Review

The Quantum Harmonic Otto Cycle

Ronnie Kosloff 1,* † and Yair Rezek 2,3,†

1 Institute of Chemistry, The Hebrew University, Jerusalem 91904, Israel
2 Computer Science Department, Technion, Haifa 32000, Israel; yair.rezek@mail.huji.ac.il
3 Tel Hai College, Kiryat Shmona 1220800, Israel
* Correspondence: ronnie@fh.huji.ac.il; Tel.: +972-2-6585-485
† These authors contributed equally to this work.

Abstract: The quantum Otto cycle serves as a bridge between the macroscopic world of heat engines and the quantum regime of thermal devices composed from a single element. We compile recent studies of the quantum Otto cycle with a harmonic oscillator as a working medium. This model has the advantage that it is analytically trackable. In addition, an experimental realization has been achieved, employing a single ion in a harmonic trap. The review is embedded in the field of quantum thermodynamics and quantum open systems. The basic principles of the theory are explained by a specific example illuminating the basic definitions of work and heat. The relation between quantum observables and the state of the system is emphasized. The dynamical description of the cycle is based on a completely positive map formulated as a propagator for each stroke of the engine. Explicit solutions for these propagators are described on a vector space of quantum thermodynamical observables. These solutions which employ different assumptions and techniques are compared. The tradeoff between power and efficiency is the focal point of finite-time-thermodynamics. The dynamical model enables the study of finite time cycles limiting time on the adiabatic and the thermalization times. Explicit finite time solutions are found which are frictionless (meaning that no coherence is generated), and are also known as shortcuts to adiabaticity. The transition from frictionless to sudden adiabats is characterized by a non-hermitian degeneracy in the propagator. In addition, the influence of noise on the control is illustrated. These results are used to close the cycles either as engines or as refrigerators. The properties of the limit cycle are described. Methods to optimize the power by controlling the thermalization time are also introduced. At high temperatures, the Novikov–Curzon–Ahlborn efficiency at maximum power is obtained. The sudden limit of the engine which allows finite power at zero cycle time is shown. The refrigerator cycle is described within the frictionless limit, with emphasis on the cooling rate when the cold bath temperature approaches zero.

Keywords: quantum thermodynamics; quantum transport; quantum; Gibbs state

PACS: 03.65.Yz; 05.70.Ln; 07.20.Pe

1. Introduction

Quantum thermodynamics is devoted to the link between thermodynamical processes and their quantum origin. Typically, thermodynamics is applied to large macroscopic entities. Therefore, to what extent is it possible to miniaturize. Can thermodynamics be applicable to the level of a single quantum device? We will address this issue in the tradition of thermodynamics, by learning from an example: Analysis of the performance of a heat engine [1]. To this end, we review recent progress in the study of the quantum harmonic oscillator as a working medium of a thermal device. The engine composed of
a single harmonic oscillator connected to a hot and cold bath is an ideal analytically solvable model for a quantum thermal device. It has therefore been studied extensively and inspired experimental realisation. Recently, a single ion heat engine with an effective harmonic trap frequency has been experimentally realised [2]. This device could roughly be classified as a reciprocating Otto engine operating by periodically modulating the trap frequency.

Real heat engines operate far from reversible conditions. Their performance resides between the point of maximum efficiency and maximum power. This has been the subject of finite time thermodynamics [3,4]. The topic has been devoted to the irreversible cost of operating at finite power. Quantum engines add a twist to the subject, as they naturally incorporate dynamics into thermodynamics [5,6].

Quantum heat engines can be classified either as continuous or reciprocating. The prototype of a continuous engine is the three-level amplifier pioneered by Scovil and Schulz-DuBois [7]. This device is simultaneously coupled to three input currents. It is therefore termed a tricycle, and can operate either as an engine or as a refrigerator. A review of continuous quantum heat engines has been published recently [8] and therefore is beyond the scope of this review.

Reciprocating engines are classified according to their sequence of strokes. The most studied cycles are Carnot [9–13] and Otto cycles [14–19]. The quantum Otto cycle is easier to analyze, and therefore it became the primary example of a reciprocating quantum heat engine. The pioneering studies of quantum reciprocating engines employed a two-level system—a qubit—as a working medium [9,10,14,20]. The performance analysis of the quantum versions of reciprocating engines exhibited an amazing resemblance to macroscopic counterparts. For example, the efficiency at maximum power of the quantum version of the endoreversible engine converges at high temperature to the Novikov–Curzon–Ahlborn macroscopic Newtonian model predictions [21,22]. The deviations were even small at low temperature, despite the fact that the heat transport law was different [9]. The only quantum feature that could be identified was related to the discrete structure of the energy levels.

Heat engines with quantum features require a more complex working medium than a single qubit weakly coupled to a heat bath. This complexity is required to obtain quantum analogues of friction and heat leaks. A prerequisite for such phenomena is that the external control part of the Hamiltonian does not commute with the internal part. This generates quantum non-adiabatic phenomena which lead to friction [23]. A working medium composed of a quantum harmonic oscillator has sufficient complexity to represent generic phenomena, but can still be amenable to analytic analysis [24].

The quantum Otto cycle is a primary example of the emerging field of quantum thermodynamics. The quest is to establish the similarities and differences in applying thermodynamic reasoning up to the level of a single quantum entity. The present analysis is based on the theory of quantum open systems [25,26]. A dynamical description based on the weak system bath coupling has been able to establish consistency between quantum mechanics and the laws of thermodynamics [25]. These links allow work and heat to be defined in the quantum regime [27]. This framework is sufficient for the present analysis.

In the strong coupling regime where the partition between system and bath is not clear, the connection to thermodynamics is not yet established—although different approaches have been suggested [28,29]. A different approach to quantum thermodynamics termed quantum thermodynamics resource theory follows ideas from quantum information resource theory, establishing a set of rules [30,31]. We will try to show how this approach can be linked to the Otto cycle under analysis.

2. The Quantum Otto Cycle

Nicolaus August Otto invented a reciprocating four stroke engine in 1861, and won a gold medal in the 1867 Paris world fair [32]. The basic components of the engine are hot and cold reservoirs, a working medium, and a mechanical output device. The cycle of the engine is defined by four segments:
1. The hot isochore: heat is transferred from the hot bath to the working medium without volume change.

2. The power adiabat: the working medium expands, producing work, while isolated from the hot and cold reservoirs.

3. The cold isochore: heat is transferred from the working medium to the cold bath without volume change.

4. The compression adiabat: the working medium is compressed, consuming power while isolated from the hot and cold reservoirs, closing the cycle.

Otto determined that the efficiency $\eta$ of the cycle is limited to

$$\eta_o \leq 1 - \left( \frac{V_c}{V_h} \right) \left( \frac{C_p}{C_v} - 1 \right),$$

where $V_c/V_h$ and $T_c/T_h$ are the volume and temperature of the working medium at the end of the hot and cold isochores, respectively. $C_p$ and $C_v$ are the heat capacities under constant pressure and constant volume [33].

As expected, Otto efficiency is always smaller than the efficiency of the Carnot cycle $\eta_o \leq \eta_c = 1 - T_c/T_h$.

The first step in learning from an example is to establish a quantum version of the Otto cycle. This is carried out by seeking analogues for each segment of the cycle. What makes the approach unique is that it is applicable to a small quantum system such as a single atom in a harmonic trap. The description is embedded in the theory of open quantum systems. Each of these segments is defined by a completely positive (CP) propagator [34] describing the change of state in the working medium:

$$\hat{\rho}_{f} = U_{i\rightarrow f} \hat{\rho}_{i},$$

where the density operator $\hat{\rho}$ describes the state of the working medium.

The quantum engine Otto cycle is therefore described as:

1. The hot isochore: heat is transferred from the hot bath to the working medium without change in the external parameter $\omega_h$. The stroke is described by the propagator $U_h$.

2. The expansion adiabat: the working medium reduces its energy scale. The harmonic frequency changes from $\omega_h$ to $\omega_c$, with $\omega_h > \omega_c$, producing work while isolated from the hot and cold reservoirs. The stroke is described by the propagator $U_{hc}$.

3. The cold isochore: heat is transferred from the working medium to the cold bath without change in the external parameter $\omega_c$. The stroke is described by the propagator $U_c$.

4. The compression adiabat: the working medium increases its energy scale. The harmonic frequencies increase from $\omega_c$ to $\omega_h$, consuming power while isolated from the hot and cold reservoirs. The stroke is described by the propagator $U_{ch}$.

The cycle propagator becomes the product of the segment propagators:

$$U_{cyc} = U_{ch} U_c U_{hc} U_h.$$  \hspace{1cm} (1)

The cycle propagator is a completely positive (CP) map of the state of the working medium [34]. The order of propagators is essential, since the segment propagators do not commute; for example, $[U_h, U_{hc}] \neq 0$. The non-commuting property of the segment propagators is not an exclusive quantum property. It is also present in stochastic descriptions of the engine where the propagators operate on a vector of populations of the energy eigenvalues. Nevertheless, it can have a quantum origin for engines with propagators with small action [35,36]. The same operators but with different parameters (such as different frequencies) can be used to describe an Otto refrigeration cycle. Figure 1 shows a schematic representation of the Otto cycle in phase space.

In the adiabatic limit when the population $N$ stays constant in the expansion and compression segments, the work per cycle becomes:

$$W_{cyc} = \hbar \Delta \omega \Delta N,$$  \hspace{1cm} (2)

where $\Delta \omega = \omega_h - \omega_c$ and $\Delta N$ is the population difference. $\Delta N = \Delta N_c = \Delta N_h$, since the cycle is periodic. Under these conditions, the efficiency becomes:

$$\eta_o = 1 - \frac{\omega_c}{\omega_h} \leq \eta_c,$$  \hspace{1cm} (3)
where $\eta_c$ is the Carnot efficiency $\eta_c = 1 - \frac{T_c}{T_h}$. At this stage, it is also useful to define the compression ratio $C = \frac{\omega_h}{\omega_c}$.

2.1. Quantum Dynamics of the Working Medium

The quantum analogue of the Otto cycle requires a dynamical description of the working medium, the power output, and the heat transport mechanism.

A particle in a harmonic potential will constitute our working medium. This choice is amenable to analytic solutions and has sufficient complexity to serve as a generic example. Even a single specimen is sufficient to realize the operation of an engine.

We can imagine a single particle in a harmonic trap $V(Q) = \frac{k}{2}Q^2$. Expansion and compression of the working medium is carried out by externally controlling the trap parameter $k(t)$. The energy of the particle is represented by the Hamiltonian operator:

$$\hat{H} = \frac{1}{2m} \hat{P}^2 + \frac{k(t)}{2} \hat{Q}^2,$$

where $m$ is the mass of the system and $\hat{P}$ and $\hat{Q}$ are the momentum and position operators. All thermodynamical quantities will be intensive; i.e., normalized to the number of particles. In the macroscopic Otto engine, the internal energy of the working medium during the adiabatic expansion is inversely proportional to the volume. In the harmonic oscillator, the energy is linear in the frequency $\omega(t) = \sqrt{k(t)/m}$ [37]. This therefore plays the role of inverse volume $\frac{1}{V}$.

The Hamiltonian (4) is the generator of the evolution on the adiabatic segments. The frequency $\omega$ changes from $\omega_h$ to $\omega_c$ in a time period $\tau_{hc}$ in the power adiabat ($\omega_h > \omega_c$) and from $\omega_c$ to $\omega_h$ in
a period $\tau_{\text{c}}$ in the compression adiabat. The dynamics of the state $\hat{\rho}$ during the adiabatic segments is unitary and is the solution of the Liouville von Neumann equation [38]:

$$\frac{d}{dt}\hat{\rho}(t) = -\frac{i}{\hbar}[\hat{H}(t), \hat{\rho}(t)],$$  \hfill (5)

where $\hat{H}$ is time dependent during the evolution. Notice that $[\hat{\rho}(t), \hat{H}(t')] \neq 0$, since the kinetic energy does not commute with the varying potential energy. This is the origin of quantum friction [23,39].

The formal solution to Equation (5) defines the propagator:

$$\hat{U}(t)\hat{\rho}(0) = \hat{U}(t)\hat{\rho}(0)\hat{U}^\dagger,$$  \hfill (6)

where $\hat{U}$ satisfies the equation:

$$i\hbar \frac{d}{dt} \hat{U} = \hat{H}(t)\hat{U},$$  \hfill (7)

with the initial condition $\hat{U}(0) = \hat{I}$.

The dynamics on the hot and cold isochores is a thermalization process of the working medium with a bath at temperature $T_h$ or $T_c$. The dynamics of an open quantum system, where the working medium is described explicitly and the bath implicitly [40–42]:

$$\frac{d}{dt}\hat{\rho}(t) = -\frac{i}{\hbar}[\hat{H}, \hat{\rho}] + \mathcal{L}_D(\hat{\rho}),$$  \hfill (8)

where $\mathcal{L}_D$ is the dissipative term responsible for driving the working medium to thermal equilibrium, while the Hamiltonian $\hat{H} = \hat{H}(\omega_{h/c})$ is static. The equilibration is not complete in typical operating conditions, since only a finite time $\tau_h$ or $\tau_c$ is allocated to the hot or cold isochores. The dissipative “superoperator” $\mathcal{L}_D$ must conform to Lindblad’s form for a Markovian evolution [40,41], and for the harmonic oscillator can be expressed as [43–45]:

$$\mathcal{L}_D(\hat{\rho}) = k_\uparrow(\hat{a}^\dagger\hat{\rho}\hat{a} - \frac{1}{2}\{\hat{a}^\dagger\hat{a}, \hat{\rho}\}) + k_\downarrow(\hat{a}\hat{\rho}\hat{a}^\dagger - \frac{1}{2}\{\hat{a}\hat{a}^\dagger, \hat{\rho}\}),$$  \hfill (9)

where anticommutator $\{\hat{A}, \hat{B}\} \equiv \hat{A}\hat{B} + \hat{B}\hat{A}$. $k_\uparrow$ and $k_\downarrow$ are heat conductance rates obeying detailed balance $k_\uparrow/k_\downarrow = e^{-\Delta S/T}$, and $T$ is either $T_h$ or $T_c$. The operators $\hat{a}^\dagger$ and $\hat{a}$ are the raising and lowering operators, respectively. Notice that they are different in the hot and cold isochores, since

$$\hat{a} = \frac{1}{\sqrt{2}}(\sqrt{\frac{m}{\hbar\omega}}\hat{Q} + i\sqrt{\frac{1}{m\omega}}\hat{P})$$

depends on $\omega$. Formally for the isochores $\hat{U}_{h/c} = \exp(\mathcal{L}\tau)$ where $\mathcal{L} = -i/\hbar[\hat{H},\cdot] + \mathcal{L}_D$.

