Nonlinear thermoelectricity in point-contacts at pitch-off: a catastrophe aids cooling

Robert S. Whitney

1 Laboratoire de Physique et Modélisation des Milieux Condensés (UMR 5493), Université Grenoble I, Maison des Magistères, B.P. 166, 38042 Grenoble, France.

Dated: September 18, 2012

We consider refrigeration and heat engine circuits based on the nonlinear thermoelectric response of point-contacts at pinch-off, allowing for electrostatic interaction effects. We show that a refrigerator can cool to much lower temperatures than predicted by the thermoelectric figure of merit $ZT$ (which is based on linear-response arguments). The lowest achievable temperature has a discontinuity, called a fold catastrophe in mathematics, at a critical driving current $I = I_c$. For $I > I_c$, one can in principle cool to absolute zero, when for $I < I_c$ the lowest temperature is about half the ambient temperature. Chargeless particles (typically phonons and photons) stop cooling at a temperature above absolute zero, and above a certain threshold also modify the discontinuity. More generally, we show that any system with a high figure of merit should have its nonlinear response analyzed; since its figure of merit gives little indication of its potential as a refrigerator or heat-engine.

**Introduction.** Nanostructures often have thermoelectric responses, with electrical-currents causing heat-currents, and vice-versa [14]. There have recently been a number of proposals for systems with large thermoelectric responses [5–12] which could have engineering applications for efficient thermoelectric power-generation and refrigeration. In particular, it is hoped that they could efficiently cool electrons well below the temperature of standard cryostats [13], which are increasingly inefficient at sub-Kelvin temperatures.

Here we calculate the fully nonlinear thermoelectric response of a point-contact at pinch off. This system’s linear (and nearly linear) thermoelectric response is well-studied experimentally [14,15] and theoretically [14,16–18]. However a refrigerator which cools to significantly below the temperature of its environment will typically be far outside the linear regime [13,19]; since the temperature difference is not small compared with the average temperature. Here, we apply the nonlinear version of the Landauer-Büttiker scattering theory [20,24] to thermoelectric heat-transport [25] through point-contacts. This shows that the dimensionless figure of merit, $ZT$, ceases to be an accurate measure of the thermoelectric response outside the linear regime. Electricity generation is worse than linear-response theory indicates, but refrigeration is better (achieving much lower temperatures than linear-response theory predicts). Indeed, the lowest temperature of the refrigerator is a discontinuous function of the electrical current. This discontinuity — a fold catastrophe in mathematical language — occurs at a critical current $I_c$, and is beneficial to refrigeration. For currents $I < I_c$ the refrigerator cannot cool below a finite temperature (of order half the ambient temperature for $I \to I_c$), while for $I > I_c$ it passes the catastrophe point and can in principle cool to absolute zero (see Fig. 1).

In practice, a thermoelectric device’s quality is reduced by heat flow carried by chargeless particles; usually phonons and photons. In the nonlinear regime, we show the weak phonon or photon effects do not significantly affect the catastrophe, $I_c$, but do stop the cooling at a temperature above absolute zero. However at a critical value of the thermal transport due to chargeless particles, the catastrophe becomes a discontinuity in the derivative of the curve lowest temperature versus $I$.

**Nearly-linear analysis and its breakdown.** The usual “nearly-linear” analysis [1] takes linear response theory plus a Joule heating term, and enables one to quantify devices in terms of their dimensionless figure of merit $ZT = GS^2T/(\Theta_{el} + \Theta_{ph})$, where $T$ is the device temperature, $S$ is its Seebeck coefficient, $G$ and $\Theta_{el}$ are electrical and thermal conductances of electrons, while $\Theta_{ph}$ is the thermal conductivity of chargeless excitations, principally phonons and photons. This nearly-linear analysis predicts electric power generation (when the island is heated) with an efficiency

$$
\eta = \frac{\sqrt{ZT+1} - 1}{\sqrt{ZT+1} + 1} \left( 1 - \frac{T_0}{T_{isl}} \right),
$$

