)$

Knowledge of $\tilde{\varepsilon}(q,V)$ allows the reconstruction of $\varepsilon(\theta)$ by summing over topological sectors and then taking $V\to\infty$ limit.

The energy was given in units of χ_0

$$\chi_0 \equiv \varepsilon''(0) = \lim_{V\to\infty}\left(\frac{\frac{1}{V}\sum_{Q\in Z} Q^2 Z_Q(V)}{\sum_{Q\in Z} Z_Q(V)}\right)$$

The topological susceptibility at zero

Sets the scale for variations of $\varepsilon(\theta)$

Knowledge of $\tilde{\varepsilon}(q,V)$ allows the reconstruction of $\varepsilon(\theta)$ by summing over topological sectors and then taking $V\to\infty$ limit.

Does knowledge of $\tilde{\varepsilon}(q)$ alllow us to reconstruct $\varepsilon(\theta)$ by summing over topological sectors (even if exact)? That is can you take the $V\to\infty$ limit before summing.

The energy was given in units of χ_0

$$\chi_0 \equiv \varepsilon''(0) = \lim_{V\to\infty}\left(\frac{\frac{1}{V}\sum_{Q\in Z} Q^2 Z_Q(V)}{\sum_{Q\in Z} Z_Q(V)}\right)$$

The topological susceptibility at zero

Sets the scale for variations of $\varepsilon(\theta)$

Knowledge of $\tilde{\varepsilon}(q,V)$ allows the reconstruction of $\varepsilon(\theta)$ by summing over topological sectors and then taking $V\to\infty$ limit.

Does knowledge of $\tilde{\varepsilon}(q)$ alllow us to reconstruct $\varepsilon(\theta)$ by summing over topological sectors (even if exact)? That is can you take the $V\to\infty$ limit before summing.

- **Intuitively yes**: Both quantities are intensive quantities independent of volume. They describe the same system. They must contain the same info.

The energy was given in units of χ_0

$$\chi_0 \equiv \varepsilon''(0) = \lim_{V\to\infty}\left(\frac{\frac{1}{V}\sum_{Q\in Z} Q^2 Z_Q(V)}{\sum_{Q\in Z} Z_Q(V)}\right)$$

The topological susceptibility at zero

Sets the scale for variations of $\varepsilon(\theta)$

Knowledge of $\tilde{\varepsilon}(q,V)$ allows the reconstruction of $\varepsilon(\theta)$ by summing over topological sectors and then taking $V\to\infty$ limit.

Does knowledge of $\tilde{\varepsilon}(q)$ alllow us to reconstruct $\varepsilon(\theta)$ by summing over topological sectors (even if exact)? That is can you take the $V\to\infty$ limit before summing.

- **Intuitively Yes**: Both quantities are intensive quantities independent of volume. They describe the same system. They must contain the same info.
- ***Intuitively No***: *The sum over Q to obtain $\mathcal{Z}(\theta, V)$ requires* an exponentially accurate knowledge (in V) of the $\mathcal{Z}_Q(V)$ while we have dropped power law terms to get infinite volume limit.

Which Intuition is correct?

Which Intuition is correct?

This is the key question for the remainder of this talk

Which Intuition is correct?

This is the key question for the remainder of this talk

To pose this mathematically let us define a new quantity:

$$\underline{\varepsilon}(\theta) = -\lim_{V\to\infty}\frac{1}{V}\log\left(\sum_{Q} e^{-\tilde{\varepsilon}\left(\frac{Q}{V}\right)V}\, e^{i\theta Q}\right)$$

Uses $\tilde{\varepsilon}\left(\frac{q}{V}\right)$ **rather than** $\tilde{\varepsilon}\left(\frac{q}{V},V\right)$ **as is used to compute** $\varepsilon(\theta)$

The issue amounts to answer does $\underline{\varepsilon}(\theta)=\varepsilon(\theta)$?

The dilute instanton gas

- To get insight we first focused on a problem that is **NOT** QCD (in any known limit) but for which
  - The θ dependence is known and most aspects of the system are easily calculable
  - The system is known not to break CP spontaneously at $\theta=\pi$.
- We considered any theory that in some controlled limit is accurately described by a dilute instanton gas with instantons of fixed action.
- The virtue of systems of this type is that the calculable nature of the problem allows us to test ideas.

The dilute instanton gas

$\mathcal{Z}_Q(V)$ Is easy to compute

$$\boxed{\mathcal{Z}_Q(V) = \mathcal{Z}_0 \sum_{n=0}^{\infty} \frac{\left(\frac{1}{2}\, c\, e^{-S_0} V\right)^{|Q|+2n}}{n!\,(n+|Q|)!} = I_Q(ce^{-S_0}V)}$$

Where c is a constant, S_0 is the action of a single instanton and I is a modified Bessel function

The dilute instanton gas

$Z_Q(V)$ Is easy to compute

$$\mathcal{Z}_Q(V) = \mathcal{Z}_0 \sum_{n=0}^{\infty} \frac{\left(\frac{1}{2}\, c\, e^{-S_0} V\right)^{|Q|+2n}}{n!\,(n+|Q|)!} = I_Q(ce^{-S_0}V)$$

Where c is a constant, S_0 is the action of a single instanton and I is a modified Bessel function

Since $$I_Q(z) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{z\cos\theta}\cos(Q\theta)d\theta$$

If follows that $$\mathcal{Z}(\theta) = \mathcal{Z}_0 \exp\left(ce^{-S_0}V\cos(\theta)\right)$$

The dilute instanton gas

$Z_Q(V)$ Is easy to compute

$$\boxed{\mathcal{Z}_Q(V) = \mathcal{Z}_0 \sum_{n=0}^{\infty} \frac{\left(\frac{1}{2}\, c\, e^{-S_0} V\right)^{|Q|+2n}}{n!\,(n+|Q|)!} = I_Q(ce^{-S_0}V)}$$

Where c is a constant, S_0 is the action of a single instanton and I is a modified Bessel function

Since $$I_Q(z) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{z\cos\theta}\cos(Q\theta)d\theta$$

If follows that $$\boxed{\mathcal{Z}(\theta) = \mathcal{Z}_0 \exp\left(ce^{-S_0} V \cos(\theta)\right)}$$

Thus

$$\boxed{\varepsilon(\theta) = \varepsilon_0 + \chi_0(1-\cos(\theta))}$$

$$\text{with } \chi_0 \equiv ce^{-S_0}$$

$$\text{and } \varepsilon_0 \equiv \frac{-\log\left(\mathcal{Z}_0\right)}{V} - ce^{-S_0}$$

Get the standard dilute instanton gas result with a cosinusoidal dependance on θ; no kink.