Equation (9) is known as a quantum Master equation [42] or L-GKS [40,41]. It is an example of a reduced description where the dynamics of the working medium is sought explicitly while the baths are described implicitly by two parameters: the heat conductivity $\Gamma = k_\downarrow - k_\uparrow$ and the bath temperature $T$. The Lindblad form of Equation (9) guarantees that the density operator of the extended system (system + bath) remains positive (i.e., physical) [40]. Specifically, for the harmonic oscillator, Equation (9) has been derived from first principles by many authors [44,46–49].

To summarize, the quantum model of the Otto cycle is composed of a working fluid of harmonic oscillators (4). The power stroke is modeled by the Liouville von Neumann Equation (5), while the heat transport via a Master Equations (8) and (9).

3. Quantum Thermodynamics

Thermodynamics is notorious for its ability to describe a process employing an extremely small number of variables. In scenarios where systems are far from thermal equilibrium, further variables have to be added. The analogue description in quantum thermodynamics is based on a minimal set
of quantum expectations $\langle \hat{X}_n \rangle$, where $\langle \hat{X}_n \rangle = \text{Tr}\{\hat{X}_n\hat{\rho}\}$. The dynamics of this set is generated by the Heisenberg equations of motion

$$\frac{d}{dt}\hat{X} = \frac{\partial \hat{X}}{\partial t} + i\hbar^{-1}[\hat{H}, \hat{X}] + \mathcal{L}_D^*(\hat{X}),$$

(10)

where the first term addresses an explicitly time-dependent set of operators, $\hat{X}(t)$.

The dynamical approach to quantum thermodynamics [25] seeks the relation between thermodynamical laws and their quantum origin. The first law of thermodynamics is equivalent to the energy balance relation. The energy expectation $E$ is obtained when $\hat{X} = \hat{H}$, i.e., $E = \langle \hat{H} \rangle$. The quantum energy partition defining the first law of thermodynamics, $dE = dW + dQ$, is obtained by inserting $\hat{H}$ into (10) [5,6,25,50]:

$$\frac{d}{dt}E = \dot{W} + \dot{Q} = \langle \frac{\partial \hat{H}}{\partial t} \rangle + \langle \mathcal{L}_D^*(\hat{H}) \rangle.$$

(11)

The power is identified as

$$\mathcal{P} = \dot{W} = \langle \frac{\partial \hat{H}}{\partial t} \rangle.$$

The heat exchange rate becomes

$$\frac{d}{dt}Q = \langle \mathcal{L}_D^*(\hat{H}) \rangle.$$

The analysis of the Otto cycle benefits from the simplification that power is produced or consumed only on the adiabats and heat transfer takes place only on the isochores.

The thermodynamic state of a system is fully determined by the thermodynamical variables. Statistical thermodynamics adds the prescription that the state is determined by the maximum entropy condition subject to the constraints set by the thermodynamical observables [51–53]. Maximizing the von Neumann entropy [38]

$$S_{VN} = -k_B\text{Tr}\{\hat{\rho} \ln(\hat{\rho})\}$$

subject to the energy constraint leads to thermal equilibrium [53]

$$\hat{\rho}_{eq} = \frac{1}{Z}e^{-\frac{\hat{H}}{k_BT}}.$$

(13)

where $k_B$ is the Boltzmann constant and $Z = \text{Tr}\{e^{-\frac{\hat{H}}{k_BT}}\}$ is the partition function.

In general, the state of the working medium is not in thermal equilibrium. In order to generalize the canonical form (13), additional observables are required to define the state of the system. The maximum entropy state subject to this set of observables [51,54] $\langle \hat{X}_i \rangle = \text{tr}\{\hat{X}_i\hat{\rho}\}$ becomes

$$\hat{\rho} = \frac{1}{Z} \exp \left( \sum_j \beta_j \hat{X}_j \right),$$

(14)

where $\beta_j$ are Lagrange multipliers. The generalized canonical form of (14) is meaningful only if the state can be cast in the canonical form during the entire cycle of the engine, leading to $\beta_j(t)$. This requirement is called canonical invariance [55]. It implies that if an initial state belongs to the canonical class of states, it will remain in this class throughout the cycle.

A necessary condition for canonical invariance is that the set of operators $\hat{X}$ in (14) is closed under the dynamics generated by the equation of motion. If this condition is satisfied, then the state of the system can be reconstructed from a small number of quantum observables $\langle \hat{X}_i \rangle(t)$. These become the thermodynamical observables, since they define the state under the maximum entropy principle.
The condition for canonical invariance on the unitary part of the evolution taking place on the adiabats is as follows: if the Hamiltonian is a linear combination of the operators in the set \( \hat{H}(t) = \sum_m h_m \hat{X}_m \) \((h_m(t) \) are expansion coefficients), and the set forms a closed Lie algebra \( [\hat{X}_j, \hat{X}_k] = \sum_l C_{jk}^l \hat{X}_l \) (where \( C_{jk}^l \) is the structure factor of the Lie algebra), then the set \( \hat{X} \) is closed under the evolution \([56]\).

For a closed Lie algebra, the generalized Gibbs state Equation (14) can always be written in a product form:

\[
\hat{\rho} = \prod_k e^{\lambda_k \hat{X}_k},
\]

where there is a one-to-one relation between \( \lambda \) and \( \beta \), depending on the order of the product form. Multiplying the equation of motion by \( \hat{\rho}^{-1} \) leads to \( \frac{d}{dt} \hat{\rho} \hat{\rho}^{-1} = \mathcal{L}(\hat{\rho}) \hat{\rho}^{-1} \). Using the product form and the Backer–Housdorff relation, the l.h.s. decomposes to a linear combination of the operator algebra. This is also true for the r.h.s \([\hat{H}, \hat{\rho}] \hat{\rho}^{-1} \), which also becomes a linear combination of the operator algebra. Comparing both sides of the equation of motion, one obtains a set of coupled differential equations for the coefficients \( \lambda_k \). Their solution guarantees that canonical invariance prevails \([54]\).

For the harmonic Otto cycle, the set of the operators \( \hat{P}^2, \hat{Q}^2, \) and \( \hat{D} = \frac{1}{2}(\hat{Q}\hat{P} + \hat{P}\hat{Q}) \) form a closed Lie algebra. Since the Hamiltonian is a linear combination of the first two operators of the set \( (\hat{P}^2 \) and \( \hat{Q}^2) \), canonical invariance will prevail on the adiabatic segments.

On the isochores, the set of operators also has to be closed to the operation of \( \mathcal{L}_D \). The set \( \hat{P}^2, \hat{Q}^2, \) and \( \hat{D} \) is closed to \( \mathcal{L}_D \), defined by (9). For canonical invariance of \( \hat{\rho} \), \( \mathcal{L}_D \hat{\rho} \hat{\rho}^{-1} \) should also be a linear combination of operators in the algebra. For the harmonic working medium and \( \mathcal{L}_D \) defined in (9), this condition is fulfilled. As a result, canonical invariance with the set of operators \( \hat{P}^2, \hat{Q}^2, \) and \( \hat{D} = \frac{1}{2}(\hat{Q}\hat{P} + \hat{P}\hat{Q}) \) prevails for the whole cycle \([24]\).

The significance of canonical invariance is that a solution of the operator dynamics allows the reconstruction of the state of the working medium during the whole cycle. As a result, all dynamical quantities become functions of a very limited set of thermodynamic quantum observables \( \langle \hat{X}_j \rangle \). The choice of a set of operators \( \{\hat{X}_j\} \) should reflect the most essential thermodynamical variables.

The operator algebra forms a vector space with the scalar product \( \langle \hat{X}_j \cdot \hat{X}_k \rangle = Tr\{\hat{X}_j\hat{X}_k\} \). This vector space will be used to describe the state \( \hat{\rho} \) and define the cycle propagators \( \mathcal{U}_l \). This description is a significant reduction in the dimension of the propagator \( \mathcal{U} \) from \( N^2 \), where \( N \) is the size of Hilbert space to \( M \) the size of the operator algebra.

Explicitly, variables with thermodynamical significance are chosen for the harmonic oscillator. These variables are time-dependent and describe the current state of the working medium:

- The Hamiltonian \( \hat{H}(t) = \frac{1}{2m} \hat{P}^2 + \frac{1}{2}m \omega(t)^2 \hat{Q}^2 \).
- The Lagrangian \( \hat{L}(t) = \frac{1}{2m} \hat{P}^2 - \frac{1}{2}m \omega(t)^2 \hat{Q}^2 \).
- The position momentum correlation \( \hat{C}(t) = \frac{1}{2} \omega(t) (\hat{Q}\hat{P} + \hat{P}\hat{Q}) = \omega(t) \hat{D} \).

These operators are linear combinations of the same Lie algebra as \( \hat{Q}^2, \hat{P}^2, \) and \( \hat{D} \). A typical cycle in terms of these variables is shown in Figure 2.

In the algebra of operators, a special place can be attributed to the Casimir operator \( \hat{C} \). This Casimir commutes with all the operators in the algebra \([57,58]\). Explicitly, it becomes:

\[
\hat{C} = \frac{\hat{H}^2 - \hat{L}^2 - \hat{C}^2}{\hbar^2 \omega^2} = -\frac{(\hat{Q}\hat{P} + \hat{P}\hat{Q})^2 + 2\hat{P}\hat{Q}\hat{P}\hat{Q} + 2\hat{Q}\hat{P}\hat{Q}}{4\hbar^2}.
\]

Since \( [\hat{H}, \hat{C}] = 0 \), \( \hat{C} \) is constant under the evolution of the unitary segments generated by \( \hat{H} \). The Casimir for the harmonic oscillator is a positive operator with a minimum value determined by the uncertainty relation: \( \langle \hat{C} \rangle \geq \frac{\hbar}{4} \) \([59]\).
Figure 2. Otto refrigeration cycle displayed in the thermodynamical variables $\hat{H}, \hat{L}, \hat{C}$. When the working medium is in contact with a hot bath, the system exhausts heat and equilibrates, spiralling downward from a high $\langle \hat{H} \rangle$ (energy) value and towards zero correlation $\langle \hat{C} \rangle$ and Lagrangian $\langle \hat{L} \rangle$ (i.e., towards thermal equilibrium). The hot isochore is marked by the red dotted line $A \to D$. On the expansion adiabat, the system spirals downwards, losing energy as it cools down—marked by the green line $D \to C$. It then spirals upwards (blue line), gaining energy from the cold bath $C \to B$. In addition, it spirals towards zero $\langle \hat{C} \rangle$ and $\langle \hat{L} \rangle$. Then, the compression adiabat (black line) takes it back to the top of the hot (red) spiral $B \to A$.

A related invariant to the dynamics is the Casimir companion [59], which for the harmonic oscillator is defined as:

$$\chi' = \frac{\langle \hat{H} \rangle^2 - \langle \hat{L} \rangle^2 - \langle \hat{C} \rangle^2}{\hbar^2 \omega^2}.$$ \tag{17}

Combining Equations (16) and (17), an additional invariant to the dynamics can be defined:

$$\frac{1}{\hbar^2 \omega^2} \left( \text{Var}(\hat{H}) - \text{Var}(\hat{L}) - \text{Var}(\hat{C}) \right) = \text{const}, \tag{18}$$

where $\text{Var}(\hat{A}) = \langle \hat{A}^2 \rangle - \langle \hat{A} \rangle^2$.

Coherence is an important quantum feature. The coherence is characterised by the deviation of the state of the system from being diagonal in energy [60,61], and it can be defined as:

$$C_0 = \frac{1}{\hbar \omega} \sqrt{\langle \hat{L}^2 \rangle + \langle \hat{C}^2 \rangle}.$$ \tag{19}

From Equation (16), we can deduce that increasing coherence has a cost in energy $\Delta E = \hbar \omega C_0$.

For the closed algebra of operators, the canonical state of the system $\hat{\rho}$ can be cast into the product form [24,62]. This state $\hat{\rho}$ is defined by the parameters $\beta, \gamma$, and $\gamma^*$:

$$\hat{\rho} = \frac{1}{Z} e^{\beta \hat{a} \hat{a}^\dagger} e^{-\beta \hat{a}^\dagger \hat{a}} e^{\gamma \hat{C}^\dagger \hat{C}} e^{-\gamma \hat{L}^\dagger \hat{L}}, \tag{20}$$

where $\hat{H} = \hbar \omega (\hat{a} \hat{a}^\dagger + \hat{a}^\dagger \hat{a})$, $\hat{C} = -i \frac{\hbar \omega}{2} (\hat{a}^2 - \hat{a}^\dagger 2)$, $\hat{L} = -\frac{\hbar \omega}{2} (\hat{a}^2 + \hat{a}^\dagger 2)$, and...
\[ Z = \frac{e^{\beta \omega}}{(e^{\beta \omega} - 1)^{\frac{1}{2}}}. \]  

Equation (21) can be inverted, leading to

\[ e^{\beta \omega} = \frac{\langle \hat{L} \rangle^2 + \langle \hat{C} \rangle^2 - \langle \hat{H} \rangle^2 + \frac{\hbar^2 \omega^2}{4}}{\langle \hat{L} \rangle^2 + \langle \hat{C} \rangle^2 - \langle \hat{H} \rangle^2}. \]  

Equations (23) and (24) relate the state of the system \( \hat{\rho} \) (by Equation 20) to the thermodynamical observables \( \langle \hat{H} \rangle, \langle \hat{L} \rangle, \) and \( \langle \hat{C} \rangle \).