Figure 1: Heat-current $J(T_{isl}, I)$ through a point-contact when driven with a current $I$, for negligible phonon or photon heating. Blue indicates cooling of the island in Fig. 2, while red indicates heating. The solid curve is the steady-state ($J = 0$), with the catastrophe at $I_c$. The straight line is the maximum current, $I_{max}$, corresponding to infinite bias.
where \( T_{\text{isl}} \) and \( T_0 \) are the island and environment temperatures. Typically, \( ZT \) is taken at the temperature \( \frac{1}{2}(T_0 + T_{\text{isl}}) \). Carnot efficiency corresponds to \( ZT \to \infty \). For refrigeration, it predicts that the lowest achievable temperature, \( T_{\text{min}} \), is given by

\[
T_{\text{min}}/T_0 = 1 - \frac{1}{2}ZT.
\]

The origin of Eq. (2) is easily seen, and enlightens us about the breakdown of the nearly-linear analysis. One starts with the heat-current out of the island [11]

\[
J(T_{\text{isl}}, I) \approx \Pi I - \Theta_+ (T_0 - T_{\text{isl}}) - \frac{1}{2}R_+ I^2,
\]

for Peltier coefficient difference \( \Pi = (\Pi_2 - \Pi_1) \), sum of thermal conductances \( \Theta_+ = \Theta_1 + \Theta_2 + \Theta_{\text{ph}} \) and sum of electrical resistances \( R_+ = G_1^{-1} + G_2^{-1} \). The first two terms are linear-response terms, while the last is the Joule heating. The steady-state curve, \( J = 0 \), shows that the lowest \( T_{\text{isl}} \) is a quadratic function of \( I \). The parabola’s minimum is \( T_{\text{min}} \) in Eq. (2) — using the Onsager relation \( \Pi = S/T \) to note \( ZT = G_+ I^2/(\Theta_+ T) \).

For our model of a point-contact at pinch-off (detailed below, linear-response scattering theory [25][29][34] gives \( G_1 = (e^2/h)(1/2), \Pi_1 = -(k_B T_0/e)2\ln(2), \) and \( \Theta_1 = (k_B T_0/h)(\pi^2/6 - 2[\ln(2)]^2) \) [35]. This gives \( ZT \approx 1.4 \), which would imply that \( \eta = 0.22(1 - T_0/T_{\text{isl}}), \quad T_{\text{min}} = 0.3 T_0. \)

However, this analysis fails when nonlinear effects are too strong to start from Eq. (3). Typically this occurs when the leading nonlinear Peltier term [17], of the form \( \Pi I^2 \), is larger than the Joule heating term \( \frac{1}{2}R_+ I^2 \). For the point-contact this is so, since \( \Pi = (h/e^2)3\ln[2] \). Then including this term in Eq. (3) changes the sign of the prefactor on \( I^2 \); which means we must look beyond \( I^2 \) to find the steady-state curve’s minimum. Our fully nonlinear analysis below, shows that expanding about equilibrium to any order [17] will not give this minimum.

**Nonlinear analysis.** The thermoelectricity literature [13] discusses \( J(T_{\text{isl}}, I) \) — as in Eq. (3) above — rather than \( J(T_{\text{isl}}, I) \) for voltage drop, \( V \). This is because different thermoelectric devices are arranged in series electrically (see Fig. 2a,b), so \( I \) is the same in all of them (unlike voltage drops). Thus it is easier to get response of a series of elements from each element’s \( J(T_{\text{isl}}, I) \) than from each element’s \( J(T_{\text{isl}}, V) \). For complicated non-linear responses, the former is straight-forward while the latter is extremely difficult; thus we consider \( J(T_{\text{isl}}, I) \).

We take the island to be classical; i.e. big enough for particles entering it to thermalize to a Fermi distribution at temperature \( T_{\text{isl}} \) before escaping. See Refs. [20][28] for cases where there is a quantum dot in place of the classical island. In our case, each point-contact can be treated by a separate Landauer-Büttiker scattering matrix analysis [25], see also Refs. [29][34]. We generalize these heat currents to the nonlinear regime [36], including electrostatic (Hartree-like) interaction effects in a self-consistent and gauge-invariant manner; as Refs. [21][24] did for charge-current, see also [37][38]. To go beyond the voltage-squared contributions to transport (which Ref. [21] treated in detail), we use a simple model of interactions, which is none the less gauge-invariant and self-consistent. The heat-current into lead \( i \) of a given nanostructure is

\[
J_i = -\int_{-\infty}^{\infty} \frac{d\epsilon}{\hbar} \sum_j (\epsilon - qV_i) A_{ij} \left( (\epsilon - qV_k) \right) f_j(\epsilon),
\]

