# The Polaron model with cut-off in the strong coupling regime

## Introduction

In the following, we study the polaron Hamilton operator and its ground state energy in the strong coupling limit, which is a model to describe the interaction of a charged particle moving trough a polarized medium. The investigation is based on the Feynman-Kac formula, which allows one to express the kernel of the imaginary time Schrödinger semi-group $\langle x|e^{-\beta(-\Delta+V)}|y\rangle$ as an expectation value. This has the advantage, that one can now apply tools from probability theory to the original problem. In the case of the polaron, Donsker and Varadhan [3] used the theory of large deviations to verify that the asymptotic behaviour of the ground state energy $E_0(\alpha)$ is given by

$$E_0(\alpha) \approx -\alpha^2 \gamma_p,$$

with a suitable constant $\gamma_p$. Their proof is based on the zero temperature limit $\beta \to \infty$. In this work we rather follow the work of [1], by performing the infinite temperature limit $\beta \to 0$. Note that in both cases, the result is not quantitative, in the sense that one obtains the limit $\alpha^{-2} E_0(\alpha)$ but not the one of $E_0(\alpha) + \alpha^2 \gamma_p$.
The goal of this write-up is, to improve the asymptotic result in the case of the polaron model with cut-off modes. We will be able to verify that $\alpha^{-q}\left(E_0(\alpha) + \alpha^2\gamma_p\right) \underset{\alpha\to 0}{\longrightarrow} 0$ for all $q > \frac{2}{3}$. It is crucial, that we discretize the original model, written in strong coupling units. Otherwise we cannot expect the finite model to feature an $\alpha^2$ asymptotic. It is worth mentioning, that the strong coupling units correspond to a semi-classical treatment of the polarized medium.

## The Model

We want to investigate the asymptotic behaviour of the ground state energy $E_0(\alpha)$ in the limit $\alpha \to \infty$ of the Hamilton operator $H_\alpha := \alpha^2 \mathfrak{h}_\alpha$, where

$$\mathfrak{h}_\alpha := -\frac{1}{2}\Delta + \frac{\epsilon_\alpha |x|^2}{2\alpha^4} + \sum_{k\in\Lambda} b_k^\dagger b_k + n^{-\frac{3}{2}} \sum_{k\in\Lambda} |k|^{-1}\left(e^{ikx} b_k^\dagger + e^{-ikx} b_k\right),$$

and $\Lambda := \{k \in \left(n^{-1}\mathbb{Z}\right)^3 : 0 < |k| \leq m\} \subset \mathbb{R}^3$. Note that $b_k$ should satisfies the $\alpha$ dependent CCR $[b_k, b_l^\dagger] = \alpha^{-2}\delta_{k,l}I$. Let us define $a_k := \alpha\ b_k$, which then satisfy the $\alpha$ independent CCR $[a_k, a_l^\dagger] = \delta_{k,l}I$. As a motivation, this Hamilton operator emerges, if one discretizes the Polaron Hamilton operator

$$-\frac{1}{2}\Delta + \frac{\epsilon_\alpha|x|^2}{2} + \int a_k^\dagger a_k\ \mathrm{d}k + \sqrt{\alpha}\int |k|^{-1}\left(e^{ikx}a_k^\dagger + e^{-ikx}a_k\right)\ \mathrm{d}k,$$

rewritten in strong coupling units. The harmonic potential $\frac{\epsilon_\alpha|x|^2}{2}$ is added for technical reasons, and we assume that $\epsilon_\alpha$ does not get too small or too large, to be precise we assume $\alpha^{-N} \leq \epsilon_\alpha \leq \alpha^{\frac{4}{3}}$ for some $N \geq 0$. Note that this especially includes the constant case $\epsilon_\alpha = \epsilon$.

If we transform the Hamiltonian according to the unitary map $T\psi(x) := \alpha^{-\frac{3}{2}}\ \psi(\alpha^{-1}x)$, we obtain

$$T^{-1}\ H_\alpha\ T = H^{(\alpha)},$$

with

$$H^{(\alpha)} = -\frac{1}{2}\Delta + \frac{\epsilon_\alpha|x|^2}{2} + \sum_{k\in\Lambda} a_k^\dagger a_k + \alpha\ n^{-\frac{3}{2}}\sum_{k\in\Lambda}|k|^{-1}\left(e^{ik\alpha x}a_k^\dagger + e^{-ik\alpha x}a_k\right).$$

By applying the Feynman-Kac formula and integrating out the field variables, we obtain

$$\langle x|\mathrm{tr}_{\mathcal{F}}\ e^{-\beta H^{(\alpha)}}|x\rangle = \left(\mathrm{tr}\ e^{-\beta\sum_{k\in\Lambda}a_k^\dagger a_k}\right)(2\pi\beta)^{-\frac{3}{2}}$$

$$\times\mathbb{E}_{0,0}\ e^{-\frac{\epsilon_\alpha}{2}\int_0^\beta|X_s+x|^2\ \mathrm{d}s}e^{\alpha^2\int_0^\beta\int_0^\beta\frac{\cosh\left(|t-s|-\frac{\beta}{2}\right)}{2\sinh\left(\frac{\beta}{2}\right)}f(\alpha X_s,\alpha X_t)\ \mathrm{d}s\mathrm{d}t},$$

where $X_s$ is a Brownian bridge from $(0,0)$ to $(\beta,0)$ and

$$f(x,y) := n^{-3}\sum_{k\in\Lambda}|k|^{-2}e^{ik(x-y)}.$$

By the substitution $\alpha X_s \leftrightarrow X_{\alpha^2 s}$ and rescaling in the integral as well as explicitly expressing the trace over the Fock space, we obtain

$$\langle x|\mathrm{tr}_{\mathcal{F}}\ e^{-\beta H^{(\alpha)}}|x\rangle$$

$$= \left(1-e^{-\beta}\right)^{-|\Lambda|}(2\pi\beta)^{-\frac{3}{2}}\ \cdot\mathbb{E}_{0,0}\ e^{-\frac{\epsilon_\alpha}{2}\int_0^\beta|X_s+x|^2\ \mathrm{d}s}e^{\alpha^{-2}\int_0^{\alpha^2\beta}\int_0^{\alpha^2\beta}\frac{\cosh\left(\alpha^{-2}|t-s|-\frac{\beta}{2}\right)}{2\sinh\left(\frac{\beta}{2}\right)}f(X_s,X_t)\ \mathrm{d}s\mathrm{d}t}$$

$$\leq \left(1-e^{-\beta}\right)^{-|\Lambda|}(2\pi\beta)^{-\frac{3}{2}}\ \cdot\mathbb{E}_{0,0}\ e^{-\frac{\epsilon_\alpha}{2}\int_0^\beta|X_s+x|^2\ \mathrm{d}s}e^{\alpha^{-2}\int_0^{\alpha^2\beta}\int_0^{\alpha^2\beta}\frac{1}{2}\coth\left(\frac{\beta}{2}\right)f(X_s,X_t)\ \mathrm{d}s\mathrm{d}t},$$

where we used $\cosh\left(|\tau|-\frac{\beta}{2}\right)\leq\cosh\left(\frac{\beta}{2}\right)$ in the last inequality. By applying Jensen's inequality to the probability measure $\beta^{-1}\int_0^\beta \cdot \,\mathrm{d}x$, the convex function $x\mapsto e^{-\beta x}$ and the random variable $s\mapsto\frac{\epsilon_\alpha}{2}|X_s+x|^2$, we obtain $e^{-\int_0^\beta\frac{\epsilon_\alpha}{2}|X_s+x|^2\,\mathrm{d}s}\leq\beta^{-1}\int_0^\beta e^{-\beta\frac{\epsilon_\alpha}{2}|X_s+x|^2}\,\mathrm{d}s$. Therefore, integration over the $x$ variable yields

$$\int_{\mathbb{R}} e^{-\frac{\epsilon_\alpha}{2}\int_0^\beta|X_s+x|^2\,\mathrm{d}s}\,\mathrm{d}x\leq\beta^{-1}\int_0^\beta\int_{\mathbb{R}} e^{-\beta\frac{\epsilon_\alpha}{2}|X_s+x|^2}\,\mathrm{d}x\,\mathrm{d}s=\int_{\mathbb{R}} e^{-\beta\frac{\epsilon_\alpha}{2}|x|^2}\,\mathrm{d}x=\left(\frac{\epsilon_\alpha\beta}{2\pi}\right)^{-\frac{3}{2}}.$$