Energy density for dilute instanton gas as a function of topological charge density can be extracted From $\mathcal{Z}_Q(V)$

$$\tilde{\varepsilon}(q,V) = -\log\left(I_{qV}(\chi_0 V)\right)$$

Known "uniform expansion" of modified Bessel function valid for $I_\nu(\nu x)$ in the limit of large V can be used to extract the infinite volume limit

$$\boxed{\tilde{\varepsilon}(q) = \varepsilon_0 + \chi_0 + q\sinh^{-1}\left(\frac{q}{\chi_0}\right) - \sqrt{\chi_0^2 + q^2}}$$

Energy density for dilute instanton gas as a function of topological charge density can be extracted From $\mathcal{Z}_Q(V)$

$$\tilde{\varepsilon}(q,V) = -\log\left(I_{qV}(\chi_0 V)\right)$$

Known "uniform expansion" of modified Bessel function valid for $I_\nu(\nu x)$ in the limit of large V can be used to extract the infinite volume limit

$$\boxed{\tilde{\varepsilon}(q) = \varepsilon_0 + \chi_0 + q \sinh^{-1}\left(\frac{q}{\chi_0}\right) - \sqrt{\chi_0^2 + q^2}}$$

We can test whether this is sufficient to reconstruct $\varepsilon(\theta)$ via direct summation over topoloical sectors:

$$\underline{\varepsilon}(\theta) = -\lim_{V\to\infty} \frac{1}{V} \log\left(\sum_Q e^{-\tilde{\varepsilon}\left(\frac{Q}{V}\right)V} e^{i\theta Q}\right)$$

Energy density for dilute instanton gas as a function of topological charge density can be extracted From $Z_Q(V)$

$$\tilde{\varepsilon}(q,V) = -\log\left(I_{qV}(\chi_0 V)\right)$$

Known "uniform expansion" of modified Bessel function valid for $I_\nu(\nu x)$ in the limit of large V can be used to extract the infinite volume limit

$$\boxed{\tilde{\varepsilon}(q) = \varepsilon_0 + \chi_0 + q\sinh^{-1}\left(\frac{q}{\chi_0}\right) - \sqrt{\chi_0^2 + q^2}}$$

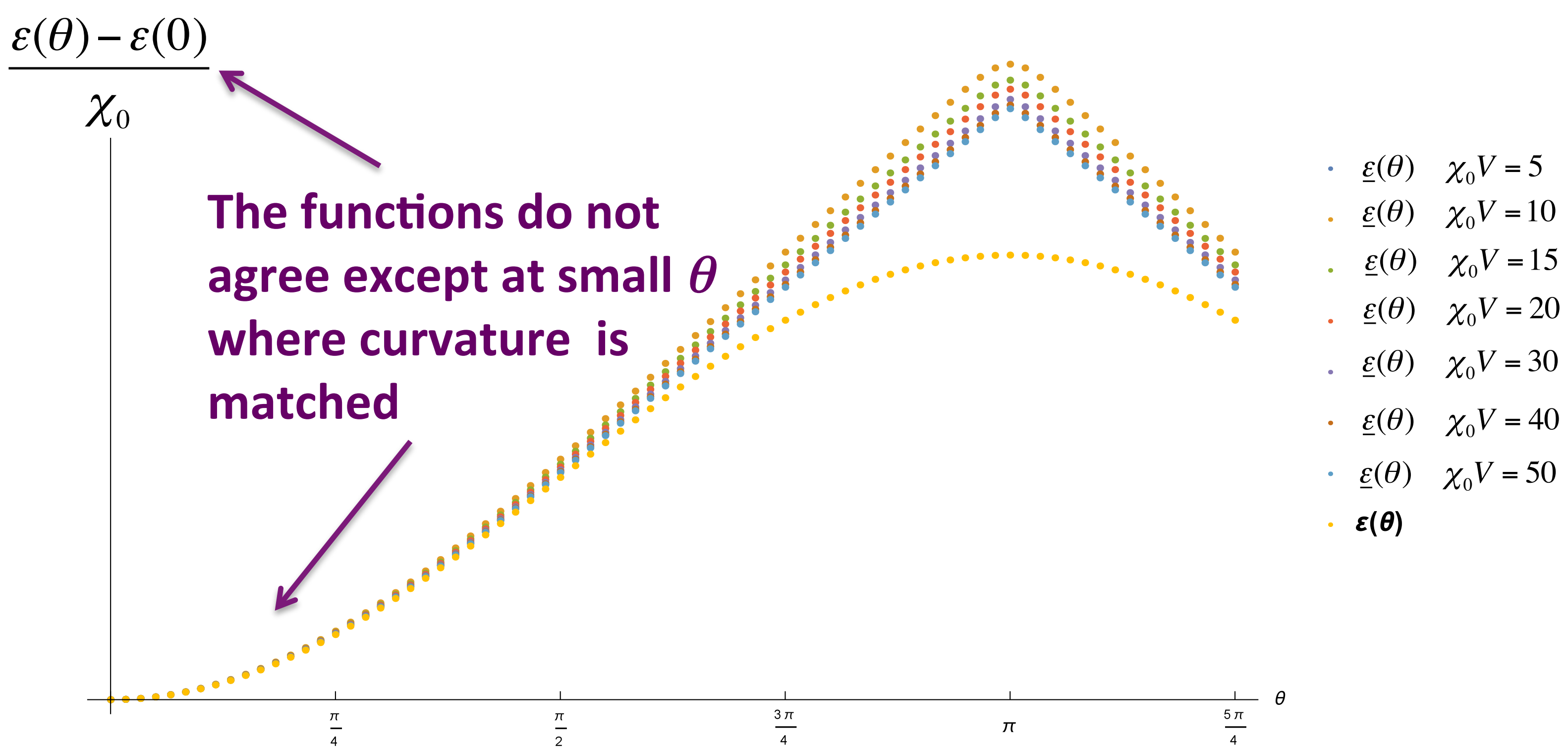
We can test whether this is sufficient to reconstruct $\varepsilon(\theta)$ via direct summation over topological sectors:

$$\underline{\varepsilon}(\theta) = -\lim_{V\to\infty}\frac{1}{V}\log\left(\sum_Q e^{-\tilde{\varepsilon}(\frac{Q}{V})V}\, e^{i\theta Q}\right)$$

Does $\underline{\varepsilon}(\theta) = \varepsilon(\theta)$?

- Ideally one could do this sum analytically for arbitrary V, obtain a functional form and then take the large V limit of the expression.
  - Unfortunately given the functional form for the energy density, it is not known how to obtain a closed form expression for this sum.
  - Instead, we can do the sum numerically for large but finite V (the relevant combination is $\chi_0 V$) and a number of values for θ.