The generalized canonical state of the system Equation (20) is equivalent to a squeezed thermal state [63]:

\[ \hat{\rho} = \mathcal{S}(\gamma) \frac{1}{Z} e^{-\beta \hat{H}} \mathcal{S}^\dag(\gamma), \]  

with the squeezing operator \( \mathcal{S}(\gamma) = \exp(\frac{i}{2} (\gamma \hat{a}^2 - \gamma^* \hat{a}^\dag \hat{a})). \) This state is an example of a generalized Gibbs state subject to non-commuting constraints [64,65]. Figure 1 shows examples of such states, which all have a Gaussian shape in phase space.

**Entropy Balance**

In thermodynamics, the entropy \( S \) is a state variable. Shannon introduced entropy as a measure of missing information required to define a probability distribution \( p \) [66]. The information entropy can be applied to a complete quantum measurement of an observable represented by the operator \( \hat{O} \) with possible outcomes \( p_j \):

\[ S_{\hat{O}} = -k_B \sum_j p_j \ln p_j, \]  

where \( p_j = \text{Tr}\{\hat{P}_j \hat{\rho}\} \). The projections \( \hat{P}_j \) are defined using the spectral decomposition theorem \( \hat{O} = \sum \lambda_i \hat{P}_j \), where \( \lambda_i \) are the eigenvalues of the operator \( \hat{O} \). \( S_{\hat{O}} \) is then the measure of information gain obtained by the measurement.

The von Neumann entropy [38] is equivalent to the minimum entropy \( S_{\hat{Y}_n} \) associated with a complete measurement of the state \( \hat{\rho} \) by the observable \( \hat{Y}_n \), where the set of operators \( \hat{Y}_n \) includes all possible non-degenerate operators in Hilbert space. The operator that minimizes the entropy commutes with the state \( [\hat{\rho}, \hat{Y}_{n\text{min}}] = 0 \). This leads to a common set of projectors of \( \hat{Y}_{\text{min}} \) and \( \hat{\rho} \); therefore,

\[ S_{VN} = -\text{Tr}\{\hat{\rho} \ln \hat{\rho}\}, \]  

which is a function of the state only. Obviously, \( S_{VN} \leq S_{\hat{O}} \). This provides the interpretation that \( S_{VN} \) is the minimum information required to completely specify the state \( \hat{\rho} \).

The primary thermodynamic variable for the heat engine is energy. The entropy associated with the measurement of energy \( S_E = S_{H} \) in general differs from the von Neumann entropy \( S_E \geq S_{VN} \). Only when \( \hat{\rho} \) is diagonal in the energy representation—such as in thermal equilibrium (13)—\( S_E = S_{VN} \).
The relative entropy between the state and its diagonal representation in the energy eigenfunctions is an alternative measure of coherence [67]:

\[
D(\hat{\rho}||\hat{\rho}_{ed}) = \text{Tr}\{\hat{\rho}(\ln\hat{\rho} - \ln\hat{\rho}_{ed})\},
\]

where \(\hat{\rho}_{ed}\) is the state composed of the energy projections which has the same populations of the energy levels as state \(\hat{\rho}\). The conditional distance \(D(\hat{\rho}||\hat{\rho}_{ed})\) is equivalent to the difference between the energy entropy \(S_E\) and the von Neumann entropy \(S_{VN}\): \(D(\hat{\rho}||\hat{\rho}_{ed}) = S_E - S_{VN} \geq 0\).

The von Neumann entropy is invariant under unitary evolution [54]. This is the result of the property of unitary transformations, where the set of eigenvalues of \(\hat{\rho}' = \hat{U}\hat{\rho}\hat{U}^\dagger\) is equal to the set of eigenvalues of \(\hat{\rho}\). Since the von Neumann entropy \(S_{VN}\) is a functional of the eigenvalues of \(\hat{\rho}\), it becomes invariant to any unitary transformation.

When the unitary transformation is generated by members of the Lie algebra, the Casimir is invariant. The von Neumann entropy of the generalized Gibbs state (20) is a function of the Casimir \(\langle \hat{G} \rangle\) [68] so that in this case it also becomes constant:

\[
S_{VN} = \ln \left( \sqrt{\langle \hat{G} \rangle} - \frac{1}{4} \right) + \sqrt{\langle \hat{G} \rangle} \ln \left( \frac{\sqrt{\langle \hat{G} \rangle}}{\langle \hat{G} \rangle - \frac{1}{4}} \right).
\]

An alternative expression for the \(S_{VN}\) entropy is calculated from the covariance matrix of Gaussian canonical states [69–71]:

\[
S_{VN} = \frac{\nu + 1}{2} \ln(\frac{\nu + 1}{2}) - \frac{\nu - 1}{2} \ln(\frac{\nu - 1}{2}),
\]

where \(\nu = \frac{2}{\pi} \sqrt{\sigma}\),

\[
\sigma = \begin{vmatrix}
\sigma_{pp} & \sigma_{pq} \\
\sigma_{qp} & \sigma_{qq}
\end{vmatrix},
\]

and \(\sigma_{ij}\) is the covariance.

The energy entropy \(S_E\) of the oscillator (not in equilibrium) is found to be equivalent to the entropy of an oscillator in thermal equilibrium with the same energy expectation value:

\[
S_E = \frac{1}{\hbar \omega} \left( \langle \hat{H} \rangle + \frac{\hbar \omega}{2} \right) \ln \left( \frac{\langle \hat{H} \rangle + \frac{\hbar \omega}{2}}{\langle \hat{H} \rangle - \frac{\hbar \omega}{2}} \right) - \ln \left( \frac{\hbar \omega}{\langle \hat{H} \rangle} - \frac{\hbar \omega}{2} \right).
\]

\(S_E\) in (30) is completely determined by the energy expectation \(E = \langle \hat{H} \rangle\). As an extreme example, for a squeezed pure state, \(S_{VN} = 0\) and \(S_E \geq 0\).

In a macroscopic working medium, the internal temperature can be defined from the entropy and energy variables \(1/T_{int} = \left( \frac{\partial S}{\partial E} \right)_V\) at constant volume. For the quantum Otto cycle, \(S_E\) is used to define the inverse internal temperature \(1/T_{int} = \left( \frac{\partial S}{\partial E} \right)_\omega\). \(T_{int}\) is a generalized temperature appropriate for non equilibrium density operators \(\hat{\rho}\). Using this definition, the internal temperature \(T_{int}\) of the oscillator working medium can be calculated implicitly from the energy expectation:

\[
E = \frac{1}{2} \hbar \omega \coth \left( \frac{\hbar \omega}{2k_B T_{int}} \right),
\]

which is identical to the equilibrium relation between temperature and energy in the harmonic oscillator. This temperature defines the work required to generate the coherence: \(W_c = k_B T_{int}(S_E - S_{VN})\) [39].
### 4. The Dynamics of the Quantum Otto Cycle

A quantum heat engine is a dynamical system subject to the tradeoff between efficiency and power. The dynamics of the reciprocating Otto cycle can be partitioned to the four strokes and later combined to generate the full cycle. Each of the segments influences the final performance: power extraction or refrigeration. The performance of the cycle can be optimized with respect to efficiency and power. Each segment can be optimized separately, and finally a global optimization is performed. The first step is to describe the dynamics of each segment in detail.

#### 4.1. Heisenberg Dynamics of Thermalisation on the Isochores

The task of the isochores is to extract and reject heat from thermal reservoirs. The dynamics of the working medium is dominated by an approach to thermal equilibrium. In the Otto cycle, the Hamiltonian \( \hat{H} \) is constant (\( \omega = \omega_{h/c} \) is constant). The Heisenberg equations of motion generating the dynamics for an operator \( \hat{X} \) become:

\[
\frac{d}{dt} \hat{X} = \frac{i}{\hbar} \{ \hat{H}, \hat{X} \} + k_\downarrow (\hat{a}^\dagger \hat{X} \hat{a} - \frac{1}{2} \{ \hat{a}^\dagger \hat{a}, \hat{X} \}) + k_\uparrow (\hat{a} \hat{X} \hat{a}^\dagger - \frac{1}{2} \{ \hat{a} \hat{a}^\dagger, \hat{X} \}). \tag{32}
\]

Equation (32) is the analogue of (8) and (9) in the Schrödinger frame.

For the dynamical set of observables, the equations of motion become:

\[
\frac{d}{dt} \begin{pmatrix} \hat{H} \\ \hat{L} \\ \hat{C} \\ \hat{I} \end{pmatrix} (t) = \begin{pmatrix} \Gamma & 0 & 0 & \Gamma \langle \hat{H} \rangle_{eq} \\ 0 & -\Gamma & -2\omega & 0 \\ 0 & 2\omega & -\Gamma & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{H} \\ \hat{L} \\ \hat{C} \\ \hat{I} \end{pmatrix} (t), \tag{33}
\]

where \( \Gamma = k_\downarrow - k_\uparrow \) is the heat conductance and \( k_\uparrow/k_\downarrow = e^{-h\omega/k_B T} \) obeys detailed balance where \( \omega = \omega_{h/c} \) and \( T = T_{h/c} \) are defined for the hot or cold bath, respectively. From (11), the heat current can be identified as:

\[
\dot{Q} = -\Gamma (\langle \hat{H} \rangle - \langle \hat{H} \rangle_{eq}) = \Gamma \frac{h\omega}{2} \left( \coth\left( \frac{h\omega}{2k_B T_B} \right) - \coth\left( \frac{h\omega}{2k_B T_{int}} \right) \right), \tag{34}
\]

where \( T_B \) is the bath temperature. In the high temperature limit, the heat transport law becomes Newtonian [22]: \( \dot{Q} = \Gamma (k_B T_B - k_B T_{int}) \).

The solution of isochore dynamics (33) generates the propagator defined on the vector space of the observables \( \hat{H}, \hat{L}, \hat{C}, \hat{I} \): The propagator on the isochore has the form [68,72]:

\[
U_{h/c} = \begin{pmatrix} R & 0 & 0 & H_{eq}(1 - R) \\ 0 & R_c & -R_s & 0 \\ 0 & R_s & R_c & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \tag{35}
\]

where \( R = e^{-\Gamma t}, c = \cos(2\omega t), s = \sin(2\omega t), \) and \( H_{eq} = \frac{\hbar \omega}{e^{\hbar \omega/k_BT} - 1} \). It is important to note that the propagator on the isochores does not generate coherence from energy \( \hat{H} \). The coherence Equation (19) is a function of the expectations of \( \hat{L}, \hat{C} \), which are not coupled to \( \hat{H} \).

#### 4.2. The Dynamics on the Adiabats and Quantum Friction

The dynamics on the adiabats is generated by a time-dependent Hamiltonian. The task is to change the energy scale of the working medium from one bath to the other. The oscillator frequency changes from \( \omega_h \) to \( \omega_c \) on the power expansion segment and from \( \omega_c \) to \( \omega_h \) on the compression segment.
The Hamiltonian—which is explicitly time-dependent—does not commute with itself at different times \([\hat{H}(t), \hat{H}(t')] \neq 0\). As a result, coherence is generated with an extra cost in energy.

The Heisenberg equations of motion (10) for the dynamical set of operators are expressed as \([68,72]\):

\[
\frac{d}{dt} \begin{pmatrix} \hat{H} \\ \hat{L} \\ \hat{C} \\ \hat{I} \end{pmatrix} (t) = \omega(t) \begin{pmatrix} \mu & -\mu & 0 & 0 \\ -\mu & \mu & -2 & 0 \\ 0 & 2 & \mu & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \hat{H} \\ \hat{L} \\ \hat{C} \\ \hat{I} \end{pmatrix} (t),
\]

(36)

where \(\mu = \frac{\omega(t)}{\omega_0}\) is a dimensionless adiabatic parameter. In general, all operators in (36) are dynamically coupled. This coupling is characterized by the non-adiabatic parameter \(\mu\). When \(\mu \to 0\), the energy decouples from the coherence and the cycle can be characterized by \(p_n\)—the probability of occupation of energy level \(e_n\).

Power is obtained from the first-law (11) as:

\[
\mathcal{P} = \mu \omega \left( \langle \hat{H} \rangle - \langle \hat{L} \rangle \right).
\]

(37)

Power on the adiabats (37) can be decomposed to the “useful” external power \(\mathcal{P}_{ex} = \mu \omega \langle \hat{H} \rangle\) and to the power invested to counter friction \(\mathcal{P}_f = -\mu \omega \langle \hat{L} \rangle\) if \(\langle \hat{L} \rangle > 0\). Under adiabatic conditions \(\mu \to 0\), \(\langle \hat{L} \rangle = 0\), since no coherence is generated; therefore, \(\mathcal{P}_f = 0\). Generating coherence consumes power when the initial state is diagonal in energy \([\hat{p}, \hat{H}] = 0\) \([73,74]\).

Insight on the adiabatic dynamics can be obtained from the closed-form solution of the dynamics when the non-adiabatic parameter \(\mu = \omega(t) / \omega_0\) is constant. This leads to the explicit time dependence of the control frequency \(\omega: \omega(t) = \omega(0) (1 - \mu/2)\). Under these conditions, the matrix in Equation (36) becomes stationary. This allows a closed-form solution to be obtained by diagonalizing the matrix. Under these conditions, the adiabatic propagator \(\mathcal{U}_a\) has the form:

\[
\mathcal{U}_a = \frac{\omega(t)}{\omega(0)} \frac{1}{\Omega} \begin{pmatrix} 4 - \mu^2 c & -\mu \Omega s & -2\mu(c-1) & 0 \\ -\mu \Omega s & \Omega^2 c & -2\Omega s & 0 \\ 2\mu(c-1) & 2\Omega s & 4c - \mu^2 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},
\]

(38)

where \(\Omega = \sqrt{4 - \mu^2}\) and \(c = \cos(\Omega \theta(0))\), \(s = \sin(\Omega \theta(0))\), and \(\theta(t) = -\frac{1}{\mu} \log \left( \frac{\omega(t)}{\omega(0)} \right)\) \([68]\). For \(|\mu| > 2\), the solutions are oscillatory. For \(|\mu| > 2\), the sin and cos functions become sinh and cosh. More on the transition point from damped to over-damped dynamics is presented in Section 4.5.