With this at hand, we can estimate the trace of the semi-group by

$$\begin{aligned}\operatorname{tr} e^{-\beta H^{(\alpha)}}&\leq\left(1-e^{-\beta}\right)^{-|\Lambda|}(\epsilon_\alpha\beta^2)^{-\frac{3}{2}}\cdot\mathbb{E}_{0,0}\, e^{\alpha^2\frac{1}{2}\coth\left(\frac{\beta}{2}\right)\alpha^{-4}\int_0^{\beta\alpha^2}\int_0^{\beta\alpha^2}f(X_s,X_t)\,\mathrm{d}s\mathrm{d}t}\\&=\left(1-e^{-\beta}\right)^{-|\Lambda|}(\epsilon_\alpha\beta^2)^{-\frac{3}{2}}\cdot\mathbb{E}_{0,0}\, e^{\alpha^2\beta F_\beta[\tau^{(\alpha^2\beta)}]},\end{aligned}$$

where $\tau^{(T)}:=T^{-1}\int_0^T\delta_{X_s}\,\mathrm{d}s$ is the normalized occupation time of $X_s$ and

$$F_\beta[\mu]:=\frac{\beta}{2}\coth\left(\frac{\beta}{2}\right)\int n^{-3}\sum_{k\in\Lambda}|k|^{-2}e^{ik(x-y)}\,(\mu\otimes\mu)\,(\mathrm{d}x,\mathrm{d}y).$$

Consequently,

$$e^{-\beta E_0(\alpha)}\leq\operatorname{tr} e^{-\beta H^{(\alpha)}}\leq\left(1-e^{-\beta}\right)^{-|\Lambda|}(\epsilon_\alpha\beta^2)^{-\frac{3}{2}}\mathbb{E}_{0,0}\, e^{\alpha^2\beta F_\beta[\tau^{(\alpha^2\beta)}]}.$$

With the definition $F_{\mathrm{pol}}(\alpha,\beta):=-\beta^{-1}\log\mathbb{E}_{0,0}\, e^{\alpha^2\beta F_\beta[\tau^{(\alpha^2\beta)}]}$, we can write this as

$$E_0(\alpha)\geq F_{\mathrm{pol}}(\alpha,\beta)+|\Lambda|\beta^{-1}\log\left(1-e^{-\beta}\right)+\frac{3}{2}\beta^{-1}\log\left(\epsilon_\alpha\beta^2\right).$$

The main effort will be, to verify the following theorem:

**Theorem 0.1.** *Let $T_0>0$. For $\alpha^2\beta\geq T_0$, the free Polaron energy can be estimated by*

$$F_{\mathrm{pol}}(\alpha,\beta)+\alpha^2\gamma_p(\beta)\geq-C\beta^{-1}\log(\alpha^2\beta)-C'\beta^{-1},$$

*where we define*

$$\gamma_p(\beta):=\sup_{\|\phi\|_{L^2}=1}\left(\frac{\beta}{2}\coth\left(\frac{\beta}{2}\right)\int\int n^{-3}\sum_{k\in\Lambda}|k|^{-2}e^{ik(x-y)}|\phi(x)|^2|\phi(y)|^2\,\mathrm{d}x\mathrm{d}y-\int|\nabla\phi(x)|^2\,\mathrm{d}x\right).$$

Furthermore, we will show:

**Theorem 0.2.** *There exists a constant $C$, such that for all $\alpha > 0$, we can estimate*

$$E_0(\alpha) + \alpha^2 \gamma_p \leq C\sqrt{\epsilon_\alpha}.$$

From this, we can deduce the following corollary:

**Corolary 0.3.** *(Estimate on the ground state energy)*

*Let $\alpha^{-N} \leq \epsilon_\alpha \leq \alpha^{\frac{4}{3}} \log \alpha$ for some $N \in \mathbb{N}$. Then there exist an $\alpha_0$ and constants $A, B$ such that for all $\alpha \geq \alpha_0$*

$$|E_0(\alpha) + \alpha^2 \gamma_p| \leq C \ \alpha^{\frac{2}{3}} \log \alpha,$$

*where $\gamma_p := \gamma_p(0)$.*

*Proof.* Using the theorem above, we obtain

$$\begin{aligned}
&E_0(\alpha) + \alpha^2 \gamma_p \\
&\geq F_{\text{pol}}(\alpha, \beta) + \alpha^2 \gamma_p(\beta) + (\alpha^2 \gamma_p(1) - \alpha^2 \gamma_p(\beta)) + |\Lambda| \beta^{-1} \log \left(1 - e^{-\beta}\right) + \frac{3}{2} \beta^{-1} \log \left(\epsilon_\alpha \beta^2\right) \\
&\geq C_1 \beta^{-1} \log(\beta) - C_2 \beta^{-1} \log(\alpha\beta) - C_4 \beta^{-1} + \alpha^2 (\gamma_p(1) - \gamma_p(\beta)).
\end{aligned}$$

Let the functional value of $\phi$ be $\epsilon$ close to $\gamma_p(\beta)$. Since $f$ is bounded by some constant $D$ and $\|\phi\|_{L^2} = 1$, we have for all $\epsilon > 0$ (and therefore also for $\epsilon = 0$)

$$\begin{aligned}
&\gamma_p(1) \geq \int \int f(x,y) |\phi(x)|^2 |\phi(y)|^2 \ \mathrm{d}x \mathrm{d}y - \int |\nabla \phi(x)|^2 \ \mathrm{d}x \\
&\geq \frac{\beta}{2} \coth \left(\frac{\beta}{2}\right) \int \int f(x,y) |\phi(x)|^2 |\phi(y)|^2 \ \mathrm{d}x \mathrm{d}y - \int |\nabla \phi(x)|^2 \ \mathrm{d}x - \left(\frac{\beta}{2} \coth \left(\frac{\beta}{2}\right) - 1\right) D \\
&= \gamma_p(\beta) - \epsilon - \left(\frac{\beta}{2} \coth \left(\frac{\beta}{2}\right) - 1\right) D_1.
\end{aligned}$$

Note that $\frac{\beta}{2} \coth \left(\frac{\beta}{2}\right) - 1 \leq D_2 \beta^2$ for a large enough constant $D_2$. Hence,

$$\gamma_p(1) \geq \gamma_p(\beta) - C_3 \beta^2.$$

Together with $\alpha^{-N} \leq \epsilon_\alpha$, this leads to the inequality

$$E_0(\alpha) + \alpha^2 \gamma_p \geq C_1 \beta^{-1} \log(\beta) - C_2 \beta^{-1} \log(\alpha\beta) - C_3 \alpha^2 \beta^2 - C_4 \beta^{-1}$$

for all $\alpha, \beta$. With the choice $\beta(\alpha) := \alpha^{-\frac{2}{3}}$ we obtain for all $\alpha$ (note that still $\alpha^2 \beta \to \infty$)

$$E_0(\alpha) + \alpha^2 \gamma_p \geq -C_1' \alpha^{\frac{2}{3}} \log(\alpha) - C_2' \alpha^{\frac{2}{3}} \log(\alpha) - C_3' \alpha^{\frac{2}{3}}.$$

The upper bound follows immediately from Theorem 0.2 together with the assumption $\epsilon_\alpha \leq \alpha^{\frac{4}{3}} \log \alpha$. ∎

## Reduction to a finite dimensional Problem

**Definition 0.4.** *Let $L \subset \Lambda$ such that $\Lambda = L \dot\cup (-L)$. Then we define the random vector $V := (V_k)_{k\in\Lambda} : \mathcal{M}(\mathbb{R}^3) \to \mathbb{R}^\Lambda$ as*

$$V_k(\tau) := \begin{cases} \int \cos(k \cdot x)\ \tau(\mathrm{d}x),\ k \in L \\ \int \sin(k \cdot x)\ \tau(\mathrm{d}x),\ k \in -L. \end{cases}$$

*Furthermore, let $f_\beta : \mathbb{R}^\Lambda \to \mathbb{R}$ bet defined as*

$$f_\beta(v) := \beta \coth\left(\frac{\beta}{2}\right) n^{-3} \sum_{k\in\Lambda} |k|^{-2} |v_k|^2.$$

Note that we can write the random variable $F_\beta$ as $F_\beta = f_\beta \circ V$. Therefore, we can reduce the original problem to a finite dimensional one. To make this precise, we will define a measure on a subset of $\mathbb{R}^\Lambda$ and investigate the large deviations of the random variable $f_\beta$ instead of $F_\beta$. For convenience, we will denote with $A^\beta_{j,n} := f_\beta^{-1}\left((\frac{j}{n}, \infty)\right)$ the upper level sets of $f_\beta$.