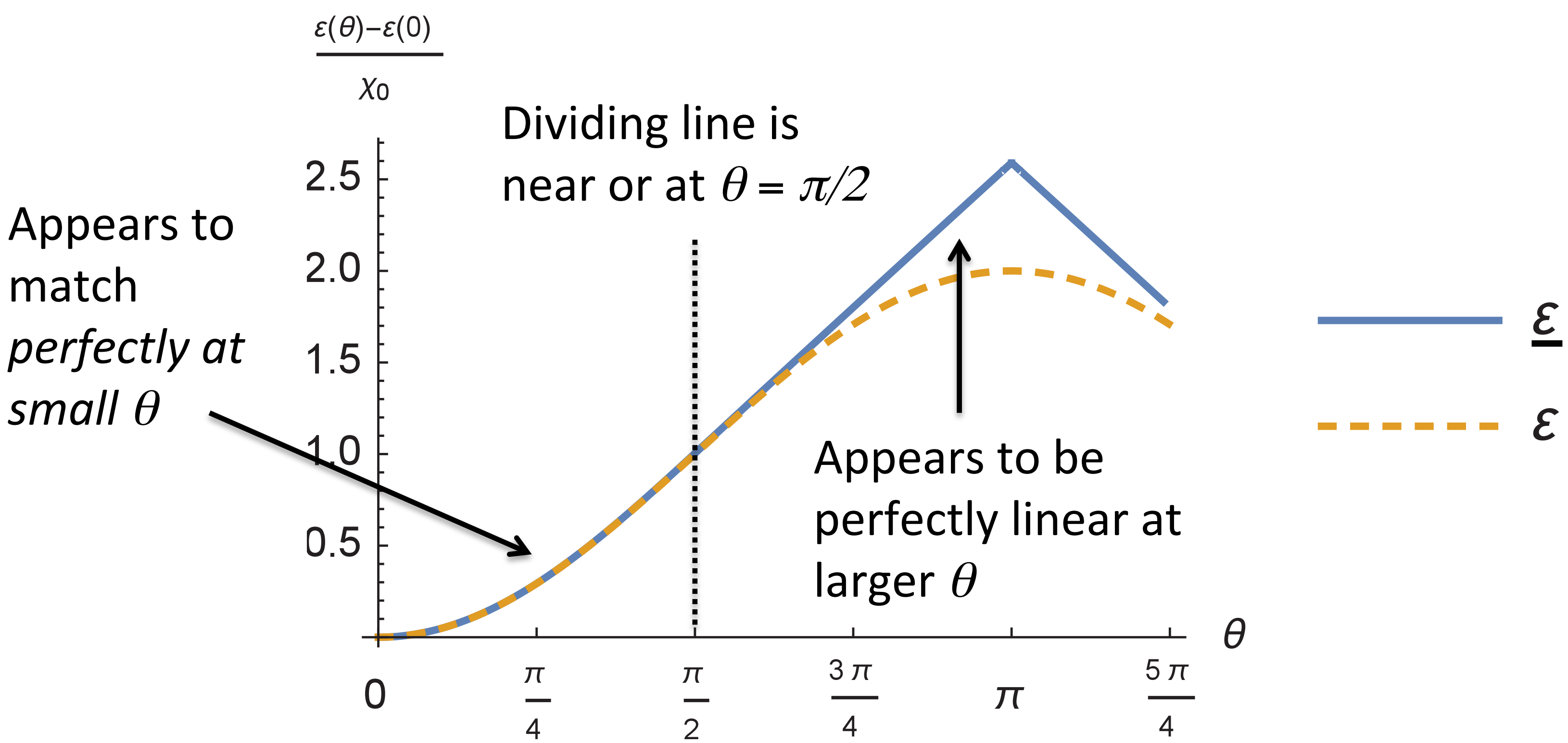
- Ideally one could do this sum analytically for arbitrary *V*, obtain a functional form and then take the large *V* limit of the expression.
  - Unfortunately given the functional form for the energy density, it is not known how to obtain a closed form expression for this sum.
  - Instead, we can do the sum numerically for large but finite *V* (the relevant combination is $\chi_0 V$) and a number of values for θ.

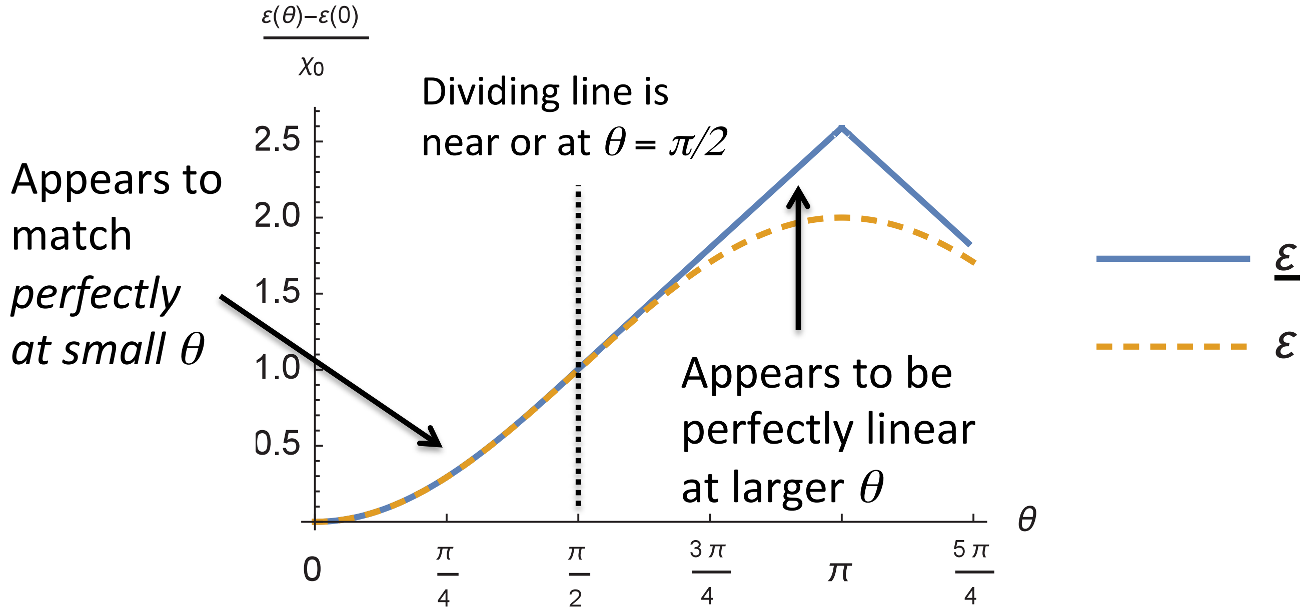


- This appears to show conclusively that the sign problem is so so severe that knowledge of the infinite volume energy density as a function of q is not enough to reconstruct $\varepsilon(\theta)$!!

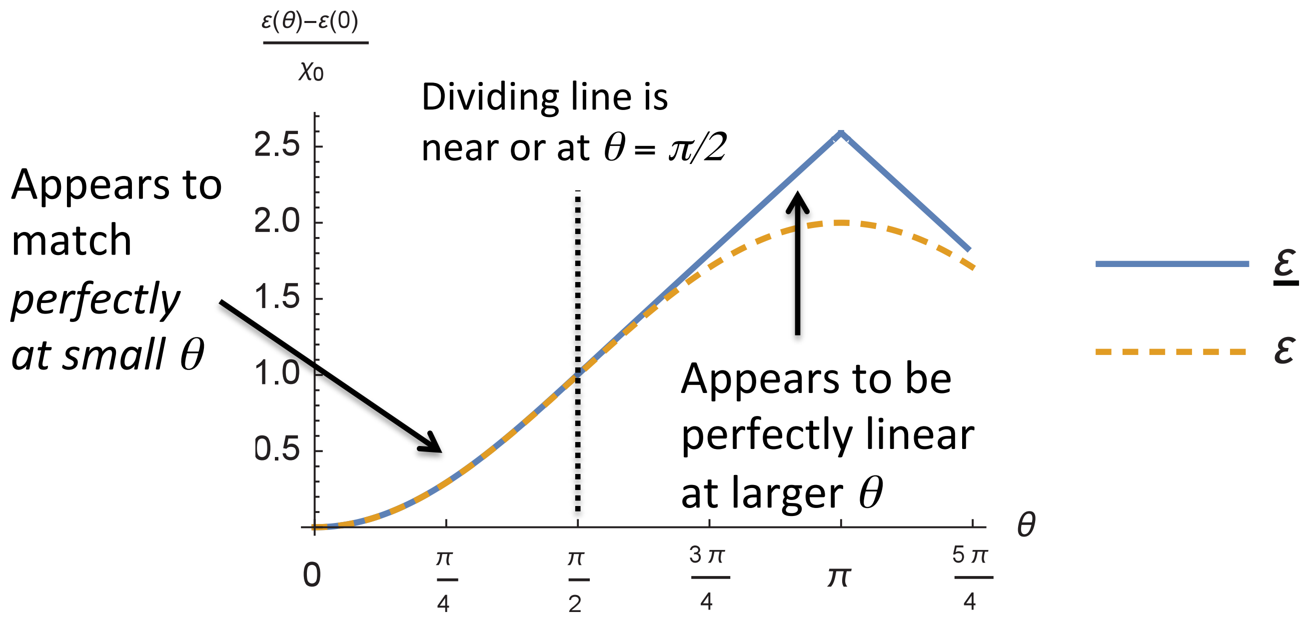
- This appears to show conclusively that the sign problem is so so severe that knowledge of the infinite volume energy density as a function of q is not enough to reconstruct $\varepsilon(\theta)$!!
- However, the situation turns out to be a bit more complicated.
  - Consider what happens if you improve numerics to be able to use VERY large volumes ($V\chi_0$=350)