where \( f_j(\epsilon) = (1 + \exp[(\epsilon - qV_j)/(k_B T_j)])^{-1} \) is the Fermi function, and \( q \) is the charge of the carriers; electrons with \( q = -e \) in point-contact 1 and holes with \( q = e \) in point-contact 2. The energy \( \epsilon \) and all voltages \( V_k \) are measured from the same external reference. The formula for the charge-current \( I_i \) into lead \( i \) is the same but with \( q \) in...
place of \((\epsilon - qV_i)\) as the first factor in the integrand. The transmission function \(A_{ij}\) of a particle through the nanostructure from lead \(j\) to lead \(i\) is 
\[
A_{ij}((\epsilon - qV_{k})) = \text{Tr} \left[ 1_s \delta_{ij} - S^\dagger_{ij}((\epsilon - qV_{k})) S_{ij}((\epsilon - qV_{k})) \right],
\]
where \(S_{ij}\) is the scattering matrix from lead \(j\) to lead \(i\), and the trace is over all modes of those leads. Here \(S_{ij}\) must be found self-consistently: it depends on the charge distribution in the nanostructure, which in turn depends on \(S_{ij}\).

Point-contact 1 is a two-lead nanostructure with electron charge-carriers \((q = -e)\). Having established the gauge-invariance, we are free to measure all energies \(\epsilon\) and all voltages \(V_{k}\) (including \(V_g\)) from the chemical potential of the island (the point-contact’s M lead). When \(V_L\) is non-zero, we assume that a proportion \((1 - a)\) of this bias is screened by the electrostatic gates a long way from the narrowest-point of the point-contact, while the rest is screened self-consistently by the electron-gas (Fig. 2c) close to the point contact. Then the screened point-contact induces a potential barrier of height, \(E_{pc}\) (measured from the island’s chemical potential), typically obeying \(E_{pc} - E_g = g_{scr}(aqV_L)\), where \(g_{scr}\) is due to screening. Here \(E_g\) can be tuned at will, since it is \(eV_g\) minus a geometry-dependent constant. Assuming a long enough point-contact that there is negligible tunnelling, one has \(A_{LM}(\epsilon - E_{pc} > 0) = -1\) (perfect transmission) and \(A_{LM}(\epsilon - E_{pc} < 0) = 0\) (no transmission) \[40\]. As an example, the Supplementary Material gives a simple model of screening for which we derive \(g_{scr}(aqV_L)\) self-consistently. However, in what follows we allow the nature of screening (both \(a\) and the form of \(g_{scr}(aqV_L)\)) to be completely arbitrary. Gauge-invariance is satisfied, as seen by noting that \(\epsilon - E_{pc}\) is a function of \((\epsilon - qV_{k})\), since \(V_{L-g}\) are biases relative to the M lead.

To operate at pinch-off, we tune \(V_g\) so that \(E_{pc} = 0\) for any given \(V_L\). If the electrostatic gates dominate screening \((a \rightarrow 0)\), then \(E_{pc}\) is independent of \(V_L\), making it easy to tune to pinch-off. Otherwise the point-contact should be calibrated prior to use; finding the pinch-off point (the \(V_g\) at which current starts to flow), as a function of \(V_L\). Then the charge and heat-currents from point-contact 1 into the metal island are

\[
I(T_{isl}, V_L) = \frac{ek_B}{\hbar} T_{isl} \ln(2) - T_0 \ln \left( 1 + e^{-\epsilon V_L/k_BT_0} \right),
\]

\[
J_1(T_{isl}, V_L) = -\frac{k^2_B T_0^2}{\hbar} T_{isl}^2 \frac{\pi^2}{12} + L_{k2} \left( -e^{-\epsilon V_L/k_BT_0} \right),
\]

for a dilogarithm function \(L_{k2}(z) = \int_0^\infty t dt/(e^t/z - 1)\). From Eqs. \[47\], we get

\[
J_1(T_{isl}, I) = -\frac{k^2_B T_0^2}{\hbar} \frac{\pi^2}{12T_0^2} + L_{k2} \left( 1 - \exp \left[ \beta(T_{isl}, I) \right] \right),
\]

where we define \(J = h \left[ I_{max}(T_{isl}) - I \right] / (ek_BT_0) \) and note that \(I \leq I_{max}(T_{isl}) = ek_BT_{isl} \ln(2)/\hbar\). This function is given by the color plot in Fig. 1 Fig. 1 for point-contact 2 (in which the carriers are holes not electrons) we take \(-e \leftrightarrow e\), then \(J_2(T_{isl}, I) = J_1(T_{isl}, I) \) since \(I_2 = -I\).