**Definition 0.5.** *We define on $\mathbb{R}^\Lambda$, equipped with the Borel algebra, the probability measure $\mathbb{P}_T := \mathrm{Law}\left(V \circ \tau^{(T)}\right)$. Note that for a bounded and measurable function $f : \mathbb{R}^\Lambda \to \mathbb{R}$*

$$\int f\ \mathrm{d}\mathbb{P}_T = \mathbb{E}_{0,0}\ f\left(V \circ \tau^{(T)}\right).$$

*Especially,*

$$\mathbb{E}_{0,0}\ e^{\alpha^2 \beta F_\beta[\tau^{(\alpha^2\beta)}]} = \int e^{\alpha^2\beta f_\beta}\ \mathrm{d}\mathbb{P}_{\alpha^2\beta}.$$

We have seen in the introduction that the Dirichlet form $\mathcal{E}[\phi] := \int |\nabla\phi|^2\ \mathrm{d}x$ plays an important role. By identifying a function $\phi$ with the measure $\mathrm{d}\mu = \phi^2 \mathrm{d}x$ we can lift the Dirichlet form to a functional defined on the space $\mathcal{M}(\mathbb{R}^3)$. This form is then usually called the Fisher information. Furthermore, we use the random vector $V$ to transport this definition to $\mathbb{R}^\Lambda$.

**Definition 0.6.** *We define the functional $I_{\mathrm{Fisher}} : \mathcal{M}(\mathbb{R}^3) \to \mathbb{R} \cup \{+\infty\}$ as*

$$I_{\mathrm{Fisher}}[\tau] := \begin{cases} \int |\nabla\phi|^2\ \mathrm{d}x,\ \mathrm{d}\tau = \phi^2 \mathrm{d}x \textit{ with } \phi \geq 0 \\ +\infty,\ \textit{otherwise.} \end{cases}$$

*With this at hand, we define $I : \mathbb{R}^\Lambda \to \mathbb{R} \cup \{+\infty\}$ by the formula*

$$I[v] := \inf_{\tau : V(\tau) = v} I_{\mathrm{Fisher}}[\tau].$$

*Furthermore, let us denote with $\Omega := [I < +\infty]$ the set where $I$ is finite and with $G := I^*$ the convex conjugate of $I$, i.e.*

$$G(a) := \sup_{v \in \mathbb{R}^\Lambda} \left( a \cdot v - I[v] \right).$$

**Lemma 0.7.** *Let $f : \mathbb{R}^\Lambda \to \mathbb{R}$. Then we have the identity*

$$\sup_{\tau \in \mathcal{M}(\mathbb{R}^3)} \Big( f\left(V(\tau)\right) - I_{\mathrm{Fisher}}[\tau] \Big) = \sup_{v \in \mathbb{R}^\Lambda} \Big( f(v) - I[v] \Big).$$

## Lower Bound on the Ground State Energy

The following three theorems are the milestones in proving Theorem 0.1. We follow the strategy of [4], with the difference that we try to obtain a more quantitative version of the asymptotic results. The proof of each of the theorems is dependent on the statement of the subsequent one. The proofs will also involve crucial lemmas from the last section. In the first theorem we will get our input information on the measure $\mathbb{P}_T$ by applying Feynman-Kac, which will yield us a Laplace principle for linear functions. From this we deduce the next theorem, which is a quantitative version of a Large Deviation result for the occupation time measure $\mathbb{P}_T$, see also [2]. This can then be used to prove the last theorem by a modified version of the Varadhan Lemma (see [5] Theorem 27.10), which is a gain a Laplace principle, but this time for the functional $f_\beta$ in which we are actually interested.

**Theorem 0.8.** *Let $T_0 > 0$. Then for all $a \in \mathbb{R}^\Lambda$ and $T \geq T_0$, we have the estimate*

$$\log \int e^{T(a\cdot v)}\ \mathbb{P}_T(\mathrm{d}v) - T\ G(a) \leq \frac{3}{2}\log T + c_1|a|.$$

**Theorem 0.9.** *Recall the definition $A^\beta_{j,n} := f_\beta^{-1}\left((\frac{j}{n},\infty)\right)$. Then, for all $M > 0$, $\beta \leq \beta_0$, $T \geq T_0$ and all $j, n$ such that $\inf_{v\in A^\beta_{j,n}} I[v] \leq M$, we have the uniform estimate*

$$\log \mathbb{P}_T\left(A^\beta_{j+1,n}\right) + T \inf_{v\in A^\beta_{j,n}} I[v] \leq c_2\log(T) + c_3\log(n) + c_4.$$

**Theorem 0.10.** *For all $\beta \leq \beta_0$ and $T \geq T_0$, we can estimate the integral*

$$\log \int e^{Tf_\beta}\ \mathrm{d}\mathbb{P}_T - T\sup_{v\in\mathbb{R}^\Lambda}\left(f_\beta(v) - I[v]\right) \leq c_5\log(T) + c_6.$$

*Proof.* (Theorem 0.8)
For $a \in \mathbb{R}^\Lambda$ let us define the potential

$$h(x) := \sum_{k\in L} a_k\cos(k\cdot x) + \sum_{k\in -L} a_k\sin(k\cdot x).$$

Applying Feynman-Kac yields

$$\begin{aligned}\langle 0|e^{-T(-\Delta-h)}|0\rangle &= (2\pi T)^{-\frac{3}{2}}\mathbb{E}_{0,0}e^{\int_0^T h(X_s)\ \mathrm{d}s} = (2\pi T)^{-\frac{3}{2}}\mathbb{E}_{0,0}e^{Ta\cdot V\left(\tau^{(T)}\right)}\\ &= (2\pi T)^{-\frac{3}{2}}\int e^{Ta\cdot v}\ \mathbb{P}_T(\mathrm{d}v).\end{aligned}$$

Let us define the modified potential $\tilde{h} := h - G(a)$. Since

$$G(a) = \sup_{v\in\mathbb{R}^\Lambda}(a\cdot v - I[v]) = \sup_{\tau\in\mathcal{M}(\mathbb{R}^3)}\Big(\int h\ \mathrm{d}\tau - I_{\text{Fisher}}[\tau]\Big) = \sup_{\|\phi\|_{L^2}=1}\Big(\int h|\phi|^2\mathrm{d}x - \mathcal{E}[\phi]\Big)$$

is the negative of the ground state energy corresponding to $-\Delta - h$, we obtain that the operator $-\Delta - \tilde{h}$ is positive. Hence, we have for $T \geq T_0$

$$\begin{aligned}&\langle 0|e^{-T(-\Delta-h)}|0\rangle = e^{TG(a)}\ \langle 0|e^{-T(-\Delta-\tilde{h})}|0\rangle \leq e^{TG(a)}\ \langle 0|e^{-T_0(-\Delta-\tilde{h})}|0\rangle\\ &= e^{TG(a)}(2\pi T_0)^{-\frac{3}{2}}\mathbb{E}_{0,0}e^{\int_0^{T_0}\tilde{h}(X_s)\ \mathrm{d}s} \leq e^{TG(a)}(2\pi T_0)^{-\frac{3}{2}}e^{2T_0\|h\|}.\end{aligned}$$

Using $\|h\|_\infty \leq c_1'|a|$, we obtain

$$\log\int e^{Ta\cdot v}\ \mathbb{P}_T(\mathrm{d}v) \leq TG(a) + \frac{3}{2}\log T + c_1T_0|a| - \frac{3}{2}\log\left(2\pi T_0\right).$$

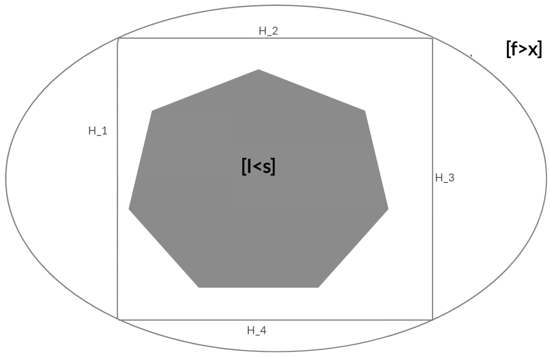
□

*Proof.* (Theorem 0.9)
Let us define the half spaces $H_a := \{v : a\cdot v \leq G(a) + s\}$. Then it is clear that,

$$[I \leq s] = \bigcap_{a\in\mathbb{R}^\Lambda} H_a.$$

For $s := \inf_{v\in A^\beta_{j,n}} I[v]$, let $a_1, .., a_p$ be as in Lemma 0.21. Then we know that $A^\beta_{j+1,n} = [f_\beta > x]$ with $x := \frac{j+1}{n}$ and $[I \leq s]$ are separated by the intersection of $p \leq Cn^{2|\Lambda|}$ half spaces $\bigcap_{i=1}^p H_i$ with $H_i := H_{a_i}$. Consequently,

$$\begin{aligned}\mathbb{P}_T(A^\beta_{j+1,n}) &\leq \sum_{i=1}^p \mathbb{P}_T(H_i^c) \leq \sum_{i=1}^p \int e^{T(a_i\cdot v - G(a_i) - s)}\ \mathbb{P}_T(\mathrm{d}v)\\ &\leq pe^{-Ts}\sup_i e^{-TG(a_i)}\int e^{Ta_i\cdot v}\ \mathbb{P}_T(\mathrm{d}v).\end{aligned}$$