- This appears to show conclusively that the sign problem is so so severe that knowledge of the infinite volume energy density as a function of q is not enough to reconstruct $\varepsilon(\theta)$!!
- However, the situation turns out to be a bit more complicated.
  - Consider what happens if you improve numerics to be able to use VERY large volumes ($V\chi_0$=350)

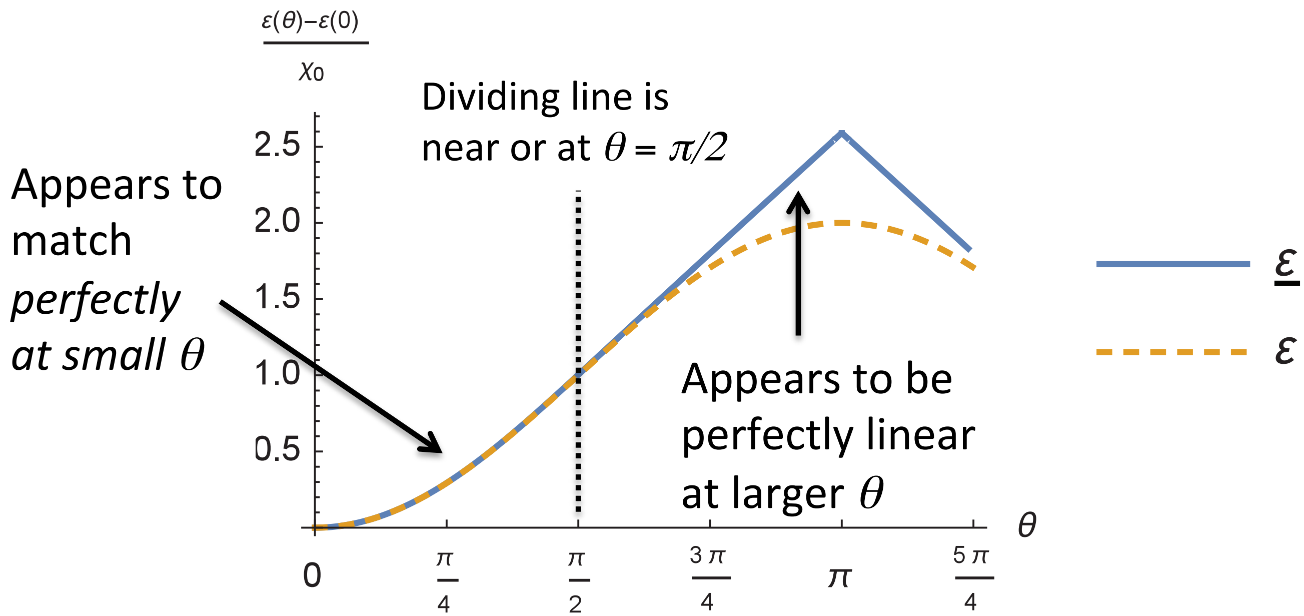




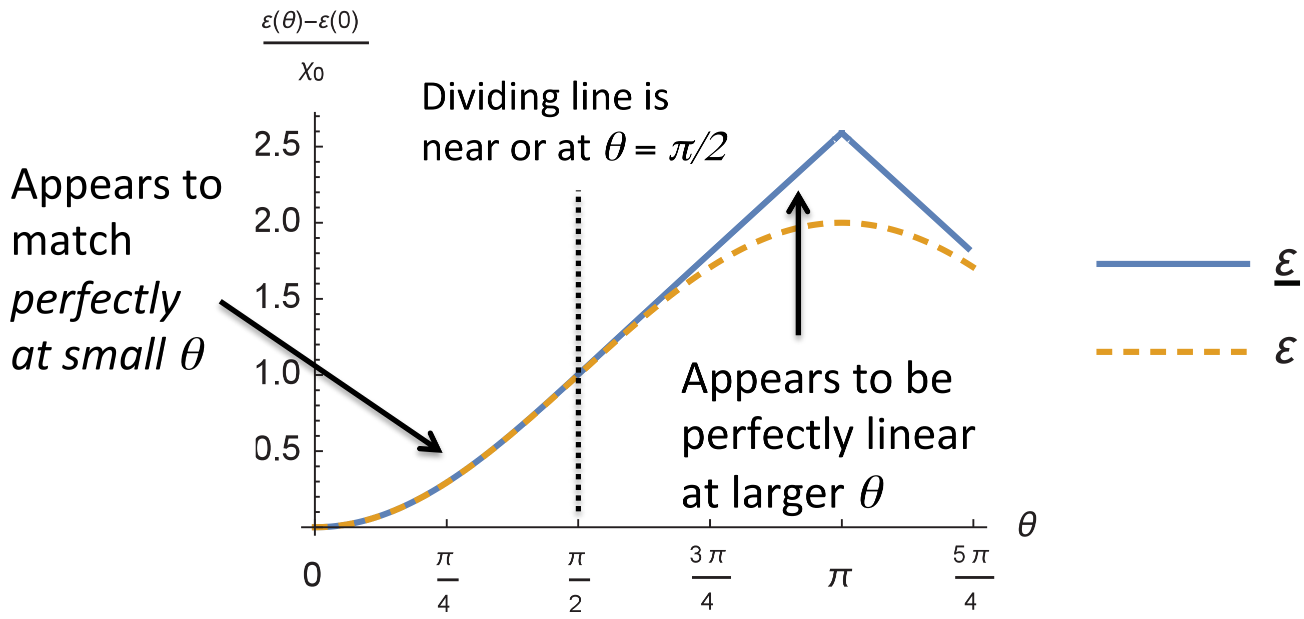
- **Some big questions:**



- **Some big questions:**
  - Why is it possible to reconstruct $\varepsilon(\theta)$ via direct summation of topological sectors in some regions of θ but not others?
    - Why doesn't exponential sensitivity of sign problem ruin things at small θ?
    - Why isn't exact knowledge of energy density as a function of topological charge density sufficient at larger θ?



- **Some big questions:**
  - Why is it possible to reconstruct $\varepsilon(\theta)$ via direct summation of topological sectors in some regions of θ but not others?
    - Why doesn't exponential sensitivity of sign problem ruin things at small θ?
    - Why isn't exact knowledge of energy density as a function of topological charge density sufficient at larger θ?
  - Why is $\theta = \pi/2$ the dividing line?



- **Some big questions:**
  - Why is it possible to reconstruct $\varepsilon(\theta)$ via direct summation of topological sectors in some regions of θ but not others?
    - Why doesn't exponential sensitivity of sign problem ruin things at small θ?
    - Why isn't exact knowledge of energy density as a function of topological charge density sufficient at larger θ?
  - Why is $\theta = \pi/2$ the dividing line?
  - Why is the function linear beyond this point?

Some Answers

- Why is $\theta = \pi/2$ the dividing line?