The difference between the expansion adiabats \(\mathcal{U}_{he}\) and compression adiabats \(\mathcal{U}_{he}\) is in the sign of \(\mu\) and the ratio \(\omega(t)/\omega(0)\). The propagator \(\mathcal{U}_a\) can be viewed as a product of a changing energy scale by the factor \(\omega(t)/\omega(0)\) and a propagation in a moving frame generated by a constant matrix. The fraction of additional work on the adiabats with respect to the adiabatic solution is causing extra energy invested in the working medium, which is defined as:

\[
\delta_f = \frac{\omega_i}{\omega_f} (\mathcal{U}_a(1,1) - \omega_f / \omega_i).
\]

(39)

For the case of \(\mu\) constant:

\[
\delta_f = \frac{2\mu^2 \sin(\frac{\theta(t)}{2})^2}{4 - \mu^2}
\]

(40)

and \(\delta_f \geq 0\).
A different approach to the deviation from adiabatic behaviour has been based on a general propagator for Gaussian wavefunctions of the form [75]:

$$\psi(x, t) = \exp \left( \frac{i}{\hbar} (a(t)x^2 + b(t)x + c(t)) \right).$$  \hfill (41)

$a(t)$ can be mapped to a time-dependent classical harmonic oscillator: $a(t) = M \dot{X}/X$, where:

$$m \frac{d^2}{dt^2} X + \omega(t)^2 X = 0.$$  \hfill (42)

The local adiabatic parameter is defined as:

$$Q^*(t) = \frac{1}{2 \omega_i \omega_f} \left( \omega_r^2 (\omega_r X(t)^2 + \ddot{X}(t)^2) + (\omega_r^2 Y(t)^2 + \ddot{Y}(t)^2) \right),$$  \hfill (43)

where $X(t)$ and $Y(t)$ are the solution of Equation (42) with the boundary conditions $X(0) = 0, \dot{X}(0) = 1$ and $Y(0) = 1, \dot{Y}(0) = 0$ for a constant frequency $Q^* = 1$. In general, the expectation value of the energy at the end of the adiabats becomes:

$$\langle \hat{H} \rangle_f = \frac{\omega_f}{\omega_i} Q^* \langle \hat{H} \rangle_i,$$  \hfill (44)

where $i/f$ correspond to the beginning and end of the stroke. In general, $Q^*(t)$ can be obtained directly from the solution of Equation (42). $Q^*(t)$ is related to $\delta_f$ by: $Q^*(t) = 1 + \delta_f$. For the case of $\mu$, constant $Q^*(t)$ can be obtained from Equation (40). In addition, $Q^*(t)$ can be obtained by the solution of the Ermakov equation (Equation (49)) [76].

The general dynamics described in Equation (38) mixes energy and coherence. As can be inferred from the Casimir Equation (16), generating coherence costs energy. This extra cost gets dissipated on the isochores, and is termed quantum friction [39,77]. The energy cost scales as $\mu^2$; therefore, slow operation (i.e., $|\mu| \ll 1$) will eliminate this cost. The drawback is large cycle times and low power. Further analysis of Equation (38) shows a surprising result. Coherence can be generated and consumed, resulting in periodic solutions in which the propagator becomes diagonal. As a result, mixing between energy and coherence is eliminated. These solutions appear when $\cos(\Omega \theta(\tau_a)) = 1$, where $\tau_a$ is the expansion or compression stroke time allocation. These periodic solutions can be characterised by a quantization relation [68]:

$$\mu^* = -\frac{2 \log(C)}{\sqrt{4 \pi^2 l^2 + \log(C)^2}},$$  \hfill (45)

where $C = \frac{\omega_f}{\omega_i}$ is the engine’s compression ratio, and $l$ the quantization number $l = 1, 2, 3, \ldots$, accompanied by the time allocation $\tau_{hc}^*$:

$$\tau_{hc}^* = \frac{1 - C}{\mu^* \omega_c}.$$  \hfill (46)

A frictionless solution with the shortest time is obtained for $l = 1$, and it scales as $\tau_{hc} \propto 1/\omega_c$. This observation raises the question: are there additional frictionless solutions in finite time? What is the shortest time that can achieve this goal?

The general solution of the dynamics depends on an explicit dependence of $\omega(t)$ on time. $\omega(t)$ can be used as a control function to optimise the performance, obtaining a state $\hat{\rho}$ diagonal in energy at the interface with the isochores. Such a solution will generate a frictionless performance. Operating at effective adiabatic conditions has been termed shortcut to adiabaticity [78–82].
The search for frictionless solutions has led to two main directions. The first is based on a time-dependent invariant operator $\tilde{I}(t)$ [82]:

$$\frac{d}{dt}\tilde{I}(t) = \frac{d}{dt}[\tilde{I}(t)] + \frac{i}{\hbar}[\tilde{H}(t), \tilde{I}(t)] = 0. \quad (47)$$

For the harmonic oscillator, the invariant is [78]:

$$\tilde{I}(t) = \frac{1}{2} \left( \frac{1}{b^2} \hat{Q}^2 m \omega_0^2 + \frac{1}{b} \hat{\beta}^2 \right), \quad (48)$$

where $\hat{\beta} = b\hat{P} - mb\hat{\dot{Q}}$. The invariant must satisfy $[\hat{H}, \tilde{I}] = 0$ for the initial $\tau_i$ and final $\tau_f$ times, then at these times the eigenstates of the invariant at the initial and final time of the adiabat are identical to those of the Hamiltonian [83,84]. This is obtained if $b(0) = 1$ and $\dot{b}(0) = 0$, $\ddot{b}(0) = 0$, as well as $b(\tau_f) = \sqrt{\omega_0/\omega_f}$ and $\dot{b}(\tau_f) = 0$, $\ddot{b}(\tau_f) = 0$. In addition, the function $b(t)$ satisfies the Ermakov equation:

$$\ddot{b} + \omega(t)^2 b = \omega_0^2/b^3. \quad (49)$$

The instantaneous frequency becomes $\omega(t) = \omega_0/b^2$ for $\dot{b}(0) = 0$. There are many solutions to the Ermakov equation, and additional constraints must be added—for example, that the frequency is at all times real and positive. These equations can be used to search for fast frictionless solutions. To obtain a minimal time $\tau_i$, some constrains have to be imposed. For example, limiting the average energy stored in the oscillator. In this case, $\tau_{hc}$ scales as $\tau_{hc} \propto 1/\sqrt{\omega_c}$ [80]. Other constraints on $\omega(t)$ have been explored. For example, the use of imaginary frequency corresponding to an inverted harmonic potential. These schemes allow faster times on the adiabat [80,85]. If the peak energy is constrained, $\tau_{hc}$ scales logarithmically with $1/\omega_c$; however, if the average energy is constrained, then the scaling becomes $\tau_{hc} \propto 1/\sqrt{\omega_c}$.

The second approach to obtain frictionless solutions is based on optimal control theory: finding the fastest frictionless solution where the control function is $\omega(t)$ [68,85–90]. Optimal control theory reveals that the problem of minimizing time is linear in the control, which is proportional to $\omega(t)$ [68]. As a result, the optimal control solution depends on the constraints $\omega_{max}$ and $\omega_{min}$. If these are set as $\omega_{max} = \omega_h$ and $\omega_{min} = \omega_c$, then the optimal time scales as $\tau_{opt} \propto \frac{1}{\omega_h \sqrt{\omega_c}}$. Other constraints will lead to faster times, but their energetic cost will diverge. This scaling is consistent considering the cost of the counter-adiabatic terms in frictionless solutions leading to the same scaling [91].

The optimal solution can be understood using a geometrical description [87]. The derivative of the change of $\langle \hat{Q}^2 \rangle$ with respect to the change in $\langle \hat{P}^2 \rangle$ becomes:

$$\frac{d}{d(\hat{P}^2)} \frac{d(\hat{Q}^2)}{d(\hat{P}^2)} = -\omega^2(t) \equiv v. \quad (50)$$

The time allocated to the change $\tau$ becomes:

$$\tau = \int \frac{d(\hat{P}^2)}{\sqrt{\langle \hat{P}^2 \rangle \langle \hat{Q}^2 \rangle - \langle \hat{G} \rangle}}, \quad (51)$$

where $\hat{G}$ is the Casimir defined in Equation (16). In addition, the control $v$ is constrained by $\omega_c^2 \leq v \leq \omega_h^2$. The initial and final $\langle \hat{P}^2 \rangle_{i/f} = mE_{i/f}$, since the initial and final $\langle \hat{L} \rangle$ and $\langle \hat{C} \rangle$ are zero. The minimum time is obtained by maximizing the product $\langle \hat{P}^2 \rangle \langle \hat{Q}^2 \rangle$ along the trajectory. The minimum time optimization leads to a bang-bang solution where the frequency is switched instantly from $\omega_h$ to $\omega_c$, as in the sudden limit Equation (58) is followed by a waiting period then switched back to $\omega_h$ until the target is reached.
and switched finally to $\omega_c$. The relation between the geometric optimization and the Ermakov equation of the shortcuts to adiabaticity has been obtained based on the geometrical optimization [92,93].

To summarize, frictionless solutions can be obtained in finite time. As a result, the engine can be completely described by the population of the energy eigenvalues or for the harmonic working medium by the expectation value of number operator $\hat{N}$. Employing reasonable constraints on the control function $\omega(t)$ results in the minimum time $\tau^*_\text{a}$ scaling as $O\left(\frac{1}{\sqrt{\omega_0/4\hbar}}\right)$.

### 4.3. The Influence of Noise on the Adiabats

The frictionless adiabat requires a very accurate protocol of $\omega(t)$ as a function of time. For any realistic devices, such a protocol will be subject to fluctuations in the external control. The controllers are subject to noise, which will induce friction-like behaviour. Can this additional friction be minimized? Insight on the effects of noise on the performance of the Otto cycle can be obtained by analysing a simple model based on the frictionless protocol with constant $\mu$ [94]. The obvious source of external noise is induced by fluctuations in the control frequency $\omega(t)$. This noise is equivalent to Markovian random fluctuations in the frequency of the harmonic oscillator. These errors are modelled by a Gaussian white noise. The dissipative Lindbland term generating such noise has the form [42,95]:

$$L_{\text{Na}}(\mathbf{A}) = -\gamma_a \omega^2 [\hat{B}, [\mathbf{B}, \hat{A}]],$$  \hspace{1cm} (52)

where $\hat{B} = m\omega \hat{Q}^2 / (2\hbar)$.

The influence of the amplitude noise generated by $L_{\text{Na}}(\mathbf{A}) = -\gamma_a \omega^2 [\hat{B}, [\mathbf{B}, \hat{A}]]$ is obtained by approximating the propagator by the product form $U_{\text{hc}} = U_a U_{\text{an}}$. The equations of motion for the amplitude noise $U_{\text{an}}$ are obtained from the interaction picture in Liouville space:

$$\frac{d}{dt} U_{\text{an}}(t) = U_a(-t) N_a(t) U_a(t) U_{\text{an}}(t) = \mathcal{W}_{\text{an}}(t) U_{\text{an}}(t),$$  \hspace{1cm} (53)

where $\mathcal{W}_{\text{an}}$ is the interaction propagator in Liouville space [94] and $U_a$ is the adiabatic propagator, Equation (38). A closed-form solution is obtained in the frictionless limit $\mu \to 0$ when $\mathcal{W}_a$ is expanded up to zero order in $\mu$:

$$W_a(t) \approx \gamma_a \omega_0 \begin{pmatrix} 1 & -c & s & 0 \\ c & -c^2 & cs & 0 \\ -s & cs & -s^2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$  \hspace{1cm} (54)

where $s = \sin(\Omega\Theta)$, $c = \cos(\Omega\Theta)$. The Magnus expansion [96] is employed to obtain the $l$ period propagator $U_{bl}(X = 2l\pi)$, where the periods are of the adiabatic propagator $U_a$ of Equation (38):

$$U_{bl}(X = 2l\pi) \approx e^{B_1 + B_2 + \ldots},$$  \hspace{1cm} (55)

where $B_1 = \int_0^{2\pi} dX W_a(X)$, $B_2 = \frac{1}{l} \int_0^{2\pi} \int_0^X dXdX' [W_a(X), W_a(X')]$, and so on. The first-order Magnus term leads to the propagator
where \( F = \left( \frac{16\omega_0^2}{1 - 16 \gamma^2} \right) \left( e^{2\pi l \mu} - 1 \right) \). For large \( l \), in Equation (45) the limit from hot to cold simplifies to: \( F = (\omega_h - \omega_c) \). The solution of Equation (53) shows that the fraction of work against friction \( \delta_f \) will diverge when \( l \to \infty \) or \( \mu \to 0 \), nulling the adiabatic solution for even a very small \( \gamma \). The best way to eliminate amplitude noise is to choose the shortest frictionless protocol. Nevertheless, some friction-like behaviors will occur.