Therefore we obtain, using Theorem 0.8

$$\log \mathbb{P}_T(A^\beta_{j+1,n}) \leq -Ts + \log(p) + \sup_i \left( -TG(a_i) + TG(a_i) + \frac{3}{2}\log T + c_1|a_i| \right)$$
$$\leq -Ts + 2|\Lambda|\log(n) + \frac{3}{2}\log T + c_1 C_M.$$

□

*Proof.* (Theorem 0.10)
Let us define $M_0 := \|f_1\|_\infty - \gamma_p(0)$ and $j_\beta(n)$ as the smallest index such that

$$\inf_{v\in A^\beta_{j_\beta(n),n}} I[v] \geq M_0.$$

By Lemma 0.22, we know that there exists a $M < \infty$, such that

$$\inf_{v\in A^\beta_{j,n}} I[v] \leq M$$

for all $j \leq j_\beta(n)+1$. Let us now define a partition of the measure space, given by $B_j := A^\beta_{j,n} \setminus A^\beta_{j+1,n}$ for $j \leq j_\beta(n)$ and $B_{j_\beta(n)+1} := A^\beta_{j_\beta(n)+1,n}$. Then we obtain

$$\int e^{Tf_\beta}\, \mathrm{d}\mathbb{P}_T \leq \sum_{j=0}^{j_\beta(n)+1} \int_{B_j} e^{Tf_\beta}\, \mathrm{d}\mathbb{P}_T \leq \sum_{j=0}^{j_\beta(n)} e^{T\frac{j+1}{n}} \mathbb{P}_T(A^\beta_{j,n}) + e^{T\|f\|_\infty}\mathbb{P}_T(A^\beta_{j_\beta(n)+1,n})$$
$$\leq n\|f_\beta\|_\infty \sup_{j\leq j_\beta(n)} e^{T\frac{j+1}{n}} \mathbb{P}_T(A^\beta_{j,n}) + e^{T\|f_\beta\|_\infty}\mathbb{P}_T(A^\beta_{j_\beta(n)+1,n}).$$

By applying Theorem 0.9, we can estimate both terms. We start by estimating the logarithm of the second one

$$\log\left(e^{T\|f_\beta\|_\infty}\mathbb{P}_T(A^\beta_{j_\beta(n)+1,n})\right) \leq T\|f_\beta\|_\infty - T\inf_{v\in A^\beta_{j_\beta(n),n}} I[v] + c_2\log(T) + c_3\log(n) + c_4$$
$$\leq T\left(\|f_\beta\|_\infty - M\right) + c_2\log(T) + c_3\log(n) + c_4.$$

We can compare $\gamma_p(0) \leq \gamma_p(\beta)$ and for $\beta \leq 1$ also $\|f_\beta\|_\infty \leq \|f_1\|_\infty$. Consequently,

$$\log\left(e^{T\|f_\beta\|_\infty}\mathbb{P}_T(A^\beta_{j_\beta(n)+1,n})\right) \leq T\gamma_p(\beta) + c_2\log(T) + c_3\log(n) + c_4.$$

For the other term, we have

$$\begin{aligned}
&\log\left(n\|f_\beta\|_\infty \sup_{j\leq j_\beta(n)} e^{T\frac{j+1}{n}}\mathbb{P}_T(A^\beta_{j,n})\right)\\
&\leq \log(n\|f_\beta\|_\infty) + T\sup_{j\leq j_\beta(n)}\left(\frac{j+1}{n} - \inf_{v\in A^\beta_{j-1,n}} I[v]\right) + c_2\log(T) + c_3\log(n) + c_4\\
&\leq T\sup_{j\leq j_\beta(n)}\sup_{v\in A^\beta_{j-1,n}}(f_\beta(v) - I[v]) + \frac{2T}{n} + c_2'\log(T) + c_3'\log(n) + c_4'\\
&\leq T\gamma_p(\beta) + \frac{2T}{n} + c_2'\log(T) + c_3'\log(n) + c_4'.
\end{aligned}$$

Using the inequality $\log(A+B) \leq \log(2) + \max\{\log(A), \log(B)\}$ leads to

$$\log\int e^{Tf_\beta}\,\mathrm{d}\mathbb{P}_T \leq T\gamma_p(\beta) + \frac{2T}{n} + c_2''\log(T) + c_3''\log(n) + c_4''.$$

With the choice $n := T$, the statement of the Theorem follows. □

## Auxiliary Tools

**Lemma 0.11.** *The functional $I : \mathbb{R}^\Lambda \to \mathbb{R} \cup \{\infty\}$ is convex.*

*Proof.* First, we verify that the Fisher information $I_{\text{Fisher}}$ is convex. To do so, let $\tau$ be an arbitrary measure with a smooth density function, i.e. a measure which can be written as $\mathrm{d}\tau = f\mathrm{d}x$. Then we can compute

$$I_{\text{Fisher}}[\tau] = \int |\nabla\sqrt{f}|^2 \,\mathrm{d}x = \int |\nabla\sqrt{f}|^2 \,\mathrm{d}x = \int \frac{|\nabla f|^2}{f} \,\mathrm{d}x.$$

Since the function $\phi(x, y) := \frac{|x|^2}{y}$ is a convex function defined on $\mathbb{R}^3 \times \mathbb{R}^+$, we obtain that $I_{\text{Fisher}}[\tau] = \int \phi(\nabla f, f) \,\mathrm{d}x$ is a convex functional.
To verify that $I$ is convex as well, let us consider a convex combination $v = t_1v_1 + t_2v_2 \in \mathbb{R}^\Lambda$. It is clear that every measure $\tau$ which can be written as $\tau = t_1\tau_1 + t_2\tau_2$ with $V(\tau_i) = v_i$, satisfies $V(\tau) = v$. Consequently,

$$\begin{aligned}
I[v] &= \inf_{\tau:V(\tau)=v} I_{\text{Fisher}}[\tau] \leq \inf_{\tau_1,\tau_2:V(\tau_i)=v_i} I_{\text{Fisher}}[t_1\tau_1 + t_2\tau_2] \\
&\leq \inf_{\tau_1,\tau_2:V(\tau_i)=v_i} \left(t_1 I_{\text{Fisher}}[\tau_1] + t_2 I_{\text{Fisher}}[\tau_2]\right) \\
&= t_1 \inf_{\tau_1:V(\tau_1)=v_1} I_{\text{Fisher}}[\tau_1] + t_2 \inf_{\tau_1:V(\tau_1)=v_1} I_{\text{Fisher}}[\tau_2] \\
&= t_1 I[v_1] + t_2 I[v_2].
\end{aligned}$$

□

**Definition 0.12.** *In the following, let us denote*

$$\mathcal{E}[\phi] := \int |\nabla\phi|^2.$$

**Lemma 0.13.** *(IMS localization formula)*
*For all $\delta, \epsilon$ with $0 < \delta < \epsilon$, we can find $C^\infty$ functions $0 \leq \chi_1, \chi_2 \leq 1$, such that $\chi_1(x) = 1$ for all $x \in [h > \|h\|_\infty - \delta]$, $\chi_1(x) = 0$ for all $x \in [h \leq \|h\|_\infty - \epsilon]$ and*

$$\chi_1^2 + \chi_2^2 = 1.$$

*Furthermore, we have for all $\phi \in L^2$*

$$\mathcal{E}[\phi] = \mathcal{E}[\chi_1\phi] + \mathcal{E}[\chi_2\phi] - \int \left(|\nabla\chi_1|^2 + |\nabla\chi_2|^2|\right) \phi^2.$$