For \(I \ll 1\), we can use \(L_{k2}(z) = z + O(z^2)\) to write

\[
J_1 = \left( k^2_B T_0^2 / h \right) \left[ 3 - (\pi^2/12)(T_{isl}/T_0)^2 + O(1) \right],
\]

so \(J_1\) as a quadratic in temperature and linear in current; the reverse of the nearly-linear theory in Eq. \[3\]. This approximation captures the features of the exact result plotted in Fig. 1 except the top-left corner. This corner is the linear-response regime (small \((T_0 - T_{isl})\) and \(I\), where one has Eq. \[6\] with \(L_{k2}(-1 + z) \approx -\pi^2/12 + \ln(2)/z\).

**Refrigeration neglecting phonons or photons.** Heat flow into the island is \(J_{total} \propto J_1\) for the devices in Fig. 2a, b; \(J_{total} = 2J_1\) for the thermocouple. The black curves in Fig. 1 are \(J_{total} = 0\), giving the steady-state temperature (solid for stable steady-states and dashed for unstable ones). Solid curves give the temperature the island will be cooled to by a current \(I\). Eq. \[9\] tells us the steady-state has \(I\) as a quadratic function of \(T_{isl}\); this approximation gives the catastrophe at \(I_{c} / (ek_BT_0) = 3(\ln(2)/\pi)^2 \approx 0.14\) with \(T_{isl}/T_0 = 6\ln(2)/\pi^2 \approx 0.42\), which is very close to the exact solution in Fig. 1.

**Refrigeration with phonons or photons.** Assume that chargeless particles carry heat ballistically from hot to cold (a reasonable assumption for phonons in clean samples below a few Kelvins \[41\]), then \(J_{ph} = \alpha(T_{isl}^2 - T_0^2)\) where \(\alpha = A \sigma_{SB}\) for an island with surface area \(A\) and emissivity \(\epsilon\), where \(\sigma_{SB}\) is the Stefan-Boltzmann constant for the phonons or photons. Fig. 3 shows that this moves the catastrophe to higher \(I\) and lower \(T_{isl}\), until at \(\alpha \propto h^3 T_0 / (k_BT_0)^2 \approx 1/8\) it is replaced by a discontinuity in the derivative of the steady-state curve. The effect remains qualitatively unchanged, if we take \(J_{ph} = \alpha(T_0^2 - T_{isl}^2)\) as for some photonic transmission \[42\].

**The transition to a continuous response.** A transition from the response in Fig 1 to a continuous one (evolving towards the nearly-linear response in Eq. \[4\]) occurs upon reducing the \(\epsilon\)-dependence of \(A_{ij}\). This reduces the thermoelectric response compared to the normal electrical response. We do this by putting the point-contact in parallel with another system with \(\epsilon\)-independent transmission; e.g. a tunnel-barrier. Fig. 3 shows the steady-state response of the point-contact in parallel with a load with conductance \(G_{load} \sim (e^2 / \hbar) \times g\) for various \(g\). For \(g \gtrsim 1\), the usual nearly-linear theory works. However for \(g = 1/4\) we see the deviations (cf. solid and dashed curves) predicted by the rule of thumb below Eq. \[4\], since \(\Pi\) of the combined system is larger than \(1/2\) when \(g < (3\ln(2) - 1)/2 \approx 0.53\) \[43\]. A transition in the steady-state’s behaviour occurs at \(g = g_c \sim 1/200\). For \(g < g_c\), the curve ceases to be single-valued, so \(T_{min}\) becomes a discontinuous function of \(I\).
Heat-engine efficiency. For $T_{isl} > T_0$, the circuit in Fig. 2 provides electrical power $P = IV$ to any load connected between L and R. Fig. 3 shows the ratio of the optimized-power (load chosen to maximize $P$) to the heat-current; a nonlinear method for finding this is outlined in the Supplementary Material.