Next, phase noise is considered. It occurs due to errors in the piecewise process used for controlling the scheduling of \( \omega \) in time. For such a procedure, random errors are expected in the duration of the time intervals. These errors are modeled by a Gaussian white noise. Mathematically, the process is equivalent to a dephasing process on the \( \text{adiabats} \) [97]. The dissipative operator \( \mathcal{L}_N \) has the form given by [42,95]:

\[
\mathcal{L}_N(\hat{\mathbf{A}}) = -\frac{\gamma_p}{\hbar^2} [\hat{\mathbf{H}}, [\hat{\mathbf{H}}, \hat{\mathbf{A}}]].
\]

In this case, the interaction picture for the phase noise \( \mathcal{U}_p \) becomes

\[
\frac{d}{dt} \mathcal{U}_p(t) = \mathcal{U}_a(-t) \mathcal{N}_p(t) \mathcal{U}_a(t) \mathcal{U}_p(t) = \mathcal{W}_p(t) \mathcal{U}_p(t),
\]

which at first order in \( \mu \) can be approximated as

\[
\mathcal{W}_p(t) \approx 2\gamma_p \omega_0 \times \begin{pmatrix} 0 & \mu s & \mu(1 - c) & 0 \\ \mu s & -(2 + \mu X) & 0 & 0 \\ \mu(c - 1) & 0 & -(2 + \mu X) & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.
\]

Again, using the Magnus expansion for one period of \( X \) leads to

\[
\mathcal{U}_p(X = 2\pi)_{B_1} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & e^{8\pi \gamma_p \omega_0} & (1 - e^{8\pi \gamma_p \omega_0}) \mu/2 & 0 \\ 0 & 0 & e^{8\pi \gamma_p \omega_0} & (1 - e^{8\pi \gamma_p \omega_0}) \mu/2 \\ 0 & 0 & 0 & 1 \end{pmatrix}.
\]

At first order in \( \mu \), this evolution operator maintains \( \delta_f(1) = 0 \), so the frictionless case holds. The second-order Magnus term leads to the noise correction

\[
\mathcal{U}_p(X = 2\pi)_{B_2} = \begin{pmatrix} \cosh \beta & -\sinh \beta & 0 & 0 \\ -\sinh \beta & \cosh \beta & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix},
\]

where \( \beta = \frac{16\omega_0^2 \pi^2}{4 + 3\gamma^2} \left( e^{2\pi l \mu} - 1 \right) \). In the limit of \( l \to \infty \), \( \beta = 4\gamma^2 (\omega_h^2 - \omega_c^2) \). The propagator \( \mathcal{U}_p(X = 2\pi)_{B_2} \) mixes energy and coherence, even at the limit \( \mu \to 0 \) and \( \tau_a \to \infty \), where one would expect frictionless solutions.

We can characterize the fraction of additional energy generated by a parameter \( \delta \). Asymptotically for amplitude noise: \( \delta_2 = e^\tau \beta/\mu > 0 \), and for phase noise \( \delta_f(1) = e^\tau \beta - 1 \approx 0 \), and the second-order correction \( \delta_f(2) = \cosh(\beta) - 1 > 0 \). Imperfect control on the \( \text{adiabats} \) will always lead to \( \delta_f > 0 \) and additional work invested in friction.
4.4. The Sudden Limit

The limit of vanishing time on the adiabats $\tau_a \ll 1/\omega_c$ leads to the sudden propagator; therefore, $\mu \to \pm \infty$. Such dynamics is termed sudden quench. The propagator $U_a$ has an explicit expression:

$$U_a = \begin{pmatrix}
\frac{1}{2}(1+\alpha) & \frac{1}{2}(1-\alpha) & 0 & 0 \\
\frac{1}{2}(1-\alpha) & \frac{1}{2}(1+\alpha) & 0 & 0 \\
0 & 0 & 1 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix}, \quad (58)$$

where $\alpha = (\omega_f/\omega_i)^2$ is related to the compression ratio $a_{ch} = C^2$ and $a_{hc} = C^{-2}$. The propagator mixes $\hat{H}$ and $\hat{L}$ when the compression ratio deviates from 1. As a result coherence is generated. The sudden propagator is an integral part of the frictionless bang-bang solutions [68,87]. Equation (58) can be employed as part of a bang-bang adiabat or as part of a complete sudden cycle.

4.5. Effects of an Exceptional Point on the Dynamics on the Adiabat

Exceptional points (EPs) are degeneracies of non-Hermitian dynamics [98,99] associated with the coalescence of two or more eigenstates. The studies of EPs have substantially grown due to the observation of (space-time reflection symmetry) PT symmetric Hamiltonians [100]. These Hamiltonians have a real spectrum, which becomes complex at the EP. The main effect of EPs (of any order) on the dynamics of PT-symmetric systems is the sudden transition from a real spectrum to a complex energy spectrum [101,102].

The adiabatic strokes are generated by a time-dependent Hamiltonian Equation (4). We therefore expect the propagator $U_a$ to be unitary, resulting in eigenvalues with the property $|u_j| = 1$. These properties are only true for a compact Hilbert space. We find surprising exceptions for the non-compact harmonic oscillator with an infinite number of energy levels. We can remove the trivial scaling $\omega(t)/\omega(0)$ in Equation (36) which originates from the diagonal part. The propagator can be written as $U_a = U_0 U_1$, where $U_0 = \omega(t)/\omega(0) I$ is a rescaling of the energy unit. The equation of motion for $U_1$ becomes:

$$\frac{d}{d\theta} U_1(\theta) = \begin{pmatrix}
0 & -\mu & 0 \\
-\mu & 0 & -2 \\
0 & 2 & 0
\end{pmatrix} U_1(\theta), \quad (59)$$

where the trivial propagation of the identity is emitted and the time is rescaled $\theta(t) = \int^t \omega'(t')dt'$. Diagonalising Equation (59) for constant $\mu$, we can identify three eigenvalues: $\lambda_1 = 0$ and $\lambda_{23} = \pm i \sqrt{4 - \mu^2}$. For $\mu \leq 2$, as expected, Equation (59) generates a unitary propagator. The three eigenvalues become degenerate when $\mu = 2$, and become real for $\mu \geq 2 \lambda_{23}$ [103]. This is possible because the generator Equation (59) is non-Hermitian. At the exceptional point, the matrix in Equation (59) has a single eigenvector corresponding to $\lambda_1 = \lambda_{23}$, which is self-orthogonal. To show this property, it is necessary to multiply the right and left eigenvectors of the non-symmetric matrix at the EP. Their product is equal to zero, showing that the eigenvector is self-orthogonal [104]. The propagator Equation (38) changes character at the EP; $\Omega = \sqrt{4 - \mu^2}$ changes from a real to an imaginary number. As a result, the dynamics at the EP changes from oscillatory to exponential [103].

This effect can also be observed in the classical parametric oscillator Equation (42). By changing the time variable $\frac{d}{d\tau} = \omega(t) \frac{d}{d\tau}$, and for constant $\mu$, the equation of motion becomes

$$\left(\frac{d^2}{d\tau^2} + \mu \frac{d}{d\tau} + 1\right) X(\tau) = 0, \quad (60)$$

which is the well-known equation of motion of a damped harmonic oscillator. Note that the original model (given by Equation (42)) does not involve dissipation, and a priori one would not expect the
appearance of an EP. The rescaling of the time coordinate allows us to identify an EP at $|\mu| = 2$, corresponding to the transition between an underdamped and an over-damped oscillator [105].

Exceptional points are also expected in the eigenvalues and eigenvectors of the total propagator $U_{cyc}$, which posses complex eigenvalues. Such points will indicate a drastic change in the cycle performance.

5. Closing the Cycle

Periodically combining the four propagators leads to the cycle propagator. Depending on the choice of parameters, we get either an engine cycle where heat flow is converted to power:

$$U_{cyc}^e = U_{ch} U_c U_{hc} U_h \quad \text{where} \quad \frac{\omega_c}{\omega_h} > \frac{T_c}{T_h},$$  \hspace{1cm} (61)

or a refrigerator cycle where power drives a heat current from the cold to the hot bath:

$$U_{cyc}^r = U_{ch} U_c U_{hc} U_h \quad \text{where} \quad \frac{\omega_c}{\omega_h} < \frac{T_c}{T_h}.$$ \hspace{1cm} (62)

In both cases, $\omega_h > \omega_c$. Frictionless cycles are either refrigerators or engines. Friction adds another possibility. When the internal friction dominates both, the engine cycle and the refrigeration cycle will operate in a dissipative mode, where power is dissipated to both the hot and cold baths. For an engine, this dissipative mode will occur when the internal temperature of the oscillator Equation (31) after the expansion adiabat (cf. Figure 3 point D) will exceed $T_h$ and in a refrigerator cycle when the internal temperature exceeds $T_c$ (cf. Figure 4 point C).

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3.png}
\caption{Typical engine cycle $S_E$ vs. $\omega$. Expansion adiabat $A \rightarrow B$. Cold isochore $B \rightarrow C$. Compression adiabat $C \rightarrow D$. Hot isochore $D \rightarrow A$. The hot and cold isotherms are indicated. The cycle parameters are $\omega_c = 0.5$, $T_c = 5$, $\omega_h = 2$, $T_h = 200$, $\tau_c = \tau_h = 2.1$, $|\mu| = 0.8$, $\Gamma = 1$.}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure4.png}
\caption{Typical frictionless refrigerator cycle $S_E$ vs. $\omega$. Expansion adiabat $D \rightarrow C$. Cold isochore $C \rightarrow B$. Compression adiabat $B \rightarrow A$. Hot isochore $A \rightarrow D$. The hot and cold isotherms are indicated. The cycle parameters are $\omega_c = 0.5$, $T_c = 1.5$, $\omega_h = 3$, $T_h = 3$, $\tau_c = \tau_h = 2.1$, $|\mu| = 0.5$, $\Gamma = 1$.}
\end{figure}
5.1. Limit Cycle

When a cycle is initiated, after a short transient time it settles to a steady-state operation mode. This periodic state is termed the limit cycle [14,106]. An engine cycle converges to a limit cycle when the internal variables of the working medium reach a periodic steady state. As a result, no energy or entropy is accumulated in the working medium. Figure 2 is an example of a periodic limit cycle. Subsequently, a balance is obtained between the external driving and dissipation. When the cycle time is reduced, friction causes additional heat to be accumulated in the working medium. The cycle adjusts by increasing the temperature gap between the working medium and the baths, leading to increased dissipation. Overdriving leads to a situation where heat is dissipated to both the hot and cold bath and power is only consumed. When this mechanism is not sufficient to stabilise the cycle, one can expect a breakdown of the concept of a limit cycle, resulting in catastrophic consequences [72].

The properties of a completely positive (CP) map can be used to prove the existence of a limit cycle. Lindblad [107] has proven that the conditional entropy decreases when applying a trace-preserving completely positive map $\Lambda$ to both the state $\hat{\rho}$ and the reference state $\hat{\rho}_{\text{ref}}$:

$$D(\Lambda \hat{\rho} || \Lambda \hat{\rho}_{\text{ref}}) \leq D(\hat{\rho} || \hat{\rho}_{\text{ref}}),$$

where $D(\hat{\rho} || \hat{\rho}') = \text{Tr}(\hat{\rho}(\log \hat{\rho} - \log \hat{\rho}'))$ is the conditional entropy distance. A CP map reduces the distinguishability between two states. This can be employed to prove the monotonic approach to steady-state, provided that the reference state $\hat{\rho}_{\text{ref}}$ is the only invariant of the CP map $\Lambda$ (i.e., $\Lambda \hat{\rho}_{\text{ref}} = \hat{\rho}_{\text{ref}}$) [108–110]. This reasoning can prove the monotonic approach to the limit cycle. The mapping imposed by the operation of a heat engine is a product of the individual evolution steps along the segments composing the cycle propagator. Each one of these evolution steps is a completely positive map, so the total evolution $U_{\text{cyc}}$ Equation (1) that represents one cycle of operation is also a CP map. If then a state $\hat{\rho}_{\text{lc}}$ is found that is a single invariant of $U_{\text{cyc}}$ (i.e., $U_{\text{cyc}} \hat{\rho}_{\text{lc}} = \hat{\rho}_{\text{lc}}$), then any initial state $\hat{\rho}_{\text{init}}$ will monotonically approach the limit cycle.

The largest eigenvalue of $U_{\text{cyc}}$ with a value of 1 is associated with the invariant limit cycle state $U_{\text{cyc}} \hat{\rho}_{\text{lc}} = \hat{\rho}_{\text{lc}}$, the fixed point of $U_{\text{cyc}}$. The other eigenvalues determine the rate of approach to the limit cycle.

In all cases studied of a reciprocating quantum heat engine, a single non-degenerate eigenvalue of 1 was the only case found. The theorems on trace preserving completely positive maps are all based on $\mathbb{C}^*$ algebra, which means that the dynamical algebra of the system is compact. Can the results be generalized to discrete non-compact cases such as the harmonic oscillator? In his study of the Brownian harmonic oscillator, Lindblad conjectured: “in the present case of a harmonic oscillator, the condition that $L$ is bounded cannot hold. We will assume this form for the generator with $\hat{H}$ and $L$ unbounded as the simplest way to construct an appropriate model” [45]. The master equation in Lindblad’s form Equation (9) is well established. Nevertheless, the non-compact character of the resulting map has not been challenged.

A nice demonstration is the study of Insinga et al. [72], which shows conditions where a limit cycle is not obtained. This study contains an extensive investigation of the limit cycles as a function of the parameters of the system [72].

5.2. Engine Operation and Performance

The engine’s cycle can operate in different modes, which are: adiabatic, frictionless, friction-dominated, and the sudden cycle. In addition, one has to differentiate between two limits: high temperature $k_B T \gg \hbar \omega$, where the unit of energy is $k_B T$, to low temperature $\hbar \omega \gg k_B T$, where the unit of energy is $\hbar \omega$.

In the adiabatic and frictionless cycles [111], the performance can be completely determined by the value of energy at the switching point between strokes.
5.2.1. Optimizing the Work per Cycle

The adiabatic limit with infinite time allocations on all segments maximises the work. No coherence is generated, and therefore the cycle can be described by the change in energy. On the expansion adiabat $E_B = \frac{\omega_e}{\omega_f} E_A$, and on the compression adiabat $E_D = \frac{\omega_e}{\omega_f} E_C$. As a result, when the cycle is closed, the heat transferred to the hot bath $Q_h = E^A - E^D$ and to the cold bath $Q_c = E_B - E_C$ are related: $\frac{Q_h}{Q_c} = \frac{\omega_e}{\omega_f}$.

The efficiency for an engine becomes the Otto efficiency:

$$\eta = \frac{W}{Q_h} = 1 - \frac{\omega_c}{\omega_f} \leq 1 - \frac{T_e}{T_h}.$$  \hspace{1cm} (63)

Choosing the compression ratio $C = \frac{\omega_c}{\omega_f} = \frac{T_c}{T_h}$ maximises the work and leads to Carnot efficiency $\eta_o = \eta_c$. Since for this limit the cycle time $\tau_{cyc}$ is infinite, the power $P = W/\tau_{cyc}$ of this cycle is obviously zero.