*Proof.* The second part of the lemma follows simply by computations, therefore let us only prove the first part. Since $X_1 := [h \leq \|h\|_\infty - \epsilon]$ is a closed set which is disjoint to the closed set $X_2 := [h \geq \|h\|_\infty - \delta]$, we can find a $r > 0$ such that even $X_1 + B_{2r}(0)$ is disjoint to $X_2 + B_{2r}(0)$. Note that we have to use the periodicity of $h$ at this point, since we cannot assume that $X_1, X_2$ are compact sets. Let us define $f_i$ as the characteristic function of the set $(X_i + B_r(0))^c$ and $g_i := \phi * f_i$, where $\phi$ is a mollifier with support contained in $B_r(0)$. It is clear that $g_i|_{X_i} = 0$ as well as $g_i(x) = 1$ for $x \in (X_i + B_{2r}(0))^c$. Since $X_1 + B_{2r}(0)$ is disjoint to $X_2 + B_{2r}(0)$, we know that either $x \in (X_1 + B_{2r}(0))^2$ or $x \in (X_2 + B_{2r}(0))^2$. Therefore,

$$g_1^2 + g_2^2 \geq 1.$$

Consequently, the functions $\chi_i := \left(g_1^2 + g_2^2\right)^{-\frac{1}{2}} g_i$ are $C^\infty$ and satisfy $0 \leq \chi_i \leq 1$ as well as

$$\chi_1^2 + \chi_2^2 = 1.$$

It is immediately clear that $\chi_1|_{X_1} = 0$. Finally, an element $x$ in $[h > \|h\|_\infty - \delta]$ is especially an element in $X_2$. Therefore, we know that $g_2(x) = 0$ and consequently

$$\chi_1(x) = \left(g_1(x)^2\right)^{-\frac{1}{2}} g_1(x) = 1.$$

□

**Lemma 0.14.** *(IMS localization formula-infinite version)*
*Recall the definition* $\Lambda := \{k \in \left(n^{-1}\mathbb{Z}\right)^3 : 0 < |k| \leq m\}$. *We define* $G := [0, 2\pi n]^3$ *as the smallest unit of periodicity, concerning the functions* $e^{ik\cdot x}$ *with* $k \in \Lambda$. *Furthermore, we denote* $G_\ell := G + (2\pi n)\ell$ *for* $\ell \in \mathbb{Z}^3$. *Then, there exist functions* $(\chi_\ell)_{\ell\in\mathbb{Z}^3}$ *such that* $0 \leq \chi_\ell \leq 1$, $\chi_\ell|_{G_\ell} = 1$, $\chi_\ell|_{G_{\ell'}} = 0$ *for all* $|\ell - \ell'| > 1$ *and*

$$\sum_\ell \chi_\ell^2 = 1.$$

*Furthermore, there exists a constant* $C$, *such that we have for all* $\phi \in L^2$

$$\mathcal{E}[\phi] \geq \sum_\ell \mathcal{E}[\chi_\ell \phi] - C \int \phi^2.$$

*Proof.* Let $0 \leq g \leq 1$ be a $C^\infty$ function, which satisfies $g|_G = 1$ and $g|_{G_\ell} = 0$ for all $|\ell| > 1$. We define $g_\ell(x) := g_\ell(x - \ell)$. It is clear, that for all $x \in \mathbb{R}^3$, there exist at most six indices $\ell_1, .., \ell_6$ such that $g_\ell(x) \neq 0$. Therefore, $0 \leq \sum_\ell g_\ell^2 \leq 6$ and this sum is a $C^\infty$ function as well, since for all compact subsets only a finite number of indices yield a non zero contribution. It is easy to check, that the functions

$$\chi_\ell := \left(\sum_{\ell'} g_{\ell'}^2\right)^{-\frac{1}{2}} g_\ell$$

have the right properties. For the last inequality of the lemma, note that for all $x \in \mathbb{R}^3$ only six indices $\ell_1, .., \ell_6$ satisfy $|\nabla \chi_\ell| \neq 0$ and since the different functions are translations of each other, we have the uniform bound $|\nabla \chi_\ell| \leq C'$ for some $C'$. If we define $C := 6C'$ we obtain

$$\mathcal{E}[\phi] \geq \sum_\ell \mathcal{E}[\chi_\ell \phi] - \int \left( \sum_\ell |\nabla \chi_\ell|^2 \right) \phi^2 \geq \sum_\ell \mathcal{E}[\chi_\ell \phi] - C \int \phi^2.$$

□

**Lemma 0.15.** *For a non zero $a \in \mathbb{R}^\Lambda$, let us define the function $h : \mathbb{R}^3 \to \mathbb{R}$*

$$h(x) := \sum_{k \in L} a_k \cos(k \cdot x) + \sum_{k \in -L} a_k \sin(k \cdot x).$$

*Then, the set $S := [h = \|h\|_\infty]$ has zero volume, i.e.*

$$\int_S 1 \, \mathrm{d}x = 0.$$

*Proof.* Let us define

$$\tilde{h} := h - \|h\|_\infty = -\|h\|_\infty \cdot 1 + \sum_{k \in L} a_k \cos(k \cdot .) + \sum_{k \in -L} a_k \sin(k \cdot .).$$

Note that $0 \notin \Lambda$, therefore the functions $1, \sin(k \cdot .), \cos(k \cdot .)$ are linearly independent. Since $a \neq 0$, this implies that $\tilde{h}$ is not constant equals 0, and therefore we know from [6] that $[\tilde{h} = 0] = [h = \|h\|_\infty]$ is a set with zero volume.

□

**Lemma 0.16.** *Let $S_\epsilon := [h \geq \|h\|_\infty - \epsilon]$. Then $S_\epsilon$ gets thinner for $\epsilon \to 0$, in the sense that $c_\epsilon$ goes to infinity for $\epsilon \to 0$, where $c_\epsilon$ is defined as*

$$c_\epsilon := \sup_\phi \mathcal{E}[\phi]$$

*and the supremum is take over all $\|\phi\| = 1$ which have their support in $S_\epsilon$.*

*Proof.* We know that from Lemma 0.15, that $S := [h = \|h\|_\infty]$ has zero volume. If we define $G_\ell^+ := \bigcup_{|\ell' - \ell| \leq 1} G_{\ell'}$ as the second unit of periodicity (compare with Lemma 0.14), then Lemma 0.15 tells us

$$\int_{G_\ell^+ \cap S_\epsilon} 1 \, \mathrm{d}x \underset{\epsilon \to 0}{\longrightarrow} \int_{G_\ell^+ \cap S} 1 \, \mathrm{d}x = 0.$$

Therefore, the constant $c'_\epsilon$ defined as

$$c'_\epsilon := \sup_\phi \mathcal{E}[\phi]$$

where the supremum is take over all $\|\phi\| = 1$ which have their support in $G_\ell^+ \cap S_\epsilon$, goes to infinity for small $\epsilon$. Note that due to the periodicity, $c'_\epsilon$ is really a $\ell$ independent quantity. Note that for all normed $\phi$ with support in $S_\epsilon$, the corresponding functions $\chi_\ell \phi$ from Lemma 0.14 have their support contained in $G_\ell^+ \cap S_\epsilon$. We conclude

$$\mathcal{E}[\phi] \geq \sum_\ell \mathcal{E}[\chi_\ell \phi] - C \geq c'_\epsilon \sum_\ell \int (\chi_\ell \phi)^2 - C = c'_\epsilon - C \underset{\epsilon \to 0}{\longrightarrow} \infty.$$

□

**Lemma 0.17.** *The set* $\Omega := [I < \infty]$ *is open and for all sequences* $v_n \underset{n\to\infty}{\longrightarrow} v \in \partial\Omega$ *with a limit in the boundary of* $\Omega$*, we have*

$$\lim_{n\to\infty} I[v_n] = \infty.$$

*Proof.* Let $v \in \partial\Omega \subset \mathbb{R}^\Lambda$. Since $I$ is a convex functional, we know that the $\Omega$ is a convex set. Therefore, there exists an $a \in \mathbb{R}^\Lambda \setminus \{0\}$ such that

$$a \cdot v = \sup_{v' \in \Omega} a \cdot v'. \tag{1}$$

With this $a$ at hand, let us define in the spirit of Lemma 0.16 the function $h : \mathbb{R}^3 \to \mathbb{R}$

$$h(x) := \sum_{k \in L} a_k \cos(k \cdot x) + \sum_{k \in -L} a_k \sin(k \cdot x).$$

Let $\tau$ be any measure, such that $V(\tau) = v$. Then we can rewrite Equation (1) as

$$\int h \, \mathrm{d}\tau = \sup_{\tau' : I_{\mathrm{Fisher}}[\tau'] < \infty} \int h \, \mathrm{d}\tau' = \|h\|_\infty.$$