Some Answers

- Why is $\theta = \pi/2$ the dividing line?
  - This point is special because it is a point of inflection.
    - It can be shown in general (i.e. not just for dilute instanton gas) that below the first point of inflection in $\varepsilon(\theta)$, $\underline{\varepsilon}(\theta) = \varepsilon(\theta)$
      - For theories without points of inflection for entire range $-\pi < \theta < \pi$, $\underline{\varepsilon}(\theta) = \varepsilon(\theta)$ over the range.
      - Such theories must have a kink at $\theta = \pm\pi$

Some Answers

- Why is $\theta = \pi/2$ the dividing line?
  - This point is special because it is a point of inflection.
    - It can be shown in general (i.e. not just for dilute instanton gas) that below the first point of inflection in $\varepsilon(\theta)$, $\underline{\varepsilon}(\theta) = \varepsilon(\theta)$
      - For theories without points of inflection for entire range $-\pi < \theta < \pi$, $\underline{\varepsilon}(\theta) = \varepsilon(\theta)$ over the range.
      - Such theories must have a kink at $\theta = \pm\pi$
    - It can also be shown in general (i.e. not just for dilute instanton gas) that above the first point of inflection in $\varepsilon(\theta)$, $\underline{\varepsilon}(\theta)$ is linear

- Why does $\underline{\varepsilon}(\theta) = \varepsilon(\theta)$ below the point of inflection?
  - The dilute instanton gas gives a major hint:

  The energy densities are related by a generalized Legendre transformation when θ is below the point of inflection.

- Why does $\underline{\varepsilon}(\theta) = \varepsilon(\theta)$ below the point of inflection?
  - The dilute instanton gas gives a major hint:

    The energy densities are related by a generalized Legendre transformation when θ is below the point of inflection.

$$\tilde{\varepsilon}(q(\theta)) = \varepsilon(\theta) + i\,\theta\,q(\theta) \quad \text{with} \quad q(\theta) = i\frac{\partial \varepsilon(\theta)}{\partial \theta}$$

$$\varepsilon(\theta(q)) = \tilde{\varepsilon}(q) - i\,\theta(q)\,q \quad \text{with} \quad \theta(q) = -i\frac{\partial \tilde{\varepsilon}(q)}{\partial q} \ .$$

Like an ordinary Legendre transformation but with a critical factor of *i* different. This necessitates analytically continuing the energy density functions into the complex plane.

- Why does $\underline{\varepsilon}(\theta) = \varepsilon(\theta)$ below the point of inflection?
  - The dilute instanton gas gives a major hint:

  The energy densities are related by a generalized Legendre transformation when θ is below the point of inflection.

$$\tilde{\varepsilon}(q(\theta)) = \varepsilon(\theta) + i\,\theta\,q(\theta) \quad \text{with} \quad q(\theta) = i\frac{\partial\varepsilon(\theta)}{\partial\theta}$$

$$\varepsilon(\theta(q)) = \tilde{\varepsilon}(q) - i\,\theta(q)\,q \quad \text{with} \quad \theta(q) = -i\frac{\partial\tilde{\varepsilon}(q)}{\partial q} \ .$$

Like an ordinary Legendre transformation but with a critical factor of *i* different. This necessitates analytically continuing the energy density functions into the complex plane.

This Legendre transform can be shown to hold in the general case (not just the dilute instanton gas) below all points of inflection. The existence of this transformation ultimately implies $\underline{\varepsilon}(\theta) = \varepsilon(\theta)$

Above the point of inflection this generalized Legendre transformation fails—as it must:

Above the point of inflection this generalized Legendre transformation fails—as it must:

it maps q onto θ and θ back onto q.

Above the point of inflection this generalized Legendre transformation fails—as it must:

it maps q onto θ and θ back onto q.

Assume the generalized Legendre transformation exists and holds for both sides of a point of inflection. Then two distinct values of θ — one above the point of inflection and one below— would be mapped to the same value of q but this value of q cannot be mapped back onto two distinct values of θ yielding a contradiction.

Above the point of inflection this generalized Legendre transformation fails—as it must:

it maps q onto θ and θ back onto q.

Assume the generalized Legendre transformation exists and holds for both sides of a point of inflection. Then two distinct values of θ — one above the point of inflection and one below— would be mapped to the same value of q but this value of q cannot be mapped back onto two distinct values of θ yielding a contradiction.

Derivation of the generalized Legendre transformation and proof that it yields $\underline{\varepsilon}(\theta) = \varepsilon(\theta)$ below an inflection point and linearity above is somewhat complicated.

It will just be sketched here.

A magical Identity (which holds for $-\pi<\theta<\pi$)

$$\underline{\varepsilon}(\theta) = -\lim_{V\to\infty} \frac{1}{V} \log \left(\sum_Q e^{-\tilde{\varepsilon}(\frac{Q}{V})V} e^{i\theta Q} \right)$$

$$= -\lim_{V\to\infty} \frac{1}{V} \log \left(\int dQ\, e^{-\tilde{\varepsilon}(\frac{Q}{V})V} e^{i\theta Q} \right)$$

$$= -\lim_{V\to\infty} \frac{1}{V} \log \left(V \int dq\, e^{-V(\tilde{\varepsilon}(q)+i\theta q)} \right)$$

A magical Identity (which holds for $-\pi<\theta<\pi$)

$$\underline{\varepsilon}(\theta) = -\lim_{V\to\infty}\frac{1}{V}\log\left(\sum_Q e^{-\tilde{\varepsilon}(\frac{Q}{V})V}\, e^{i\theta Q}\right)$$

$$= -\lim_{V\to\infty}\frac{1}{V}\log\left(\int dQ\, e^{-\tilde{\varepsilon}(\frac{Q}{V})V}\, e^{i\theta Q}\right)$$

$$= -\lim_{V\to\infty}\frac{1}{V}\log\left(V\int dq\, e^{-V(\tilde{\varepsilon}(q)+i\theta q)}\right)$$

One can replace the sum by an integral. This seems trivial—but it is anything but trivial—recall that the Q=0 term in the sum is exponentially bigger than the total. **It really is magical—but can be proved!**

A magical Identity (which holds for $-\pi<\theta<\pi$)

$$\underline{\varepsilon}(\theta) = -\lim_{V\to\infty}\frac{1}{V}\log\left(\sum_Q e^{-\tilde{\varepsilon}(\frac{Q}{V})V}\,e^{i\theta Q}\right)$$

$$= -\lim_{V\to\infty}\frac{1}{V}\log\left(\int dQ\, e^{-\tilde{\varepsilon}(\frac{Q}{V})V}\,e^{i\theta Q}\right)$$

$$= -\lim_{V\to\infty}\frac{1}{V}\log\left(V\int dq\, e^{-V(\tilde{\varepsilon}(q)+i\theta q)}\right)$$

One can replace the sum by an integral. This seems trivial—but it is anything but trivial—recall that the Q=0 term in the sum is exponentially bigger than the total. **It really is magical—but can be proved!**

At large V this is the kind of integral for which the saddles point/stationary phase approximation becomes accurate—if a saddle point exists. If one does exist it will be on the imaginary axis.

Assuming a saddle point along the imaginary axis exists

$$\underline{\varepsilon}(\theta) = \tilde{\varepsilon}(q_\theta^{\mathrm{sp}}) - i\theta q_\theta^{\mathrm{sp}} \quad \textbf{with} \quad \left.\frac{\partial\left(\tilde{\varepsilon}(q) - i\theta q\right)}{\partial q}\right|_{q=q^{\mathrm{sp}}} = 0$$

This is the form of a generalized Legendre transformation!!
It is for $\underline{\varepsilon}$ rather ε however.

Assuming a saddle point along the imaginary axis exists

$$\underline{\varepsilon}(\theta) = \tilde{\varepsilon}(q_\theta^{\rm sp}) - i\theta q_\theta^{\rm sp} \quad \textbf{with} \quad \left.\frac{\partial\left(\tilde{\varepsilon}(q) - i\theta q\right)}{\partial q}\right|_{q=q^{\rm sp}} = 0$$

This is the form of a generalized Legendre transformation!!
It is for $\underline{\varepsilon}$ rather ε however.