5.2.2. Optimizing the Performance of the Engine for Frictionless Conditions

Frictionless solutions allow finite time cycles with the same efficiency $\eta_o = 1 - \frac{\omega_c}{\omega_f}$ as the adiabatic case. A different viewpoint is to account as wasted work the average energy invested in achieving the frictionless solution, termed superadiabatic drive [112]:

$$\eta = \frac{W}{Q_h + \langle H_{ch} \rangle + \langle H_{hc} \rangle},$$  \hspace{1cm} (64)

where $\langle H_{ch} \rangle$ is the average additional energy during the adiabatic stroke. Using for example Equation (38), the average additional energy becomes $\langle H_{ch} \rangle = \frac{\omega_c}{\omega_f} E_C \frac{\mu^2}{4 - \mu^2}$, which vanishes as the non-adiabatic parameter $\mu \rightarrow 0$. This additional energy $\langle H_{ch} \rangle$ in the engine is the price for generating coherence. Coherence is exploited to cancel friction. This extra energy is not dissipated, and can therefore be viewed as a catalyst. For this reason, we do not accept the viewpoint of [112].

In our opinion, what should be added to the accounting is the additional energy generated by noise on the controls:

$$\eta = \frac{W}{Q_h + E_C \delta_{ch} + E_A \delta_{hc}} \leq \eta_o,$$  \hspace{1cm} (65)

where $\delta_{hc} = \delta_a + \delta_f$ for the power adiabat and $\delta_a$ and $\delta_f$ are generated by amplitude and phase noise on the controller (cf. Section 4.3). A similar relation is found for the compression adiabat.

Optimizing power requires a finite cycle time $\tau_{cyc}$. Optimisation is carried out with respect to the time allocations on each of the engine’s segments: $\tau_a$, $\tau_{hc}$, $\tau_c$, and $\tau_{ch}$. This sets the total cycle time $\tau_{cyc} = \tau_a + \tau_{hc} + \tau_c + \tau_{ch}$. The time allocated to the adiabats is constrained by the frictionless solutions $\tau_{hc}^*$ and $\tau_{ch}^*$. The resulting optimization is very close to the unconstrained optimum [72], especially in the interesting limit of low temperatures. The frictionless conditions are obtained either from Equation (45) or from other shortcuts to adiabaticity methods Equation (47). In the frictionless regime, the number operator is fixed at both ends of the adiabat. The main task is therefore to optimize the time allocated to thermalisation on the isochores. This heat transport is the source of entropy production.

The time allocations on the isochores determine the change in the number operator $N = \langle \hat{N} \rangle = \frac{1}{\hbar \omega_c} \langle \hat{H} \rangle$ (cf. Equation (33)): $N^B = e^{-\Gamma_{is} \hbar} \left( N^A - N^h_{eq} \right) + N^h_{eq}$ on the hot isochore, where $N^B$ is the number expectation value at the end of the hot isochore, $N^A$ at the beginning, and $N^h_{eq}$ is the equilibrium value point $E$. A similar expression exists for the cold isochore.

Work in the limit cycle becomes

$$W_q = E^C - E^B + E^A - E^D = \hbar (\omega_c - \omega_f) (N^B - N^D),$$  \hspace{1cm} (66)

where $\omega_c = \omega_f$ for the power adiabat and $\omega_c$ and $\omega_f$ are generated by amplitude and phase noise on the controller (cf. Section 4.3).
where the convention of the sign of the work for a working engine is negative, in correspondance with Callen [33], and we use the convention of Figure 3 to mark the population and energy at the corners of the cycle.

The heat transport from the hot bath becomes

$$Q_h = E^B - E^D = \hbar \omega_h (N^B - N^D).$$  \hspace{1cm} (67)

In the limit cycle for frictionless conditions, $N^B = N^A$, which leads to the relation

$$N^B = \frac{\left( e^{T_c \tau_c} - 1 \right)}{1 - e^{T_c \tau_c + T_h \tau_h}} (N^c_{eq} - N^h_{eq}) + N^h_{eq}. \hspace{1cm} (68)$$

In the periodic limit cycle, the number operator change $N^B - N^D$ is equal on the hot and cold isochores, leading to the work per cycle:

$$\mathcal{W}_q = \hbar (\omega_h - \omega_c) (N^h_{eq} - N^c_{eq}) \frac{(e^{T_c \tau_c} - 1)(e^{T_h \tau_h} - 1)}{1 - e^{T_c \tau_c + T_h \tau_h}} \equiv -G_W(T_c, \omega_c, T_h, \omega_h) F(x_c, x_h),$$

where the scaled time allocations are defined $x_c \equiv \frac{T_c}{x_c}$ and $x_h \equiv \frac{T_h}{x_h}$. The work $\mathcal{W}_q$ Equation (69) becomes a product of two functions: $G_W$, which is a function of the static constraints of the engine, and $F$, which describes the heat transport on the isochores. Explicitly, the function $G_W$ is

$$G_W(T_c, \omega_c, T_h, \omega_h) = \frac{\hbar}{2} (\omega_h - \omega_c) \left( \coth \left( \frac{\hbar \omega_h}{2k_B T_h} \right) - \coth \left( \frac{\hbar \omega_c}{2k_B T_c} \right) \right). \hspace{1cm} (70)$$

The function $F$ in Equation (69) is bounded $0 \leq F \leq 1$; therefore, for the engine to produce work, $G_W \geq 0$. The first term in (70) is positive. Therefore, $G_W \geq 0$ requires that $\frac{\omega_c}{\omega_h} \geq \frac{T_h}{T_c}$, or in terms of the compression ratio, $1 \leq C \leq \frac{T_h}{T_c}$. This is equivalent to the statement that the maximum efficiency of the Otto cycle is smaller than the Carnot efficiency $\eta_0 \leq \eta_c$.

In the high temperature limit when $\frac{\hbar \omega}{k_B T_c} \ll 1$, $G_W$ simplifies to

$$G_W = k_B T_c (1 - C) + k_B T_h (1 - C^{-1}). \hspace{1cm} (71)$$

In this case, the work $\mathcal{W}_q = -G_W F$ can be optimized with respect to the compression ratio $C = \frac{\omega_h}{\omega_c}$ for fixed bath temperatures. The optimum is found at $C = \sqrt{\frac{T_h}{T_c}}$. As a result, the efficiency at maximum power for high temperatures becomes

$$\eta_q = 1 - \frac{T_c}{T_h} \hspace{1cm} (72)$$

which is the well-known efficiency at maximum power of an endo-reversible engine [6,9,13,21,22,113]. Note that these results indicate greater validity to the Novikov–Curzon–Ahlbourn result from what their original derivation [22] indicates.

The function $F$ defined in (69) characterizes the heat transport to the working medium. As expected, $F$ maximizes when infinite time is allocated to the isochores. The optimal partitioning of the time allocation between the hot and cold isochores is obtained when:

$$\Gamma_h (\cosh(\Gamma_c \tau_c) - 1) = \Gamma_c (\cosh(\Gamma_h \tau_h) - 1). \hspace{1cm} (73)$$

If (and only if) $\Gamma_h = \Gamma_c$, the optimal time allocations on the isochores becomes $\tau_h = \tau_c$.

Optimising the total cycle power output $P$ is equivalent to optimizing $F/\tau_{cyc}$, since $G_W$ is determined by the engine’s external constraints. The total time allocation $\tau_{cyc} = \tau_{iso} + \tau_{adi}$ is partitioned...
to the time on the adiabats $\tau_{adi}$, which is limited by the adiabatic frictionless condition, and the time $\tau_{iso}$ allocated to the isochores.

Optimising the time allocation on the isochores subject to (73) leads to the optimal condition

$$\Gamma_c \tau_{cyc} (\cosh(\Gamma_h \tilde{\tau}_h) - 1) = \sinh(\Gamma_h \tilde{\tau}_h + \Gamma_c \tau_c) - \sinh(\Gamma_c \tau_c) - \sinh(\Gamma_h \tilde{\tau}_h).$$

(74)

When $\Gamma_h = \Gamma_c \equiv \Gamma$, this expression simplifies to:

$$2x + \Gamma_{adi} = 2 \sinh(x),$$

(75)

where $x = \Gamma_c \tau_c = \Gamma_h \tilde{\tau}_h$. For small $x$, Equation (75) can be solved, leading to the optimal time allocation on the isochores: $\tau_c = \tilde{\tau}_h \approx (\Gamma_{adi}/3)^{\frac{1}{2}} / \Gamma$. Considering the restriction due to frictionless condition [86], this time can be estimated to be: $\tau_c \approx \tilde{\tau}_h \approx \frac{1}{2} \left( \frac{\Gamma}{\sqrt{\omega_c \omega_h}} \right)^{\frac{1}{2}}$. When the heat transport rate $\Gamma$ is sufficiently large, the optimal power conditions lead to the bang-bang solution where vanishingly small time is allocated to all segments of the engine [14] and $\tau_{cyc} \approx 2 \tau_{adi}$.

The entropy production $\Delta S_u$ reflects the irreversible character of the engine. In frictionless conditions, the irreversibility is completely associated with the heat transport. $\Delta S_u$ can also be factorized to a product of two functions:

$$\Delta S_u = G_S(T_c, \omega_c, T_h, \omega_h) F(x_c, x_h),$$

(76)

where $F$ is identical to the $F$ function defined in (69). The function $G_S$ becomes:

$$G_S(T_c, \omega_c, T_h, \omega_h) = \frac{1}{2} \left( \frac{\hbar \omega_h}{k_B T_h} - \frac{\hbar \omega_c}{k_B T_c} \right) \left( \coth \left( \frac{\hbar \omega_c}{2k_B T_c} \right) - \coth \left( \frac{\hbar \omega_h}{2k_B T_h} \right) \right).$$

(77)

Due to the common $F(x_c, x_h)$ function, the entropy production has the same dependence on the time allocations $\tilde{\tau}_h$ and $\tau_c$ as the work $W$ [114]. As a consequence, maximizing the power will also maximize the entropy production rate $\Delta S_u / \tau_{cyc}$. Note that entropy production is always positive, even for cycles that produce no work, as their compression ratio $C$ is too large, which is a statement of the second law of thermodynamics.

The dependence of the $G_S$ function on the compression ratio can be simplified in the high temperature limit, leading to:

$$G_S = C \frac{T_c}{T_h} + C^{-1} \frac{T_h}{T_c} - 2,$$

(78)

which is a monotonic decreasing function in the range $1 \leq C \leq \frac{T_h}{T_c}$ that reaches a minimum at the Carnot boundary when $C = \frac{T_h}{T_c}$. When power is generated, the entropy production rate in the frictionless engine is linearly proportional to the power:

$$S_u = \left( \frac{\hbar \omega_h}{k_B T_h} - \frac{\hbar \omega_c}{k_B T_c} \right) \left( \frac{1}{\omega_h - \omega_c} \right) \mathcal{P}.$$

(79)

Frictionless harmonic cycles have been studied under the name of superadiabatic driving [111]. The frictionless adiabats are obtained using the methods of shortcut to adiabaticity [82] and the invariant Equation (47). An important extension applies shortcuts to adiabaticity to working mediums composed of interacting particles in a harmonic trap [76,115–117].

A variant of the Otto engine is an addition of projective energy measurements before and after each adiabat. This construction is added to measure the work output [118]. As a result, the working medium is always diagonal in the energy basis. In the frictionless case, the cycle is not altered by this projective measurement of energy.
5.2.3. The Engine in the Sudden Limit

The extreme case of the performance of an engine with zero time allocation on the adiabats is dominated by the frictional terms. These terms arise from the inability of the working medium to adiabatically follow the external change in potential. A closed-form expression for the sudden limit can be derived based on the adiabatic branch propagator $U_{hc}$ and $U_{ch}$ in Equation (58).

To understand the role of friction, we demand that the heat conductance terms $\Gamma$ are very large, thus eliminating the thermalisation time. In this limiting case, the work per cycle becomes:

$$W_s = \frac{(\omega_c - \omega_h)(\omega_c + \omega_h)}{4\omega_c\omega_h} \left( h\omega_c \coth\left( \frac{h\omega_h}{2k_BT_h} \right) - h\omega_h \coth\left( \frac{h\omega_c}{2k_BT_c} \right) \right).$$

(80)

The maximum produced work $-W_s$ can be optimised with respect to the compression ratio $C$.

At the high temperature limit:

$$W_s = \frac{1}{2}k_BT_h(C^2 - 1)(\frac{T_c}{T_h} - \frac{1}{C^2}).$$

(81)

For the frictionless optimal compression ratio $C = \sqrt{\frac{T_c}{T_h}}$, $W_s$ is zero. The optimal compression ratio for the sudden limit becomes: $C = \left( \frac{T_h}{T_c} \right)^{1/4}$, leading to the maximal work in the high temperature limit

$$W_s = -\frac{1}{2}k_BT_c \left( 1 - \sqrt{\frac{T_h}{T_c}} \right)^2.$$

(82)

The efficiency at the maximal work point becomes:

$$\eta_s = \frac{1 - \sqrt{\frac{T_h}{T_c}}}{2 + \sqrt{\frac{T_h}{T_c}}}. $$

(83)

Equation (83) leads to the following hierarchy of the engine’s maximum work efficiencies:

$$\eta_s \leq \eta_q \leq \eta_c.$$  

(84)

Equation (84) leads to the interpretation that when the engine is constrained by friction its efficiency is smaller than the endo-reversible efficiency, where the engine is constrained by heat transport that is smaller than the ideal Carnot efficiency. At the limit of $T_c \to 0$, we have $\eta_s = \frac{1}{2}$ and $\eta_q = \eta_c = 1$ [119].