By the definition of $c_\epsilon$ in Lemma 0.16, it is clear that $I_{\mathrm{Fisher}}[\tau] \geq c_\epsilon$, and therefore we have by the statement of Lemma 0.16

$$I_{\mathrm{Fisher}}[\tau] \geq c_\epsilon \underset{\epsilon \to 0}{\longrightarrow} \infty.$$

Consequently, $v \notin \Omega$ since there is no realization $V(\tau) = v$ with finite Fisher information. This tells us, that $\Omega$ is disjoint to its boundary, and hence it is an open set.
Let us now consider an arbitrary sequence $v_n \underset{n\to\infty}{\longrightarrow} v$ converging to $v$. It is clear that $a \cdot v_n \underset{n\to\infty}{\longrightarrow} a \cdot v = \|h\|_\infty$. Therefore, for all $t > 0$ there exists a $n_0$ such that for all $n \geq n_0$, we have $a \cdot v_n \geq \|h\|_\infty - t$. This means for any realization $V(\tau) = v_n$ of $v_n$, we

have $\int h \, \mathrm{d}\tau \geq \|h\|_\infty - t$. For arbitrary $\epsilon > 0$ and $q > 0$ we can find $t$ small enough, such that this leads to

$$\tau \left( [h \leq \|h\|_\infty - \frac{\epsilon}{2}] \right) < q.$$

In the following let $\chi_1, \chi_2$ be as in Lemma 0.13 for $\delta := \frac{\epsilon}{2}$ and let us define $C_\epsilon := \|\nabla \chi_1\|_\infty + \|\nabla \chi_2\|_\infty$. If $I_{\text{Fisher}}[\tau]$ is finite, then we can write $\mathrm{d}\tau = \phi^2 \mathrm{d}x$ and we define $\phi_1 := \chi_1 \phi$ as well as $\phi_2 := \chi_2 \phi$. Note that the support of $\phi_1$ is contained in the set $S_\epsilon$ from Lemma 0.16, and therefore $\mathcal{E}[\phi_1] \geq c_\epsilon \|\phi_1\|^2$. We also know that $\|\phi_1\|^2 + \|\phi_2\|^2 = 1$ as well as $\|\phi_2\|^2 \leq q$ since we know that $\tau \left( [h \leq \|h\|_\infty - \frac{\epsilon}{2}] \right) < q$ as well as that the support of $\chi_2$ is contained in $[h \leq \|h\|_\infty - \frac{\epsilon}{2}]$. If we combine these results and use that $\nabla \chi_i$ has a support contained in $[h \leq \|h\|_\infty - \frac{\epsilon}{2}]$, we obtain

$$I_{\text{Fisher}}[\tau] = \mathcal{E}[\phi] \geq \mathcal{E}[\phi_1] + \mathcal{E}[\phi_2] - C\tau \left( \left[ h \leq \|h\|_\infty - \frac{\epsilon}{2} \right] \right) \geq c_\epsilon (1 - q) - C_\epsilon q.$$

If we chose $q$ small enough, we obtain for all $n \geq n(\epsilon)$ big enough and all $\tau$ which satisfy $V(\tau) = v_n$, that

$$I_{\text{Fisher}}[\tau] \geq \frac{1}{2} c_\epsilon - 1.$$

Since the right hand side goes to $\infty$ for $\epsilon \to 0$, we conclude

$$\lim_{n \to \infty} I[v_n] = \infty.$$

$\square$

**Corolary 0.18.** *The functional $I$ is lower semi-continuous. Therefore, the convex conjugate acts as an involution $I^{**} = I$.*

**Lemma 0.19.** *The sets $[f_\beta \leq x]$ and $[f_\beta > x + \delta]$ can be separated by at most $C\delta^{-2|\Lambda|}$ half spaces, where $C$ does not depend on $x$ as long as $x + \delta \leq x_0$ for some $x_0$. This means, there exist half spaces $Q_1, .., Q_m$ with $m \leq C\delta^{-2|\Lambda|}$ such that*

$$[f_\beta \leq x] \subset \bigcap_{i=1}^{m} Q_i \subset [f_\beta \leq x + \delta].$$

*Proof.* Let us define the linear map $L : \mathbb{R}^\Lambda \to \mathbb{R}^\Lambda$ by $L(v) := \left( \frac{\beta}{2} \coth \left( \frac{\beta}{2} \right) |k|^2 v_k \right)_k$ as well as the sets $X := L \left( [f_\beta \leq x] \right)$ and $Y := L \left( [f_\beta \leq x + \delta] \right)$. By the definition of $f_\beta$ it is clear that $X = B_{\sqrt{x}}(0)$ and $Y = B_{\sqrt{x+\delta}}(0)$ are balls with radius $r_1 := \sqrt{x}$, respectively $r_2 := \sqrt{x + \delta}$. Since $x + \delta \leq x_0$, we have $r_2 - r_1 \geq (2\sqrt{x_0})^{-1} \delta$. A ball of radius $r_1$ and one of radius $r_2$ in a $d := |\Lambda|$ dimensional space can be separated by $D(r_2 - r_1)^{-2d}$ half spaces $V_1, .., V_m$ with $m \leq D(r_2 - r_1)^{-2d} \leq \tilde{D}\delta^{-2|\Lambda|}$. Therefore, the half spaces $Q_i := L^{-1}(V_i)$ satisfy the claim of the Lemma.

□

**Lemma 0.20.** *Let $P := \bigcap_{i=1}^{M} Q_i \subset \mathbb{R}^d$ be a non empty intersection of finitely many half spaces $Q_i$ and $Q$ another half space with $P \subset Q$. Then, there exist $i_1, .., i_{d+1} \in J$ which already satisfy*

$$\bigcap_{k=1}^{d+1} Q_{i_k} \subset Q.$$

*Proof.* Let us write the half spaces as $Q = \{v : a \cdot v \leq s\}$ and $Q_i = \{v : a_i \cdot v \leq s_i\}$. With the definition $s^* := \sup_{v \in P} (a \cdot v)$ and $Q^* := \{v : a \cdot v \leq s^*\}$ we obtain that there exists a $v_0 \in \partial P$ with $a \cdot v_0 = s^*$ as well as the inclusion $P \subset Q^* \subset Q$. Furthermore, we denote with $N_{v_0} \subset \mathbb{R}^d$ the normal cone of $P$ at $v_0$, i.e. the set of all directions $b$ such that $b \cdot (v_0 - v) \geq 0$ for all $v \in P$. It is clear that $a \in N_{v_0}$.
Note that $P$ is a convex polytope, defined by the collection of linear inequalities $a_i \cdot v \leq s_i$. This implies, that the normal cone is the convex cone generated by the elements $a_i$, which satisfy $a_i \cdot v_0 = s_i$, i.e.

$$N_{v_0} = \mathrm{c}\left(\{a_i : i \in I(v_0)\}\right)$$

where $I(v_0)$ is the set of all $i$ which satisfy $a_i \cdot v_0 = s_i$. Let $\Gamma$ be a possible infinite triangulation of this cone, i.e. $\Gamma = \{\gamma_j : j \in J\}$ is a collection of $d+1$ simplices $\gamma_j$ such that the union of all simplices is $N_{v_0}$ and the extreme points of the $\gamma_j$ are subsets of $\{\lambda a_i : i \in I(v_0), \lambda \geq 0\}$. Furthermore, the $\gamma_j$ also satisfy a "disjointness" property, which we do not need in the following. Since $a \in N_{v_0}$, there exists a simplex $\gamma \in \Gamma$ with $a \in \gamma$. From the properties of a triangulation follows, that we can write $\gamma = \mathrm{conv}(\lambda_1 a_{i_1}, .., \lambda_{d+1} a_{i_{d+1}})$ as the convex set, generated by the points $\lambda_1 a_{i_1}, .., \lambda_{d+1} a_{i_{d+1}}$. Therefore, there exist $0 \leq t_1, .., t_n$ such that $a = t_1 a_{i_1} + .. + t_{d+1} a_{i_{d+1}}$. Consequently,

$$Q_{i_1} \cap .. \cap Q_{i_{d+1}} \subset Q^* \subset Q.$$

□

**Lemma 0.21.** *Let $s := \inf_{v \in A^{\beta}_{j,n}} I[v] \leq M$. Then we can find $a_1, .., a_p$ with $p \leq Cn^{2|\Lambda|}$ as well as $|a_i| \leq C_M$, such that the half spaces*

$$H_i := \{v : a_i \cdot v \leq G(a_i) + s\}$$

*separate the convex stets $[I \leq s]$ and $A^{\beta}_{j+1,n}$.*

*Proof.* Let us define $x := \frac{j+\frac{1}{2}}{n}$, $y := \frac{j+1}{n}$ and $\delta := y - x = \frac{1}{2n}$. Then it is clear that $[f_\beta > y] = A^\beta_{j+1,n}$ and