For $T_{isl} \gg T_0$ we find that both the heat-flow carried by electrons and the electrical power go like $(T_{isl}/T_0)^2$, and the optimal efficiency saturates at $1 - \sqrt{1 - 6\ln[2]/\pi} \approx 15.9\%$ in the absence of phonons or photons. We have no simple argument why this curve is slightly peaked at $T_{isl}/T_0 \approx 0.88$. For finite coupling to ballistic phonons or photons, the curves decay to zero at large $T_{isl}$, because $J_{ph}$ goes like $(T_{isl}/T_0)^4$. If instead $J_{ph} \to \alpha(T_{isl}/T_0)^2$ [12], the optimal efficiency saturates at $1 - \sqrt{1 - 6\ln[2]^2/[\pi^2 + 6\alpha]}$, less than Eq. 1 predicts.

Concluding remark. The principle experimental challenges to observing the discontinuity in Fig. 4 are minimizing phonon and photon effects, and keeping the point-contact at pinch-off. However, irrespective of whether this is easily achievable, our results give an important message for all works on systems [8][12] showing large figures of merit, $ZT$. Having $ZT \gg 1$ is a strong hint that the nonlinear Peltier coefficient $\Pi$ may be larger than $\frac{1}{2} R$. If so, optimal response will occur deep in the non-linear regime, where $ZT$ ceases to give insight into the device’s true potential as a refrigerator or heat-engine. This is already the case for the point-contact whose $ZT$ is only 1.4; so we can certainly expect the same for systems with $ZT \gg 1$. In some cases (such as the point-contacts studied here) the device may be much better than its $ZT$ would imply.

Acknowledgements. I thank F. Hecking, G. Rastelli and C. Stafford, with particular thanks going to Ph. Jacquod, for their insightful comments at various stages of this work.

[1] H.J. Goldsmid, *Thermoelectric Refrigeration* (Temple Press, London, 1964). H.J. Goldsmid, *Introduction to Thermoelectricity* (Springer, Heidelberg, 2009).
[2] F.J. DiSalvo, Science **285**, 703 (1999).
[3] A. Shakouri and M. Zebarjadi Chapt 9 of *Thermal nanosystems and nanomaterials*, S. Volz (Ed.) (Springer, Heidelberg, 2009). A. Shakouri , Annu. Rev. Mater. Res. **41**, 399 (2011).
[4] F. Giazotto, T.T. Heikkila, A. Luukanen, A.M. Savin, J.P. Pekola, Rev. Mod. Phys. **78**, 217 (2006). J.T. Muhonen, M. Meschke, J.P. Pekola, Rep. Prog. Phys. **75**, 046501 (2012).
[5] G. Casati, C. Mejía-Monasterio, and T. Prosen, Phys. Rev. Lett. **101**, 016601 (2008).
[6] D. Nozaki, H. Sevincli, W. Li, R. Gutierrez, and G. Cuniberti, Phys. Rev. B **81**, 235406 (2010).
[7] K.K. Saha, T. Markussen, K.S. Thygesen, and B.K. Nikolić, Phys. Rev. B **84**, 041412(R) (2011).
[8] M. Wierzbicki and R. Swirkowicz, Phys. Rev. B **84**, 075410 (2011).
[9] O. Karlström, H. Linke, G. Karlström, and A. Wacker, Phys. Rev.. B **84**, 113415 (2011).
[10] T. Gunst, T. Markussen, A.-P. Jauho, and M. Brandbyge, Phys. Rev. B **84**, 155449 (2011).
[11] G. Rajput, and K.C. Sharma, J. Appl. Phys. **110**, 113723 (2011).
[12] P. Trocha and J. Barnaś, Phys. Rev. B **85**, 085408 (2012).
[13] S. Rajauria, P.S. Luo, T. Fournier, F.W.J. Hecking, H. Courtois, and B. Pannetier, Phys. Rev. Lett. **99**, 047004 (2007). S. Rajauria , P. Gandit , F.W.J. Hecking, B. Pannetier , H. Courtois, J. Low Temp. Phys. **154**, 211 (2009).
[14] L.W. Molenkamp, Th. Gravier, H. van Houten, O.J.A. Buikj, M.A.A. Mabesoone, and C.T. Foxon, Phys. Rev. Lett. **68**, 3765 (1992). H. van Houten, L.W. Molenkamp, C.W.J. Beenakker, and C.T. Foxon, Semicond. Sci. Technol. **7**, B215 (1992).
[15] U. Ghoshal, S. Ghoshal, C. McDowell, L. Shi, S. Cordes,
and M. Farinelli, Appl. Phys. Lett. 80, 3006 (2002)

[16] E.N. Bogachev, A.G. Scherbakov and U. Landman, Solid State Comm. 108, 851 (1998). This work calculates the non-linear differential Peltier coefficient \( \Pi_{\text{diff}} = dJ/dI \), even though some formulas assume linear behaviour.