An upper limit to the work invested in friction $W_f$ is obtained by subtracting the maximum work in the frictionless limit Equation (69) from the maximum work in the sudden limit Equation (80). In both these cases, infinite heat conductance is assumed, leading to $N^B = N_{eq}^{h}$ and $N^D = N_{eq}^{c}$. Then, the upper limit of work invested to counter friction becomes:

$$W_f = h\omega_h \frac{(C - 1)^2(1 + C + 2CN_{eq}^{c} + 2N_{eq}^{h})}{4C^2}.$$  

(85)

At high temperature, Equation (85) changes to:

$$W_f = \frac{1}{2}k_BT_h(C^2 - 1)^2(\frac{C}{C^2} + \frac{T_c}{T_h}).$$

(86)
The maximum produced work at the high temperature limit of the frictionless and sudden limits differ by the optimal compression ratio. For the frictionless case, \( C^* = \sqrt{\frac{T_h}{T_c}} \), and for the sudden case, \( C^* = \left( \frac{T_h}{T_c} \right)^{1/4} \).

The work against friction \( W_f \) (Equation (86)) is an increasing function of the temperature ratio. For the compression ratio that optimises the frictionless limit, the sudden work is zero. At this compression ratio, all the useful work is balanced by the work against friction \( W_f = W_r \). Beyond this limit, the engine transforms to a dissipator, generating entropy at both the hot and cold baths. This is in contrast to the frictionless limit, where the compression ratio \( C = \frac{T_h}{T_c} \) leads to zero power.

The complete sudden limit assumes short time dynamics on all segments including the \textit{isochores}. These cycles with vanishing cycle times approach the limit of a continuous engine. The short time on the \textit{isochores} means that coherence can survive. Friction can be partially avoided by exploiting this coherence, which—unlike the frictionless engine—is present in the four corners of the cycle. The condition for such cycles is that the time allocated is much smaller than the natural period set by the frequency \( \tau_s, T_h \ll 2\pi/\omega \) and by heat transfer \( \tau_s, T_h \ll 1/\Gamma \). The heat transport from the hot and cold baths in each stroke becomes very small. For simplicity, \( \Gamma_h T_h = \Gamma_c T_c \) is chosen to be balanced. Under these conditions, the cycle propagator becomes:

\[
U_{\text{cyc}} = \begin{pmatrix}
(1-g)^2 & 0 & 0 & \left[ g\left(1-g\right)\frac{1}{2}(1+C^2)\hbar\omega c N_{eq}^h + g\hbar\omega_c N_{eq}^c \right] \\
0 & (1-g)^2 & 0 & \left(1-g\right)\frac{1}{2}(1-C^2)\hbar\omega_c N_{eq}^c \\
0 & 0 & (1-g)^2 & 0 \\
0 & 0 & 0 & 1
\end{pmatrix},
\tag{87}
\]

where the degree of thermalisation is \( g = 1 - R \approx \Gamma_h T_h \). Observing Equation (87), it is clear that the limit cycle vector contains both \( \hat{H} \) and \( \hat{L} \).

The work output per cycle becomes:

\[
W_S = -\hbar\omega_c \frac{8}{2 - g} \frac{C^2 - 1}{2C^2} \left( N_{eq}^h - CN_{eq}^c \right).
\tag{88}
\]

Extractable work is obtained in the compression range of \( 1 < C < \frac{N_{eq}^h}{N_{eq}^c} \). The maximum work is obtained when \( C^* = \left( \frac{N_{eq}^h}{N_{eq}^c} \right)^{1/4} \). At high temperature, the work per cycle simplifies to:

\[
W_S \approx -k_B T_c \Gamma \frac{C^2}{2} - 1 \left( \frac{T_h}{T_c} - C^2 \right).
\tag{89}
\]

The work vanishes for the frictionless compression ratio \( C = \sqrt{\frac{T_h}{T_c}} \). The optimal compression ratio at the high temperature limit becomes: \( C^* = \left( \frac{T_h}{T_c} \right)^{1/4} \).

The entropy production becomes:

\[
\Delta S_u = \hbar\omega_c \frac{8}{2 - g} \frac{C^2 - 1}{2C^2} \left( N_{eq}^h \left( \frac{1 + C^2}{2CT_c} - \frac{C}{T_h} \right) + N_{eq}^c \left( \frac{1 + C^2}{2T_h} - \frac{1}{T_c} \right) \right).
\tag{90}
\]

Even for zero power (e.g., \( C = 1 \)), the entropy production is positive, reflecting a heat leak from the hot to cold bath.

The power of the engine for zero cycle time \( \tau_h \to 0 \) and \( \tau_c \to 0 \) is finite:

\[
P_S = -\frac{\hbar\omega_c \Gamma \frac{C^2}{2} - 1}{2C^2} \left( N_{eq}^h - CN_{eq}^c \right).
\tag{91}
\]
This means that we have reached the limit of a continuously operating engine. This observation is in accordance with the universal limit of small action on each segment [35,36]. When additional dephasing is added to Equation (87), no useful power is produced and the cycle operates in a dissipator mode.

The efficiency of the complete sudden engine becomes:

$$\eta_S = \frac{C^2 - 1}{2C^2} \cdot \frac{1 - C N_{eq}^{ch} / N_{eq}^{th}}{2 - (1 + C^2) \left[ \frac{N_{eq}^{ch}}{N_{eq}^{th}} \right]}.$$ (92)

The extreme sudden cycle is a prototype of a quantum phenomenonan engine that requires global coherence to operate. At any point in the cycle, the working medium state is non-diagonal in the energy representation.

5.2.4. Work Fluctuation in the Engine Cycle

Fluctuations are extremely important for a single realisation of a quantum harmonic engine. The work fluctuation can be calculated from the fluctuation of the energy at the four corners of the cycle [120,121]. The energy fluctuations for a generalised Gibbs state (Equation (20)) is related to the internal temperature (Equation (31)) $Var(E) = (k_B T_{int})^2$. For frictionless cycles, the variance of the work becomes:

$$Var(W) = (k_B T_{int}^h)^2 \left( 1 + 1/C^2 \right) + (k_B T_{int}^c)^2 \left( 1 + C^2 \right),$$ (93)

where $T_{int}^h$ and $T_{int}^c$ are the internal temperatures at the end of the hot and cold thermalisation. For the case of complete thermalisation when the oscillator reaches the temperature of the bath, the work variance is smallest for the Carnot compression ratio $C = T_h / T_c$. Generating coherence will increase the energy variance (cf. Equation (18)), and with it the work variance [120].

5.2.5. Quantum Fuels: Squeezed Thermal Bath

Quantum fuels represent a resource reservoir that is not in thermal equilibrium due to quantum coherence or quantum correlations. The issue is how to exploit the additional out-of-equilibrium properties of the bath. The basic idea of quantum fuels comes from the understanding that coherence can reduce the von Neumann entropy of the fuel. In principle, this entropy can be exploited to increase the efficiency of the engine without violating the second-law [122]. An example of such a fuel is supplied by a squeezed thermal bath [123–130]. Such a bath delivers a combination of heat and coherence. As a result, work can be extracted from a single heat bath without violating the laws of thermodynamics. An additional suggestion for a quantum fuel is a non-Markovian hot bath [131].

The model of this engine starts from a squeezed boson hot bath where $\hat{H}_B = \sum_k \hbar \Omega_k \hat{b}_k \hat{b}_k^\dagger$. This bath is coupled to the working medium by the interaction $\hat{H}_{SB} = \sum_k i \hbar \gamma_k (\hat{a} \hat{b}_k^\dagger - \hat{a}^\dagger \hat{b}_k)$. As a result, the master equation describing thermalisation (Equation (9)) is modified to [126,132]:

$$\mathcal{L}_{D}(\hat{\rho}) = k_1 (\hat{s}^\dagger \hat{\rho} \hat{s} - \frac{1}{2} \{ \hat{s} \hat{s}^\dagger, \hat{\rho} \}) + k_1 (\hat{s} \hat{\rho} \hat{s}^\dagger - \frac{1}{2} \{ \hat{s}^\dagger \hat{s}, \hat{\rho} \}),$$ (94)

where $\hat{s} = \hat{a} \text{cosh}(\gamma) + \hat{a}^\dagger \text{sinh}(\gamma) = \hat{S} \hat{a} \hat{S}^\dagger$. $\hat{S}$ is the squeezing operator (Equation (25)) and $\gamma$ the squeezing parameter.
Under squeezing, the equation of motion of the hot isochore thermalisation (Equation (33)) is modified to:

\[
\frac{d}{dt} \begin{pmatrix}
\hat{H} \\
\hat{L} \\
\hat{C} \\
\hat{I}
\end{pmatrix}(t) = \begin{pmatrix}
-\Gamma & 0 & 0 & \Gamma \langle \hat{H} \rangle_{sq} \\
0 & -\Gamma & -2\omega & 0 \\
0 & 2\omega & -\Gamma & \Gamma \langle \hat{C} \rangle_{sq} \\
0 & 0 & 0 & 0
\end{pmatrix} \begin{pmatrix}
\hat{H} \\
\hat{L} \\
\hat{C} \\
\hat{I}
\end{pmatrix}(t),
\]

(95)

where \( \Gamma = k_r - k_i \) is the heat conductance and \( k_r/k_i = e^{-\hbar\omega h/k_BT_h} \) obeys detailed balance. The difference from the normal thermalisation dynamics Equation (33) is in the equilibrium values:

\( \langle \hat{H} \rangle_{sq} = \cosh^2(\gamma) \langle \hat{H} \rangle_{eq} + \sinh^2(\gamma) \hbar \omega h k_i \),

where \( \langle \hat{H} \rangle_{eq} \) is the equilibrium value of the oscillator at temperature \( T_h \). In addition, the invariant state of Equation (95) contains coherence:

\( \langle \hat{C} \rangle_{sq} = -\sinh(2\gamma) \frac{\hbar}{k_BT_h} \). This coherence is accompanied by additional energy that is transferred to the system. The squeezed bath delivers extra energy to the working fluid as if the hot bath has a higher temperature, since \( \langle \hat{H} \rangle_{sq} \geq \langle \hat{H} \rangle_{eq} \). This temperature can be calculated from Equation (31). The thermalization to the squeezed bath generates mutual correlation between the system and bath [132].

The coherence transferred to the system \( \langle \hat{C} \rangle \leq \langle \hat{C} \rangle_{sq} \) can be cashed upon to increase the work of the cycle. This requires an adiabatic protocol which is similar to the frictionless case. In the frictionless case, the protocol of \( \omega(t) \) was chosen to cancel the coherence generated during the stroke and to reach a state diagonal in energy. This protocol can be modified to exploit the initial coherence and to reach a state diagonal in energy but with lower energy, thus producing more work. The coherence thus serves as a source of quantum availability, allowing more work to be extracted from the system [130,133,134]. For example, using the propagator on the adiabat Equation (38) based on \( \mu = \text{constant} \), the stroke period \( \tau_{st} \) can be increased from the frictionless value to add a rotation \( \cos(\Omega \omega(t)) \), which will null the coherence and reduce the final energy. Other frictionless solutions could be modified to reach the same effect.

5.3. Closing the Cycle: The Performance of the Refrigerator

A refrigerator or heat pump employs the working medium to shuttle heat from the cold to hot reservoir. A prerequisite for cooling is that the expansion adiabat should cause the temperature of the working medium to be lower than the cold bath. In addition, at the end of the compression adiabat, the temperature should be hotter than the hot bath (cf. Figure 4). To generate a refrigerator, we use the order of stroke propagators in Equation (62). The heat extracted from the cold bath becomes:

\( Q_c = E^B - E^C = \hbar \omega_c (N^B - N^C) \).

(96)

The interplay between efficiency and cooling power is the main theme in the performance analysis. The efficiency of a refrigerator is defined by the coefficient of performance (COP):

\[ \text{COP} = \frac{Q_c}{\dot{W}^\text{in}} = \frac{\omega_c}{\omega_h - \omega_c} \leq \frac{T_c}{T_h - T_c}. \]

(97)

The cooling power \( R_c \) is defined as:

\[ R_c = \frac{Q_c}{\tau_{cyc}}. \]

(98)

Optimising the performance of the refrigerator can be carried out by a similar analysis to the one employed for the heat engine. Insight into the ideal performance can be gained by examining the expansion adiabat. The initial excitation should be minimized, requiring the hot bath to cool the working medium to its ground state. This is possible if \( \hbar \omega_h \gg k_B T_h \). Next, the expansion should be as adiabatic as possible so that at the end the working medium is still as close as possible to its ground
state $E_C \approx \frac{\hbar \omega}{2}$. The frictionless solutions found in Section 4.2 can be employed to achieve this task in minimum time.

5.3.1. Frictionless Refrigerator

The adiabatic refrigerator is obtained in the limit of infinite time $\mu \to 0$, leading to constant population $N$ and $S_E$. Then, $E_C = \frac{\hbar \omega}{2} E^D$. At this limit, since $\tau \to \infty$, the cooling rate vanishes $\mathcal{R}_c = 0$. The Carnot efficiency can be obtained when $C = \frac{T_h}{T_c}$.

Frictionless solutions require that the state $\hat{\rho}$ is diagonal in energy in the beginning and at the end of the adiabat. The analytic propagator on the expansion adiabat (Equation (38)) describes the expansion adiabat:

$$E^C = \frac{1}{C} \frac{1}{\Omega^2} \left( 4 - \mu^2 \right) E^D ,$$

where $c = \cos(\Omega \theta_c)$ and $\theta_c = -\frac{1}{\mu} \log(C)$.

Other faster frictionless solutions can be obtained using the protocols of Section 4.2, such as the superadiabatic protocol or by applying optimal control theory [86]. Both cases lead to the scaling of the adiabatic expansion time as $\tau_{hc} \propto \frac{1}{\sqrt{\omega_c \omega_h}}$.

Once the time allocation on the adiabats is set, the time allocation on the isochores is optimised for the thermalisation using the method of [24], and the optimal cooling power becomes:

$$\mathcal{R}_c^* = -\hbar \omega_c \Gamma \left( N_{c}^{eq} - N_{h}^{eq} \right) ,$$

where $z = \Gamma_h \tau_h = \Gamma_c \tau_c$. The optimal $z$ is determined by the solution of the equation $2z + \Gamma(\tau_{hc} + \tau_{ch}) = 2 \sinh(z)$.