$$[I \leq s] \subset [f_\beta < x].$$

Furthermore, we denote $H_a := \{v : a \cdot v \leq G(a) + s\}$. Since $[I \leq s] = \bigcap_{a \in \mathbb{R}^\Lambda} H_a$ is a compact set which is included in the open and bounded set $[f_\beta < x]$ we obtain by a compactness argument, that there exist finitely many $H_{a_1}, .., H_{a_M}$ such that already $P := \bigcap_{i=1}^M H_{a_i} \subset [f_\beta < x]$. By Lemma 0.19, we know that there exist half spaces $Q_1, .., Q_m$ with $m \leq Cn^{2|\Lambda|}$ and

$$[f_\beta < x] \subset \bigcap_{l=1}^m Q_l \subset A^\beta_{j+1,n}.$$

Especially, $P \subset Q_l$ for all $l = 1, .., m$. Therefore, we can apply Lemma 0.20 for $P = \bigcap_{i=1}^M H_{a_i}$ and each subspace $Q_l$, which yields us the existence of indices $a_{i_{l,k}}$ with $l = 1, .., m$ and $k = 1, .., |\Lambda| + 1$, such that $\bigcap_{k=1}^{|\Lambda|+1} H_{a_{i_{l,k}}} \subset Q_l$. Consequently, the collection of half spaces $H_{a_{i_{l,k}}}$ consists of at most $C(|\Lambda| + 1)n^{2|\Lambda|}$ elements and indeed separates $[I \leq s]$ from $A^\beta_{j+1,n}$, since

$$[I \leq s] \subset P \subset \bigcap_{l,k} H_{a_{i_{l,k}}} \subset \bigcap_{l=1}^m Q_l \subset A^\beta_{j+1,n}.$$

Finally, we have to verify that all $a$ involved in this collection are bounded in a suitable way. First we define $K := [I \leq M]$, where $M$ is given by the assumption of the Lemma. Since this is a compact set contained in the open set $\Omega := [I < \infty]$, there exists an $\epsilon > 0$, such that the closure of $K^+ := K + B_\epsilon(0)$ is still contained in $\Omega$. As a consequence, $\sup_{a \in K^+} I[v] < \infty$. Without loss of generality, we can assume that all $H_a$ involved in the separation of $[I \leq s]$ from $A^\beta_{j+1,n}$ are approximately tangent spaces, in the sense that there exists a $v \in \mathbb{R}^\Lambda$ such that $a \cdot v = G(a) + s$ and $\operatorname{dist}(v, [I \leq s]) < \frac{\epsilon}{2}$. Therefore,

$$a \cdot v \geq G(a) = \sup_{w \in \mathbb{R}^\Lambda} (a \cdot w - I[w]) \geq a \cdot \left( v + \frac{\epsilon}{2|a|} a \right) - I \left[ v + \frac{\epsilon}{2|a|} a \right].$$

Rearranging the inequality above and using the assumption $s \leq M$ yields

$$|a| \leq \frac{2}{\epsilon} I \left[ v + \frac{\epsilon}{2|a|} a \right] \leq \frac{2}{\epsilon} \sup_{a \in K^+} I[v] =: C_M < \infty.$$

□

**Lemma 0.22.** *For a given $M_0 < \infty$ let $j_\beta(n)$ be the smallest index such that*

$$\inf_{v \in A^\beta_{j_\beta(n),n}} I[v] \geq M_0.$$

*Then, one can can find a $M < \infty$ and $\beta_0, n_*$, such that for all $\beta \leq \beta_0$, $n \geq n*$ and $j \leq j_\beta(n) + 1$*

$$\inf_{v \in A^\beta_{j,n}} I[v] \leq M.$$

*Proof.* Let us define $y' := \max f_0\left([I \leq M_0]\right)$. If we take $\epsilon$ small enough and define $y := y' + \epsilon$, we know that $M := \inf_{f_0(v) > y} I[v] < \infty$. Let us take $\beta_0$ small enough, such that $\|f_\beta - f_0\| < \frac{\epsilon}{3}$ for all $\beta \leq \beta_0$ and $n_*$ big enough such that $\frac{1}{n} < \frac{\epsilon}{3}$ for all $n \geq n_*$. If we can show that $\frac{j_\beta(n)+1}{n} \leq y$, we are done since this implies for all $j \leq j_\beta(n) + 1$

$$\inf_{f_\beta(v) > \frac{j}{n}} I[v] \leq \inf_{f_0(v) > y} I[v] = M.$$

We are going to verify $\frac{j_\beta(n)+1}{n} \leq y$ by contradiction. Therefore, let us assume $\frac{j_\beta(n)+1}{n} > y$. This implies $\frac{j_\beta(n)-1}{n} > y' + \frac{\epsilon}{3}$. Since $\max f_\beta\left([I \leq M_0]\right) \leq y' + \frac{\epsilon}{3}$, it is clear that for all $v \in A^\beta_{j_\beta(n)-1,n}$ we have $f_\beta(v) > \max f_\beta\left([I \leq M_0]\right)$. Therefore, we have already for the index $j_\beta(n) - 1$ the estimate

$$\inf_{v \in A^\beta_{j_\beta(n)-1,n}} I[v] \geq M_0.$$

A contradiction to the definition of the index $j_\beta(n)$. □

## Upper Bound on the Ground State Energy

In this section, we want to prove the upper bound on the ground state energy given by Theorem 0.2. We will achieve this by computing the energy of suitable test functions. Before we can turn to the proof of the theorem, we need to prove the following Lemma:

**Lemma 0.23.** *Let* $g : \mathbb{R}^3 \to \mathbb{R}$ *be a function with* $\|g\| = 1$. *Furthermore, let* $L = k2\pi n$ *be a multiple of the periodicity* $2\pi n$. *Then we can find a modified version* $\tilde{g}$ *with* $\operatorname{supp}(\tilde{g}) \subset [-2L, 2L]^3$, $\|\tilde{g}\| = 1$, $\int |\nabla \tilde{g}|^2 \, \mathrm{d}x \leq \int |\nabla g|^2 \, \mathrm{d}x + CL^{-2}$ *and*

$$\int\int n^{-3} \sum_{k\in\Lambda} |k|^{-2} e^{ik(x-y)} |\tilde{g}(x)|^2 |\tilde{g}(y)|^2 \, \mathrm{d}x\mathrm{d}y = \int\int n^{-3} \sum_{k\in\Lambda} |k|^{-2} e^{ik(x-y)} |g(x)|^2 |g(y)|^2 \, \mathrm{d}x\mathrm{d}y.$$

*Proof.* Similar to the IMS formula in 0.14, we construct a smooth quadratic partition of the whole space. First of all, let $h$ be a smooth function which is one on $[-1,1]^3$ and has a support contained in $[-2,2]^3$. For $\ell \in \mathbb{Z}^3$ we define the translated version $h_\ell(x) := h(x-\ell)$ and $H_\ell(x) := \left(\sum_{\ell'} h_{\ell'}(x)\right)^{-\frac{1}{2}} h_\ell(x)$. It is clear that this these functions form a quadratic partition of the whole space and $\sum_\ell |\nabla H_\ell|^2 \leq C$ for some constant $C$. Now we define the final partition as $\chi_\ell(x) := H_\ell(L^{-1}x)$. It is clear that this is again a smooth quadratic partition, now with the support property $\operatorname{supp}(\chi_\ell) \subset [-2L, 2L]^3$. The gradient can be estimated by

$$\sum_\ell |\nabla \chi_\ell|^2 \leq CL^{-2}.$$

Let us define the measures $\tau, \tau_\ell, \tau', \tau'_\ell$ by $\mathrm{d}\tau = g^2 \mathrm{d}x$, $\mathrm{d}\tau_\ell = (\chi_\ell g)^2 \, \mathrm{d}x$ as well as the translated versions $\tau'_\ell := \tau_\ell(. + \ell L)$ and $\tau' := \sum_\ell \tau'_\ell$. Note that it is clear that $\tau = \sum_\ell \tau_\ell$ and that $\tau'$ is like $\tau$ a probability measure. The advantage of the rearranged measure $\tau'$ is that it has a support contained in $[-2L, 2L]^3$. Let us verify the bounds on the Fisher information, which is just the Dirichlet form on the level of measures. From the properties of a quadratic partition of unity, it is clear that

$$\sum_\ell I_{\text{Fisher}}[\tau'_\ell] = I_{\text{Fisher}}[\tau] + \int \sum_\ell |\nabla \chi_\ell|^2 \, \mathrm{d}\tau \leq I_{\text{Fisher}}[\tau] + CL^{-2},$$

where we used that the Fisher information is translation invariant $I_{\text{Fisher}}[\tau'_\ell] = I_{\text{Fisher}}[\tau_\ell]$. Furthermore, we know that $I_{\text{Fisher}}$ is a convex functional and it is obvious that it scales like $I_{\text{Fisher}}[\lambda\mu] = \lambda I_{\text{Fisher}}[\mu]$ for a positive number $\lambda$, if we allow it to be defined on all measures and not only probability measures. Therefore, $I_{\text{Fisher}}$ is even sub-linear, which implies

$$I_{\text{Fisher}}[\tau'] \leq \sum_\ell I_{\text{Fisher}}[\tau'_\ell] \leq I_{\text{Fisher}}[\tau] + CL^{-2}.$$