[17] M.A. Çiğiloğlu, S. Turgut and M. Tomak, Phys. Stat. Sol. (b) 241, 2075 (2004).

[18] N. Nakpathomkun, H.Q. Xu, and H. Linke, Phys. Rev. B 82 235428 (2010).

[19] A.S. Vasenko, E.V. Bezuglyi, H. Courtois, F.W.J. Hekking, Phys. Rev. B 81, 094513 (2010). A. S. Vasenko, F. W. J. Hekking, J. Low Temp. Phys. 154, 221 (2009).

[20] M. Moskalets, JETP Lett. 62, 719 (1995).

[21] T. Christen and M. Büttiker, Europhys. Lett. 35, 523 (1996).

[22] D. Sanchez and M. Buttiker, Phys. Rev. Lett. 93, 106802 (2004).

[23] J. Mair and Ph. Jacquod, preprint [arXiv:1205.2705].

[24] D. Sanchez, and R. Lopez, preprint [arXiv:1209.1264].

[25] P.N. Butcher, J. Phys.: Condens. Matter, 2, 4869 (1990).

[26] B. Sothmann, R. Sánchez, A.N. Jordan, and M. Büttiker, Phys. Rev. B 85, 205301 (2012).

[27] B. Sothmann, and M. Büttiker, EPL 99, 27001 (2012).

[28] R. Sánchez and M. Büttiker, preprint [arXiv:1207.2557].

[29] H.-L. ENGQUIST and P.W. ANDERSON, Phys. Rev. B 24, 1151 (1981).

[30] U. Sivan and Y. Imry, Europhys. Lett. 91, 67009 (2010).

[31] Ph. Jacquod and R.S. Whitney, Europhys. Lett. 91, 67009 (2010).

[32] Y.-S. Liu, B.C. Hsu, and Y.C. Chen, J. Phys. Chem. C 115, 6111 (2011).

[33] Ph. Jacquod, R.S. Whitney, J. Mair, and M. Büttiker, preprint [arXiv:1207.1629].

[34] \( q \) is the thermal conductance for \( I = 0 \) (not \( V = 0 \)).

[35] R.S. Whitney, in preparation.

[36] M. Büttiker, J. Phys. Condens. Matter, 5, 9361 (1993). M. Büttiker, A. Prêtre, and H. Thomas, Phys. Rev. Lett. 70, 4114 (1993); Z. Phys. B, 94, 133 (1994). T. Christen and M. Büttiker, Phys. Rev. Lett. 77, 143 (1996).

[37] C. Petitjean, D. Waltner, J. Kuipers, I. Adagideli, K. Richter, Phys. Rev. B, 80,115310 (2009).

[38] J.S. Heron, T. Fournier, N. Mingo, O. Bourgeois, Nano Lett. 9, 1861 (2009).

[39] The results are unchanged if a little bit of tunnelling means that \( A_{\text{LM}}(\epsilon) \) goes smoothly from 0 to \( -1 \) on a scale less than \( k_{\text{B}}T_{\text{is}} \).

[40] W. Holmes, J.M. Gildemeister, P.L. Richards, and V. Kotsubo, Appl. Phys. Lett. 72, 2250 (1998). H.F.C. Hoovers, M.L. Ridder, A. Germeau, M.P. Brujin, P.A.J. de Korte, and R.J. Wiegerink, Appl. Phys. Lett. 86, 251903 (2005).

[41] L.M.A. Pascal, H. Courtois, and F.W.J. Hekking, Phys. Rev. B 83, 125113 (2011).

[42] This is similar to the para- to ferro-magnetic crossover of a magnet in a B-field upon reducing temperature. In \( \phi^2 \)-theory, the cross-over occurs when the \( \phi^2 \) term becomes negative, so the minimum is given by terms beyond \( \phi^2 \).