5.3.2. The Sudden Refrigerator

Short adiabats generally lead to the excitation of the oscillator and result in friction (cf. Section 4.2). Nevertheless, a refrigerator can still operate at the limit of vanishing cycle time. In a similar fashion to the sudden engine, coherence can be exploited and leads to a finite cooling power when $\tau_{cyc} \to 0$.

The cooling power for the sudden limit becomes:

$$\mathcal{R}_c = -\hbar \omega_c \Gamma \left( N_{c}^{eq} - \frac{1}{2C} (1 + C^2) N_{h}^{eq} \right) .$$

Note that the cooling rate in this sudden-limit becomes zero at a sufficiently low cold bath temperature so that $N_{c}^{eq} = (1 + C^2) N_{h}^{eq}$. This formula loses its meaning and should not be used below this temperature.

5.3.3. The Quest to Reach Absolute Zero

The quantum harmonic refrigerator can serve as a primary model to explore cooling at very low temperatures. A necessary condition is that the internal temperature of the oscillator (Equation (31)) should be lower than the cold bath. $T_{int} < T_c$ when $T_c \to 0$. This condition imposes very high compression ratios $C \gg \frac{1}{T_c}$ so that at point $D$ (cf. Figure 4), at the end of the hot thermalization, the oscillator is very close to the ground state.
An important feature of the model is that the cooling power vanishes as $T_c$ approaches zero. Qualitatively, $R_c \to 0$ means that the adiabatic expansion from point $D \to C$ for high compression ratios requires a significant amount of time. Another issue is the rate of cold thermalisation $C \to B$ and its scaling with $T_c$ when the oscillator extracts heat from the cold bath. These issues can be made quantitative by exploring the scaling exponent $\alpha$ of the optimal cooling power with the cold bath temperature $T_c$:

$$R_c = \frac{Q_c}{\tau_{cyc}} \propto T_c^{1+\alpha}.$$  (102)

The vanishing of the cooling power $R_c$ as $T_c \to 0$ is related to a dynamical version of the third-law of thermodynamics [25,135].

Walther Nernst formulated two independent formulations of the third-law of thermodynamics [136–138]. The first is a purely static (equilibrium) one, also known as the “Nernst heat theorem”, phrased:

• The entropy of any pure substance in thermodynamic equilibrium approaches zero as the temperature approaches zero.

The second formulation is dynamical, known as the unattainability principle [135,139–142]:

• It is impossible by any procedure—no matter how idealised—to reduce any assembly to absolute zero temperature in a finite number of operations [138].

The second law of thermodynamics already imposes a restriction on $\alpha$ [25,135,143]. In steady-state, the entropy production rate is positive. Since the process is cyclic, it takes place only in the baths: $\sigma = \dot{S}_c + \dot{S}_h \geq 0$. When the cold bath approaches the absolute zero temperature, it is necessary to eliminate the entropy production divergence at the cold side because $\dot{S}_c = \frac{R_c}{T_c}$. Therefore, the entropy production at the cold bath when $T_c \to 0$ scales as:

$$\dot{S}_c \sim -T_c^\alpha, \quad \alpha \geq 0.$$  (103)

For the case when $\alpha = 0$, the fulfillment of the second-law depends on the entropy production of the other baths, which should compensate for the negative entropy production of the cold bath. The first formulation of the third-law slightly modifies this restriction. Instead of $\alpha \geq 0$, the third-law imposes $\alpha > 0$, guaranteeing that the entropy production at the cold bath is zero at absolute zero: $\dot{S}_c = 0$. This requirement leads to the scaling condition of the heat current $R_c \sim T_c^{\alpha+1}$, $\alpha > 0$.

The second formulation of the third-law is a dynamical one, known as the unattainability principle: no refrigerator can cool a system to absolute zero temperature at finite time. This formulation is more restrictive, imposing limitations on the system bath interaction and the cold bath properties when $T_c \to 0$ [135]. The rate of temperature decrease of the cooling process should vanish according to the characteristic exponent $\zeta$:

$$\frac{dT_c(t)}{dt} \sim -T_c^\zeta, \quad T_c \to 0.$$  (104)

In order to evaluate Equation (104), the heat current can be related to the temperature change:

$$J_c(T_c(t)) = -c_V(T_c(t)) \frac{dT_c(t)}{dt}.$$  (105)

This formulation takes into account the heat capacity $c_V(T_c)$ of the cold bath. $c_V(T_c)$ is determined by the behaviour of the degrees of freedom of the cold bath at low temperature. Therefore, the scaling exponents can be related $\zeta = 1 + \alpha - \eta$, where $c_V \sim T_c^\eta$ when $T_c \to 0$.

The harmonic quantum refrigerator is a primary example to explore the emergence of quantum dynamical restrictions that result in cooling power consistent with the third-law of thermodynamics. Analysis of the adiabatic expansion will lead to insight on the cooling rate $R_c$ and the exponent $\alpha$. 
The frictionless solutions lead to an upper bound on the optimal cooling rate (Equation (98)). For the limit $T_c \to 0$, $\Gamma_\tau$ is large; therefore, $z$ is large, leading to:

\[ R^*_c \approx \frac{\Gamma (\tau_{hc} + \tau_{ch})}{(1 + \Gamma_{hc})^2} \Gamma \omega_c (N_{c}^{eq} - N_{\mu}^{eq}) \].

(106)

At high compression ratio, $N_{\mu}^{eq} \to 0$, and in addition $\omega_c \ll \Gamma$ one obtains:

\[ R^*_c \approx \frac{1}{\tau_{hc}} \Gamma \omega_c N_{c}^{eq} \].

(107)

Optimizing $R_c$ with respect to $\omega_c$ leads to a linear relation between $\omega_c$ and $T_c$, $\omega_c = k_B T_c$; therefore:

\[ R_c \leq A \omega_c^\nu N_{c}^{eq} \]

(108)

where $A$ is a constant and the exponent $\nu$ is either $\nu = 2$ for the $\mu = const$ solution or $\nu = \frac{3}{2}$ for the optimal control solution. Therefore:

\[ R^*_c \approx h \omega_c^2 N_{c}^{eq} \]

(109)

for the $\mu = const$ frictionless solution, and

\[ R^*_c \approx \frac{1}{2} h \omega_c^3 \sqrt{\omega_h N_{c}^{eq}} \]

(110)

for the optimal control frictionless solution. Due to the linear relation between $\omega_c$ and $T_c$, Equations (109) and (110) determine the exponent $\alpha$, where $\alpha = 1$ for the frictionless scheduling with constant $\mu$, and $\alpha = \frac{1}{2}$ for the optimal control frictionless scheduling. In all cases, the dynamical version of Nernst’s heat law is observed based only on the adiabatic expansion.

If one is forced to spend less time on the adiabat than the minimal time required for a shortcut solution, the oscillator cannot reach arbitrarily low energies or temperatures at the end of the expansion [92,93]. At the limit, one approaches the sudden adiabat limit. In this case, the refrigerator cannot cool below a minimal ($T^*_c > 0$) temperature, and the refrigerator thus satisfies the unattainability principle trivially.

The unattainability principle is related to the scaling of the heat transport $\Gamma_c$ with $T_c$. This issue has been explored in [25,135], and is related to the scaling of the heat conductivity with temperature. The arguments of [135] are applicable to the quantum harmonic refrigerator.

6. Overview

Learning from example has been one of the major sources of insight in the study of thermodynamics. A good example can bridge the gap between concrete and abstract theory. The harmonic oscillator quantum Otto cycle serves as a primary example of a quantum thermal device inspiring experimental realisation [2,144]. On the one hand, the model is very close to actual physical realisations in many scenarios [19]. On the other hand, many features of the model can be obtained as closed-form analytic solutions.

Many of the features obtained for the quantum harmonic Otto engine have been observed in stochastic thermodynamics [121,144–147]. The analytic properties of the harmonic oscillator—in particular, the Gaussian form of the state—have motivated studies of classical stochastic models of harmonic heat engines [144,148,149]. When comparing the two theories, the results seem identical in many cases. Observing the Heisenberg equations of motion for the thermodynamical variables, $\hbar$ does not appear. Planck’s constant in the commutators is cancelled by the inverse Planck constant in the equation of motion. This raises the issue of what is quantum in the quantum harmonic oscillator, or a related issue—what is quantum in quantum thermodynamics [31]?
In this review, we emphasized the power of the generalized Gibbs state in allowing a concise description of an out-of-equilibrium situation of non-commuting operators. Using properties of Lie algebra of operators, we could obtain a dynamical description of the state based on only three variables: \( \hat{H}, \hat{L}, \) and \( \hat{C} \). In the spirit of open quantum systems, we could describe the cycle propagator as a catenation of stroke propagators. All these propagators were cast in the framework of the operator algebra, showing the power of Heisenberg representation. The quantum variables were chosen to have direct thermodynamical relevance as energy and coherence. In this review, we emphasized the connections between the algebraic approach and other popular methods that have been employed to obtain insight on the harmonic engine.

This formalism allows the cycles to be classified according to the role of coherence. If the coherence vanishes at the points where the strokes meet, frictionless cycles are obtained. Such cycles require special scheduling of \( \omega(t) \) so that the coherence generated at the beginning of the stroke can be cashed upon at the end. We reviewed the different approaches to obtain such scheduling and the minimum time that such moves can be generated. This period is related to quantum speed limits [150–154], which are in turn related to the energy resources available to the system. We chose the geometric mean \( \tau_a \propto \frac{1}{\sqrt{\omega_h \omega_c}} \) to represent the minimum time allocation. Faster scheduling requires unreasonable constraints on the stored energy in the oscillator during the stroke. We also assume that this extra energy required to achieve the fast control is not dissipated and can be accounted for as a catalyst.

For these frictionless solutions on the adiabats, the optimal time allocation for thermalisation is finite, leading to incomplete thermalisation. This allows the minimum cycle time for frictionless cycles to be estimated. Avoiding friction completely is an ideal that practically cannot be obtained. Using a simple noise model, we show that some friction will always be present.

The model demonstrates the fundamental tradeoff between efficiency and power. The frictionless solutions are a demonstration that quantum coherence which is related to friction can be cashed upon, using interference to cancel this friction. As a result, the maximum efficiency of the engine can be obtained in finite cycle time. Nevertheless, the Otto efficiency is smaller than the reversible Carnot efficiency \( \eta_o = 1 - \frac{\omega_c}{\omega_h} \leq \eta_c = 1 - \frac{T_c}{T_h} \), and operating at the Carnot efficiency will lead to zero power. The entropy production can be associated with the heat transport, and for this case the entropy production is linearly related to the power. Maximum power also implies maximum entropy production. This finding is consistent with the study of Shiraishi et al. [155]. Any finite power cycle requires out-of-equilibrium setups that lead to dissipation. In the sudden limit, there is no reversible choice. Even at zero power the entropy production is positive. This could be the cost of maintaining coherence.

Beyond a minimum time allocation on the adiabats \( \tau_a \), friction cannot be avoided. The transition point is the exceptional point of the non-hermitian degeneracy on the adiabatic propagator [103]. These short time cycles are in the realm of the sudden cycles. The sudden cycles are an example of an engine or refrigerator with no classical analogue. Power production requires coherence. A sudden model without coherence operates as a dissipator generating entropy on both the hot and cold baths. The sudden cycle is composed of non-commuting propagators with small action. Such cycles are universal and have a common continuous limit [35,36]. In the continuous limit, friction and heat leaks cannot be avoided [39,156].

An obvious direction to look for quantum effects is to go to low temperatures where the unit of energy changes from \( k_B T \) to \( \hbar \omega \). The adiabatic expansion is the bottleneck for cooling to extremely low temperatures. The zero point energy plays an important role. We can approach the ground state on the hot side by increasing the frequency, leading to the minimal initial energy \( E_A = \frac{1}{2} \hbar \omega_0 \). This is a large amount of energy compared to the cold side, which has to be eliminated adiabatically or by using frictionless protocols. Any small error in these protocols will null the cooling.

The Otto quantum refrigerator is a good example for gaining insight into the limits of cooling when operating at extremely low temperatures. Such refrigerators are an integral part of any quantum technology. In the reciprocating Otto cycle, the cooling power is restricted either by the adiabatic
expansion or by vanishing of the heat transport when $T_c \to 0$ [135]. The adiabatic expansion time is an intrinsic property of the working medium. For optimal frictionless solutions, it scales as $\tau_{hc} = O(T_c^2)$, which gives a maximum rate of entropy production $\sigma = O(T_c^2)$, thus vanishing when $T_c \to 0$. This is a demonstration of a dynamical version of the Nernst heat law [25,68,135].

The quantum harmonic Otto cycle has been a template for many models of quantum heat devices due to its analytic properties—for example, Otto cycles with interacting particles [115,116] or operating with many modes [157]. The protocols developed for the harmonic case are generalised to eliminate friction in many-body dynamics.

The quantum harmonic Otto cycle has been a source of inspiration for theory and experiment. The model incorporates generic features of irreversible operation which includes friction and heat transport. The system can bridge the conceptual gap between a single microscopic device to a macroscopic heat engine.

**Acknowledgments:** We thank Amikam Levy, Tova Feldmann, Raam Uzdin and Erik Torrontegui for sharing their ideas and insights. We also thank Peter Salamon, Gonzales Muga, Robert Alicki for fruitful discussions. We acknowledge funding by the Israel Science Foundation, and COST Action MP1209 *Thermodynamics in the quantum regime*.

**Conflicts of Interest:** The authors declare no conflict of interest.

**Abbreviations**

The following abbreviations are used in this manuscript:

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>EP</td>
<td>Exceptional Point</td>
</tr>
<tr>
<td>L-GKS</td>
<td>Lindblad, Goirini Kossakowski Sudarshan</td>
</tr>
<tr>
<td>COP</td>
<td>Coefficient of Performance</td>
</tr>
<tr>
<td>CP</td>
<td>Completely Positive</td>
</tr>
</tbody>
</table>

**References**


(c) 2017 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).