But one can repeat the analysis for ε assuming a saddle point exists.

$$\begin{aligned}
\varepsilon(\theta) &= -\lim_{V\to\infty}\frac{1}{V}\log\left(\sum_Q e^{-\tilde{\varepsilon}(\frac{Q}{V},V)V}\, e^{i\theta Q}\right) \\
&= -\lim_{V\to\infty}\frac{1}{V}\log\left(\int dQ\, e^{-\tilde{\varepsilon}(\frac{Q}{V},V)V}\, e^{i\theta Q}\right) \\
&= -\lim_{V\to\infty}\frac{1}{V}\log\left(V\int dq\, e^{-V(\tilde{\varepsilon}(q,V)+i\theta q)}\right) \\
&= \tilde{\varepsilon}(q_\theta^{\rm sp}) - i\theta q_\theta^{\rm sp}
\end{aligned}$$

with

$$\lim_{V\to\infty}\left.\frac{\partial\left(\tilde{\varepsilon}(q,V) - i\theta q\right)}{\partial q}\right|_{q=q^{\rm sp}} = \left.\frac{\partial\left(\tilde{\varepsilon}(q) - i\theta q\right)}{\partial q}\right|_{q=q^{\rm sp}} = 0$$

Assuming a saddle point along the imaginary axis exists

$$\underline{\varepsilon}(\theta) = \tilde{\varepsilon}(q_\theta^{\rm sp}) - i\theta q_\theta^{\rm sp} \quad \textbf{with} \quad \left.\frac{\partial\left(\tilde{\varepsilon}(q) - i\theta q\right)}{\partial q}\right|_{q=q^{\rm sp}} = 0$$

This is the form of a generalized Legendre transformation!!
It is for $\underline{\varepsilon}$ rather ε however.

But one can repeat the analysis for ε assuming a saddle point exists.

$$\begin{aligned}\varepsilon(\theta) &= -\lim_{V\to\infty}\frac{1}{V}\log\left(\sum_Q e^{-\tilde{\varepsilon}(\frac{Q}{V},V)V}\, e^{i\theta Q}\right) \\ &= -\lim_{V\to\infty}\frac{1}{V}\log\left(\int dQ\, e^{-\tilde{\varepsilon}(\frac{Q}{V},V)V}\, e^{i\theta Q}\right) \\ &= -\lim_{V\to\infty}\frac{1}{V}\log\left(V\int dq\, e^{-V(\tilde{\varepsilon}(q,V)+i\theta q)}\right) \\ &= \tilde{\varepsilon}(q_\theta^{\rm sp}) - i\theta q_\theta^{\rm sp}\end{aligned}$$

with

$$\lim_{V\to\infty}\left.\frac{\partial\left(\tilde{\varepsilon}(q,V) - i\theta q\right)}{\partial q}\right|_{q=q^{\rm sp}} = \left.\frac{\partial\left(\tilde{\varepsilon}(q) - i\theta q\right)}{\partial q}\right|_{q=q^{\rm sp}} = 0$$

ε is given by a generalized Legendre transformation –the same one as $\underline{\varepsilon}$ assuming it has a saddle point !!

- **Thus ε and $\underline{\varepsilon}$ are equal when a saddle point exists to dominate the relevant integral for $\underline{\varepsilon}$.**
  - In these circumstances the generalized Legendre form emerges naturally.
  - This explains why $\varepsilon = \underline{\varepsilon}$ at small θ

- **Thus ε and $\underline{\varepsilon}$ are equal when a saddle point exists to dominate the relevant integral for $\underline{\varepsilon}$.**
  - In these circumstances the generalized Legendre form emerges naturally.
  - This explains why $\varepsilon=\underline{\varepsilon}$ at small θ

We also know quite a bit about when the generalized Legendre breaks down for $\underline{\varepsilon}$. It holds for all values of q for along the imaginary axis for which $\frac{\partial \underline{\varepsilon}}{\partial q}$ is well-defined and finite.

- **Thus ε and $\underline{\varepsilon}$ are equal when a saddle point exists to dominate the relevant integral for $\underline{\varepsilon}$.**
  - In these circumstances the generalized Legendre form emerges naturally.
  - This explains why $\varepsilon=\underline{\varepsilon}$ at small θ

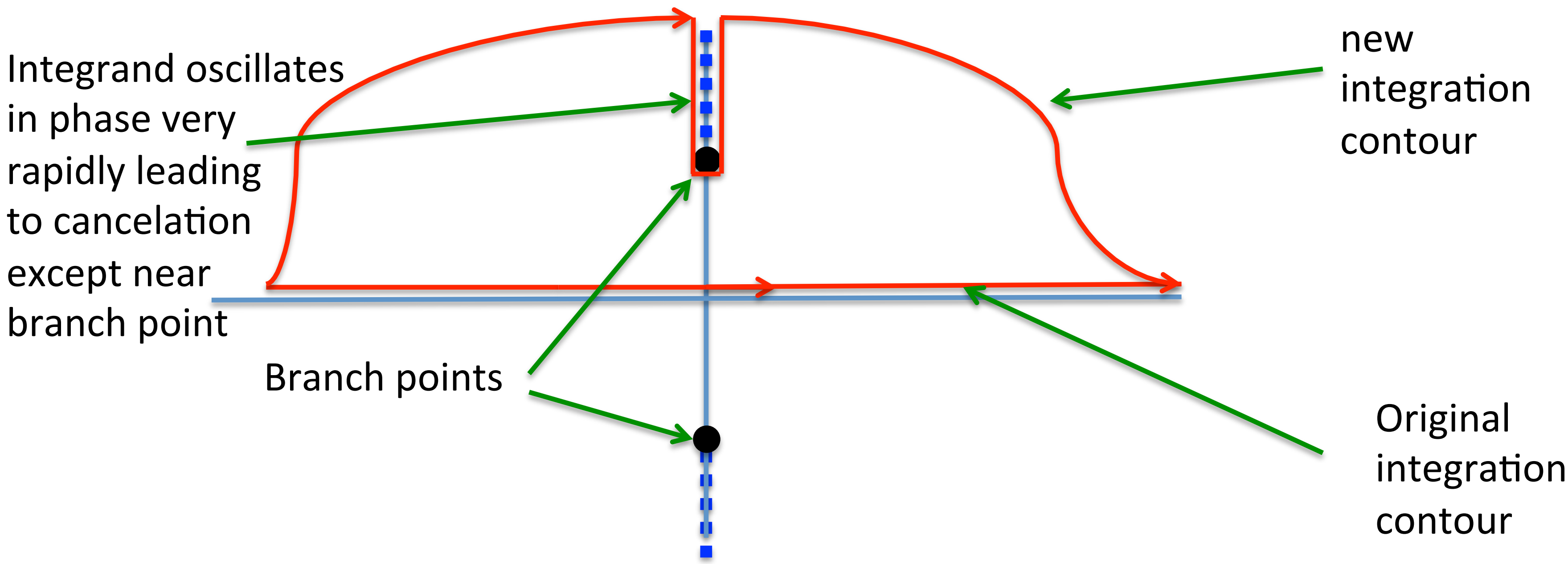
We also know quite a bit about when the generalized Legendre breaks down for $\underline{\varepsilon}$. It holds for all values of q for along the imaginary axis for which $\frac{\partial \underline{\varepsilon}}{\partial q}$ is well-defined and finite.

Thus, the structure will break down at points along the imaginary axis, where $\underline{\varepsilon}$ becomes nonanalytic and beyond which it is not well-defined. This occurs at branch points.