SUPPLEMENTARY MATERIAL

Self-consistent solution for linear screening. We model the point-contact as a one-dimensional scattering problem (along the x-axis), with the potential sketched in Fig. 2L. Close to the point-contact, this takes the form \( q V(x) = E_g - \kappa x^2 + q V_{\text{scr}}(x - x_{\text{pc}}) \) with energy measured from the chemical potential of the island. The transverse confinement generates the \( (E_g - \kappa x^2)\)-term, where \( E_g \) can be tuned, since it equals \( eV_g \) minus a geometry-dependent constant. The \( q V_{\text{scr}} \)-term is the screening due to the electron gas, which we take to be of the form

\[
V_{\text{scr}}(x) = \begin{cases} 
   a V_L (l_{\text{scr}} - x) / (2l_{\text{scr}}) & \text{for } x < -l_{\text{scr}} \\
   0 & \text{for } x > l_{\text{scr}} 
\end{cases}
\]

with \( x_{\text{pc}} \) self-consistently determined peak of \( q V(x) \). A little algebra gives \( x_{\text{pc}} = -aqV_L / (4l_{\text{scr}}) \), thus the energy at the peak is \( E_{\text{pc}} = q V_L (x_{\text{pc}}) = E_g + \frac{a}{2} q V_L (1 - a q V_L / (8k \sqrt{\kappa})). \) Finally we note that both \( a \) and \( l_{\text{scr}} \) depend on the scattering matrix of the junction, which in turn depends on \( E_{\text{pc}} \). To solve this problem self-consistently, we assume we are in the regime where \( eV_g = E_{\text{pc}} - E_g \) is small enough to approximate \( a = a_0 (1 + b_0 eV_g) \) and \( l_{\text{scr}} = l_{\text{scr0}} (1 + b_0 eV_g) \). If necessary, \( a_0, l_{\text{scr0}}, b_0 \) can be found by simulating Poisson’s equation; typically \( eV_g \) is small for small \( a \). Then \( E_{\text{pc}} \) is equal to a linear function of itself; re-arranging this gives

\[
E_{\text{pc}} = E_g + \frac{aqV_L / 2 - C(qV_L)}{1 - a_0 b_0 qV_L + 2C(qV_L) [b_0 - b_1]},
\]

where we define \( C(qV_L) = (aqV_L / l_{\text{scr0}})^2 / (16\kappa) \). As mentioned earlier, we assume that tunnelling at energies \( \epsilon < E_{\text{pc}} \) is negligible, so \( A_{\text{LM}}(\epsilon - E_{\text{pc}} > 0) = -1 \) and \( A_{\text{LM}}(\epsilon - E_{\text{pc}} < 0) = 0 \). To see that this respects gauge-independence, we recall that \( \epsilon, E_{\text{pc}}, V_{\text{L,k}} \) are all measured relative to the island’s potential, and replace them by quantities measured from a fixed external reference, so the island is at \( V_M \). For clarity, here (unlike in the paragraph containing Eq. 5) it is necessary to use a tilde to explicitly indicate quantities measured from the external reference. We make the replacement \( qV_L = (\tilde{\epsilon} - qV_M) - (\epsilon - qV_L) \). From this, we see that \( A_{\text{LM}}(\epsilon - E_{\text{pc}}) \) is only a function of the set of differences \( \{\tilde{\epsilon} - qV_k\} \), and so respects gauge-invariance.

Voltage and power dependence of cooling. Earlier we explained that it is best to consider the nonlinear response of thermoelectric devices as a function of current. However, for completeness we give the response of a single point-contact as a function of voltage or power. In Fig. 5L, the dashed-lines are contours of \( V(I, T_{\text{is}}) \) and solid-curves are contours of \( P(I, T_{\text{is}}) = IV(I, T_{\text{is}}) \); both \( V \) and \( P \) increase from top to bottom (with infinite voltage and power on the boundary of the red-triangle).

\[
E_{\text{pc}} = E_g + \frac{aqV_L / 2 - C(qV_L)}{1 - a_0 b_0 qV_L + 2C(qV_L) [b_0 - b_1]},
\]
as in Fig. 1, but with white used for all state, \( J = 0 \), marked by the black-curve. The color-scales are as in Fig. 1, but with white used for all \( hJ/(k_B T_0)^2 \) < -0.25, and the darkest color used for all \( hJ/(k_B T_0)^2 \) > 0.18.