Finally, note that the function $f(x,y) := n^{-3}\sum_{k\in\Lambda}|k|^{-2}e^{ik(x-y)}$ is periodic, in the sense that $f(x-\ell L, y-\ell' L) = f(x,y)$. Therefore,

$$\int f(x,y)\ \tau'\otimes\tau'(\mathrm{d}x,\mathrm{d}y) = \sum_{\ell,\ell'}\int f(x,y)\ \tau'_\ell\otimes\tau'_{\ell'}(\mathrm{d}x,\mathrm{d}y) = \sum_{\ell,\ell'}\int f(x-\ell,y-\ell')\ \tau_\ell\otimes\tau_{\ell'}(\mathrm{d}x$$

$$= \sum_{\ell,\ell'}\int f(x,y)\ \tau_\ell\otimes\tau_{\ell'}(\mathrm{d}x,\mathrm{d}y) = \int f(x,y)\ \tau\otimes\tau(\mathrm{d}x,\mathrm{d}y).$$

From these properties on $\tau'$, it is clear that the square root of the density function

$$\tilde{g} := \sqrt{\frac{\mathrm{d}\tau}{\mathrm{d}x}} = \chi_0(x)\sqrt{\sum_\ell g^2(x+\ell L)}$$

satisfies the desired properties.

□

*Proof.* (Proof of Theorem 0.2)
Let us consider the test function $\Psi := g\otimes\phi$, which should be the product of a scalar function $\|g\|_{L^2} = 1$ and the coherent state in the many particle space given by

$$\phi := g(x)\ e^{-\frac{\alpha}{\sqrt{2}}\sum_{k\in\Lambda}\left(c_k a_k^\dagger - \bar{c}_k a_k\right)}\cdot|0\rangle\,,$$

where we define the coefficients as $c_k := -\alpha n^{-\frac{3}{2}}|k|^{-1}\int|g(x)|^2e^{i\alpha kx}\ \mathrm{d}x$. Note that $\Psi$ has norm one. Furthermore, we compute

$$\langle\Psi|\ a_k^\dagger a_k\cdot\Psi\rangle = e^{-|c_k|^2}\sum_{n=0}^{n} n\frac{|c_k|^2}{n!}\int|g(x)|^2\ \mathrm{d}x = |c_k|^2.$$

Let us write $e^{i\alpha k\hat{x}}$ for the multiplication operator by $e^{i\alpha kx}$ and compute

$$\langle\Psi|\ \alpha n^{-\frac{3}{2}}|k|^{-1}e^{i\alpha k\hat{x}}a_k^\dagger\cdot\Psi\rangle = e^{-|c_k|^2}\sum_{n,m}\delta_{m,n+1}\sqrt{n+1}\frac{\bar{c}_k^m c_k^n}{\sqrt{n!m!}}\ \alpha n^{-\frac{3}{2}}|k|^{-1}\int e^{i\alpha kx}|g(x)|^2\ \mathrm{d}x$$

$$= e^{-|c_k|^2}\sum_n\frac{\left(|c_k|^2\right)^n\bar{c}_k}{n!}(-c_k) = -|c_k|^2.$$

Since this is a real number, it is clear that $\langle\Psi|\ \alpha n^{-\frac{3}{2}}|k|^{-1}e^{-i\alpha k\hat{x}}a_k\cdot\Psi\rangle = -|c_k|^2$ as well. If we combine the results, we obtain the following upper bound on the ground state energy

$$E_0(\alpha)\le\langle\Psi|H^\alpha|\Psi\rangle = \int|\nabla g|^2\ \mathrm{d}x + \int\frac{\epsilon_\alpha}{2}|x|^2|g(x)|^2\ \mathrm{d}x - \sum_k|c_k|^2$$

$$= \int|\nabla g|^2\ \mathrm{d}x + \int\frac{\epsilon_\alpha}{2}|x|^2|g(x)|^2\ \mathrm{d}x - \sum_{k\in\Lambda}\alpha^2n^{-3}|k|^{-2}\int\int e^{ik\alpha(x-y)}|g(x)|^2|g(y)|^2\ \mathrm{d}x\mathrm{d}y.$$

Since the inequality above holds for all $g$ with $\|g\| = 1$, we obtain by rescaling

$$E_0(\alpha) \leq \alpha^2 \left( \int |\nabla g|^2 \, \mathrm{d}x + \int \frac{\epsilon_\alpha}{2\alpha^4} |x|^2 |g(x)|^2 \, \mathrm{d}x - \int \int n^{-3} \sum_{k \in \Lambda} |k|^{-2} e^{ik(x-y)} |g(x)|^2 |g(y)|^2 \, \mathrm{d}x\mathrm{d}y \right).$$

Now we need to argue why we can omit the $\epsilon_\alpha$. To do so, let $g$ be a function which minimizes up to an error $\delta$ the functional $\int |\nabla g|^2 \, \mathrm{d}x - \int \int n^{-3} \sum_{k \in \Lambda} |k|^{-2} e^{ik(x-y)} |g(x)|^2 |g(y)|^2 \, \mathrm{d}x\mathrm{d}y$, i.e.

$$\int |\nabla g|^2 \, \mathrm{d}x - \int \int n^{-3} \sum_{k \in \Lambda} |k|^{-2} e^{ik(x-y)} |g(x)|^2 |g(y)|^2 \, \mathrm{d}x\mathrm{d}y < -\gamma_p + \delta.$$

From Lemma 0.23 we know that there exists for all $L$ a function $\tilde{g}$ with $\|\tilde{g}\| = 1$, $\mathrm{supp}(\tilde{g}) \subset [-2L, 2L]^3$ as well as $\int |\nabla \tilde{g}|^2 \, \mathrm{d}x \leq \int |\nabla g|^2 \, \mathrm{d}x + CL^{-2}$, while satisfying

$$\int \int n^{-3} \sum_{k \in \Lambda} |k|^{-2} e^{ik(x-y)} |\tilde{g}(x)|^2 |\tilde{g}(y)|^2 \, \mathrm{d}x\mathrm{d}y = \int \int n^{-3} \sum_{k \in \Lambda} |k|^{-2} e^{ik(x-y)} |g(x)|^2 |g(y)|^2 \, \mathrm{d}x\mathrm{d}y.$$

If we use $\tilde{g}$ as a test function, we obtain for all $\delta > 0$ and $L$

$$E_0(\alpha) \leq \alpha^2 \left( -\gamma_p + \delta + CL^{-2} + \frac{6\epsilon_\alpha}{\alpha^4} L^2 \right).$$

Sending $\delta \to 0$ and taking the optimal $L := \alpha \left(6\epsilon_\alpha\right)^{-\frac{1}{4}}$ yields

$$E_0(\alpha) + \alpha^2 \gamma_p \leq D\sqrt{\epsilon_\alpha}.$$

□

## References


[1] Janusz Adamowski, Bernd Gerlach, and Hajo Leschke. Strong-coupling limit of polaron energy, revisited. *Physics Letters A*, 79(2):249 – 251, 1980.

[2] David Brydges, Remco van der Hofstad, and Wolfgang König. Joint density for the local times of continuous-time markov chains. *Ann. Probab.*, 35(4):1307–1332, 07 2007.

[3] M. D. Donsker and S. R.S. Varadhan. Asymptotics for the polaron. July 1983.

[4] Jürgen Gärtner. On large deviations from the invariant measure. *Theory of Probability and Its Applications*.

[5] Olav Kallenberg. *Foundations of modern probability*. Probability and its Applications (New York). Springer-Verlag, New York, second edition, 2002.

[6] Boris Mityagin. The zero set of a real analytic function. *Mathematical Notes*, 107, 12 2015.