A theorem: Whenever a saddle point exists $\left.\frac{\partial^2 \varepsilon(\theta)}{\partial \theta^2}\,\frac{\partial^2 \tilde{\varepsilon}(q)}{\partial q^2}\right|_{q_\theta^{\rm sp}} = 1$

Thus as a point of inflection in $\varepsilon(\theta)$ is approached a nonanalytic point in q is also approached. It is easy to see that it is a branch point. One can also prove the converse—whenever a branch point is approached in q is a point of inflection in θ is approached.

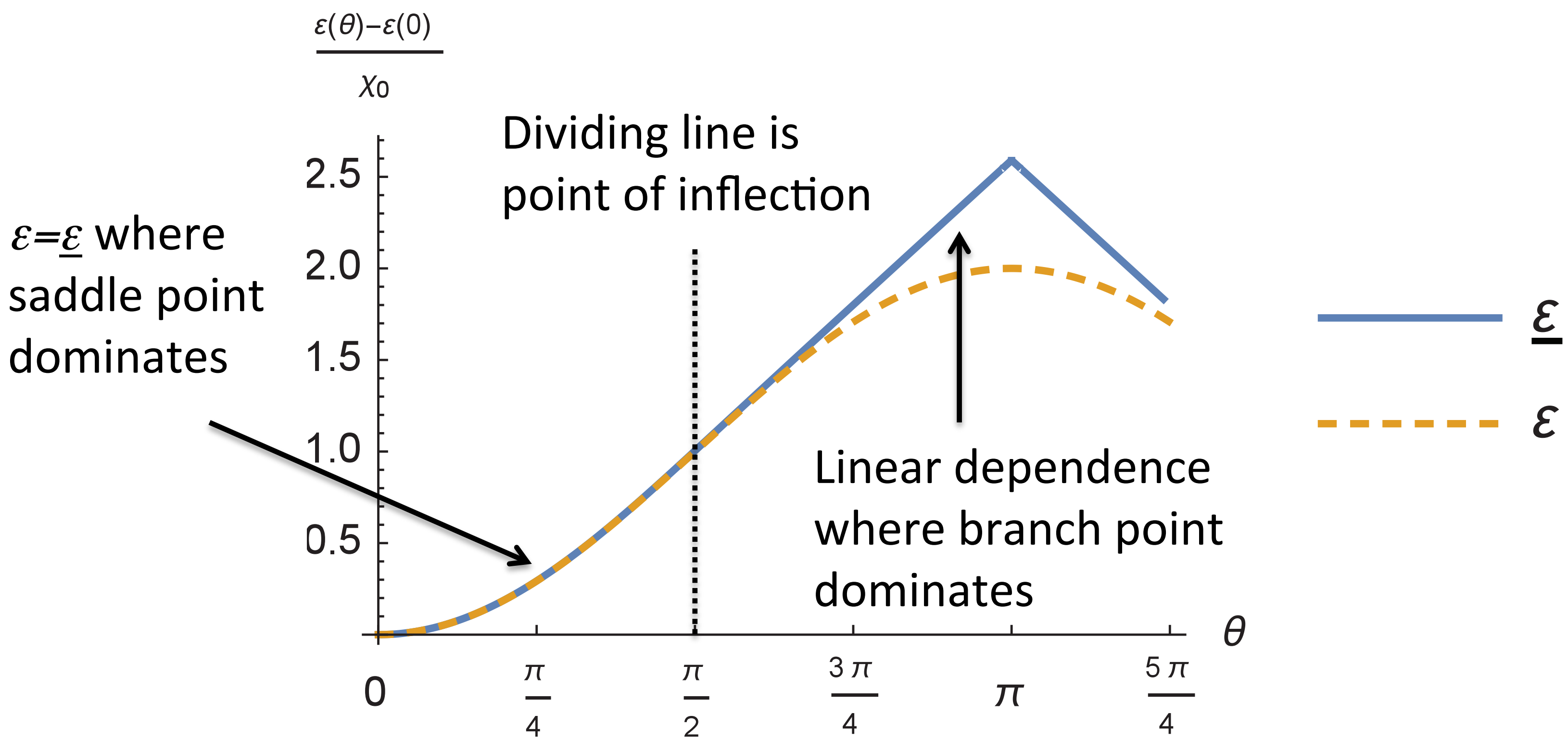
Note that just as saddle points can dominate integrals, so can branch points.



- **In the absence of a saddle point a branch point will dominate integral for $\underline{\varepsilon}$.**
  - Note that the position of branch point is independent of the value of θ.
  - Thus. where branch point dominates (i.e. beyond point of inflection in $\varepsilon(\theta)$), $\underline{\varepsilon}(\theta) = \tilde{\varepsilon}(q_\theta^{\mathrm{bp}}) - i\theta q_\theta^{\mathrm{bp}}$: a linear dependence

- **In the absence of a saddle point a branch point will dominate integral for $\underline{\varepsilon}$.**
  - Note that the position of branch point is independent of the value of θ.
  - Thus. where branch point dominates (i.e. beyond point of inflection in $\varepsilon(\theta)$), $\underline{\varepsilon}(\theta) = \tilde{\varepsilon}(q_\theta^{\rm bp}) - i\theta q_\theta^{\rm bp}$: a linear dependence

Qualitative behavior seen in dilute instanton gas, follows from this general analysis.



All of this analysis depended on the magical identity which allowed the replacement of the sum by an integral.

Thus it is useful to see why this identity works.

As a first step one can equate a periodic sum sum of δ functions to a sum of complex exponentials

$$\sum_{n=-\infty}^{\infty} \delta(x-n) = \sum_{k=-\infty}^{\infty} \exp\left(i2\pi kx\right)$$

From this it follows that

A sum

$$\sum_{n=-\infty}^{\infty} \exp\left(-\lambda f\left(\frac{n}{\lambda}\right)\right) \exp(i\theta\, n) =$$

$$\sum_{k=-\infty}^{\infty} \int_{-\infty}^{\infty} dx \exp\left(-\lambda f\left(\frac{x}{\lambda}\right)\right) \exp(i\theta x) \exp(i2\pi kx) =$$

A sum of integrals

$$\sum_{k=-\infty}^{\infty} \lambda \int_{-\infty}^{\infty} dy \exp\left(-\lambda\left(f(y) + i(\theta + 2\pi k)\, y\right)\right)$$

A sum

$$\sum_{n=-\infty}^{\infty} \exp\left(-\lambda f\left(\frac{n}{\lambda}\right)\right)\exp(i\theta\, n) =$$

$$\sum_{k=-\infty}^{\infty} \int_{-\infty}^{\infty} dx \exp\left(-\lambda f\left(\frac{x}{\lambda}\right)\right)\exp(i\theta x)\exp(i2\pi k x) =$$

A sum of integrals

$$\sum_{k=-\infty}^{\infty} \lambda \int_{-\infty}^{\infty} dy \exp\left(-\lambda\left(f(y) + i(\theta + 2\pi k)\, y\right)\right)$$

Note each integral is of the type where at large λ the integral is dominated by a special point—eg. a saddle point or branch point and, up to power law corrections, they are given at large λ by $\exp[-\lambda\{f(y_{sp}+iy_{sp}(\theta+2\pi k)\}]$.

The sum of integrals is exponentially dominated by the term with the smallest exponential; this is typically the k=0 term when $-\pi<\theta<\pi$. In infinite λ limit only one term contributes.

$$\lim_{\lambda \to 0} \frac{1}{\lambda} \log \left(\sum_{n=-\infty}^{\infty} \exp\left(-\lambda f\left(\frac{n}{\lambda}\right)\right) \exp(i\theta\, n) \right)$$

$$= \lim_{\lambda \to 0} \frac{1}{\lambda} \log \left(\lambda \int_{-\infty}^{\infty} dy \exp\left(-\lambda\left(f(y) + i(\theta)\, y\right)\right) \right)$$

The sum is replaced by a single integral; the magical identity is established!

This validates the analysis and justifies the previous conclusions about the role of the subtle interplay of the sign problem and the infinite volume limit.

Conclusions

- **The answer to the question "Does the infinite volume expression for the energy density as a function of *q* yield the correct $\varepsilon(\theta)$ via direct summation over topological sectors or is the sign problem so severe as to prevent this?" is "It depends".**

Conclusions

- **The answer to the question "Does the infinite volume expression for the energy density as a function of q yield the correct $\varepsilon(\theta)$ via direct summation over topological sectors or is the sign problem so severe as to prevent this?" is "It depends".**
  - **Below a point of inflection in $\varepsilon(\theta)$ the answer is "yes".**
    - **This is tied to the existence of a saddle point which dominates the relevant integral and yields a generalized Legendre transformation**

Conclusions

- **The answer to the question "Does the infinite volume expression for the energy density as a function of q yield the correct $\varepsilon(\theta)$ via direct summation over topological sectors or is the sign problem so severe as to prevent this?" is "It depends".**
  - **Below a point of inflection in $\varepsilon(\theta)$ the answer is "yes".**
    - **This is tied to the existence of a saddle point which dominates the relevant integral and yields a generalized Legendre transformation**
  - **Above the point of inflection in $\varepsilon(\theta)$ the answer is "no".**
    - **$\varepsilon(\theta)$ is linear in this region. This is tied to the the relevant integral being dominated by a branch point.**

Conclusions

- **The answer to the question "Does the infinite volume expression for the energy density as a function of *q* yield the correct $\varepsilon(\theta)$ via direct summation over topological sectors or is the sign problem so severe as to prevent this?" is "It depends".**
  - **Below a point of inflection in $\varepsilon(\theta)$ the answer is "yes".**
    - **This is tied to the existence of a saddle point which dominates the relevant integral and yields a generalized Legendre transformation**
  - **Above the point of inflection in $\varepsilon(\theta)$ the answer is "no".**
    - **$\underline{\varepsilon}(\theta)$ is linear in this region. This is tied to the the relevant integral being dominated by a branch point.**
  - **Theories without points of inflection in $\varepsilon(\theta)$ for $0<\theta<\pi$, the answer is "yes"for whole region; such theories have kinks at $\theta=\pm\pi$.**

- In this talk subtleties of the infinite volume limit for a calculation with a sign problem were clarified. Can we use this understanding to make money on actual calculations of in QCD?

- In this talk subtleties of the infinite volume limit for a calculation with a sign problem were clarified. Can we use this understanding to make money on actual calculations of in QCD?

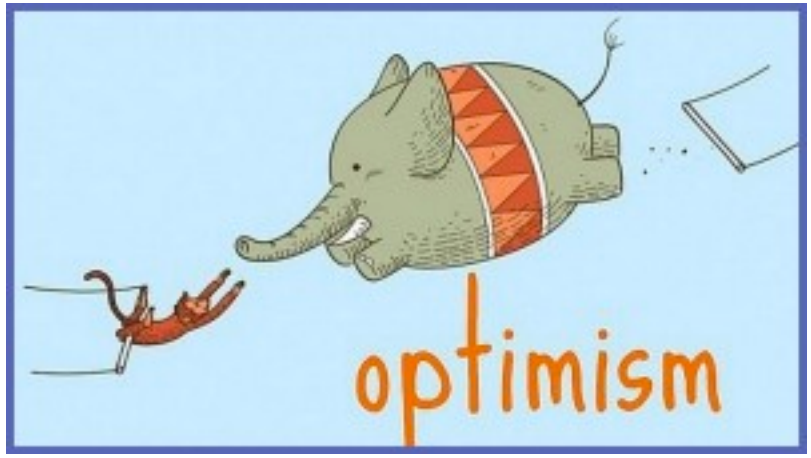
- In this talk subtleties of the infinite volume limit for a calculation with a sign problem were clarified. Can we use this understanding to make money on actual calculations in QCD?

- Are there implications for efficient calculations of $\varepsilon(\theta)$?

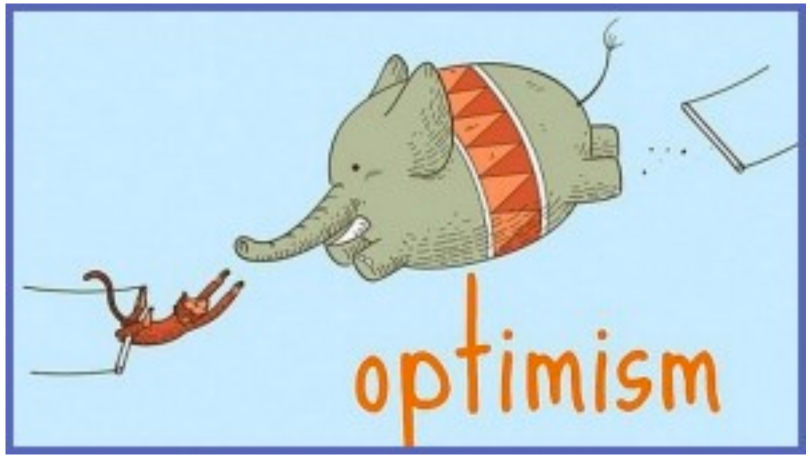
- In this talk subtleties of the infinite volume limit for a calculation with a sign problem were clarified. Can we use this understanding to make money on actual calculations in QCD?

- Are there implications for efficient calculations of $\varepsilon(\theta)$?

An optimistic take on things

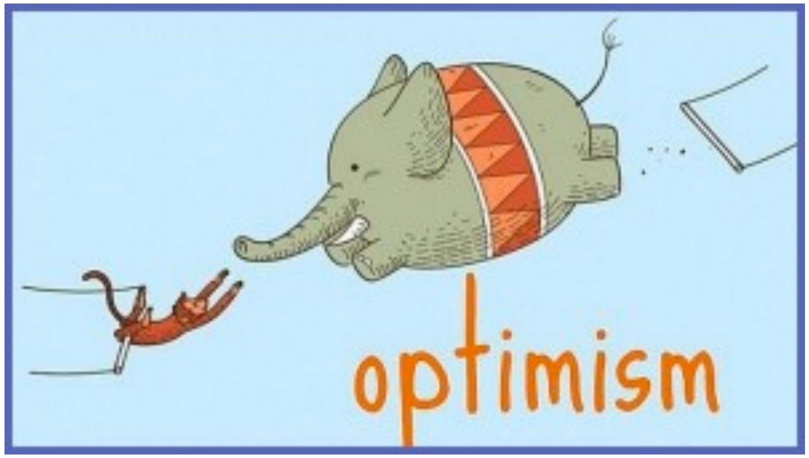


An optimistic take on things



Despite the apparent need for exponentially accuracy, we can drop power law terms in some circumstances and still get a reasonable result—at least in some cases.

An optimistic take on things



Despite the apparent need for exponentially accuracy, we can drop power law terms in some circumstances and still get a reasonable result—at least in some cases.

This suggests that maybe with enough cleverness perhaps the problem can be solved without the need for exponential accuracy.

- **Can we make money on actual calculations in QCD?**

- **Are there implications for calculations with a chemical potential?**

- Can we make money on actual calculations in QCD?

- Are there implications for calculations with a chemical potential?