These curves show that the steady-state is a continuous function of \( V \), but a discontinuous function of power; the former is clearly seen in the plot of Eq. (7) in Fig. S1. However one cannot extract the voltage-dependence of cooling for a circuit from this \( J(T_{\text{isl}}, V) \), because the relative voltage drop across each thermoelectric element depends nonlinearly on \( T_{\text{isl}} \). In general, the only way to get that response is to use \( J_{\text{total}} = \sum_j J_j(T_{\text{isl}}, I) \) where we sum \( j \) over each thermoelectric element. One can then use \( V(T_{\text{isl}}, I) \) for each element, to find the total voltage drop that generates \( I \) at a given \( T_{\text{isl}} \).

**Heat-engine efficiency in the non-linear regime.**

To calculate the maximum electrical power \( P(T_{\text{isl}}, I) \) that a heat engine can extract from a heat flow \( J(T_{\text{isl}}, I) \), one assumes a Ohmic load — so \( V(T_{\text{isl}}, I) = I/G_{\text{load}} \) is connected across its terminals, and adjust \( G_{\text{load}} \) to optimize the ratio of the power at the load \( P(T_{\text{isl}}, I) = I V(T_{\text{isl}}, I) \) to the heat flow \( J(T_{\text{isl}}, I) \). This corresponds to finding the \( I = I_{\text{opt}} \) which maximizes \( P(T_{\text{isl}}, I)/J(T_{\text{isl}}, I) \). The saddlepoint of \( P/J \) is given by \( P'J = PJ' \) where primed indicates \( (\text{d}/\text{d}I) \). If we have \( V(T_{\text{isl}}, I) \) and \( J(T_{\text{isl}}, I) \) we can solve this to find the optimal value of the charge-current \( I_{\text{opt}} \). The optimal efficiency is \( \eta = P(I_{\text{opt}})/J(I_{\text{opt}}) \); although one also has \( \eta = P'(I_{\text{opt}})/J'(I_{\text{opt}}) \), which is often easier to evaluate.

As a warm-up, we consider the usual linear problem, with \( V(T_{\text{isl}}, I) = S(T_{\text{isl}} - T_0) - G^{-1} I \) and \( J(T_{\text{isl}}, I) = \Theta(T_{\text{isl}} - T_0) + \Pi I \), with \( \Pi = \Gamma/G \) and \( S = B/G \). Using these equations to calculate the optimal efficiency in the manner described above, we find

\[
I_{\text{opt}} = \left( \Theta/\Pi \right) \sqrt{Z(T_{\text{isl}})T_{\text{isl}} + 1 - 1} (T_{\text{isl}} - T_0). \quad \text{Assuming } ZT \text{ is approximately } T \text{-independent, we get Eq. (1).}
\]

Now we use the same method to get the efficiency in the nonlinear regime. Unfortunately, in general we cannot get an analytic solution of \( P'J = PJ' \) from Eqs. (9,8), so we solve it numerically to find \( I_{\text{opt}} \) for different \( T_{\text{isl}}/T_0 \), and plot \( \eta \) against \((T_{\text{isl}} - T_0)/T_{\text{isl}} \) in Fig. 3c. However we can get an analytic result for large \((T_{\text{isl}}/T_0)\), using the fact (confirmed by the numerics) that in this limit

\[
V_1(T_{\text{isl}}, I) \approx \frac{k_B T_0}{e} \left[ -t_{\text{isl}} \ln[2] + \frac{h I}{e k_B T_0} \right],
\]

\[
J_1(T_{\text{isl}}, I) \approx -\frac{(k_B T_0)^2}{h} \left[ \frac{\kappa^2}{2} \ln[2] + \ln[2] \left( \frac{h I}{e k_B T_0} \right) t_{\text{isl}} - \frac{1}{2} \left( \frac{h I}{e k_B T_0} \right)^2 \right],
\]

where we define \( t_{\text{isl}} = T_{\text{isl}}/T_0 \) and \( \kappa = \pi^2/6 - \ln[2] \approx 1.16 \). The heat-current from the hot source into the device is \( J(T_{\text{isl}}, I) = -J_1(T_{\text{isl}}, I) \), and \( P = -V_1(T_{\text{isl}}, I) I \) (given that \( V_1 < 0 \)). In this case \( P'(T_{\text{isl}}, I_{\text{opt}}) J(T_{\text{isl}}, I_{\text{opt}}) = P(T_{\text{isl}}, I_{\text{opt}}) J'(T_{\text{isl}}, I_{\text{opt}}) \) is a quadratic equation for \( I_{\text{opt}} \), which we solve to find

\[
\frac{h I_{\text{opt}}}{e k_B T_0} = \frac{\kappa}{\ln[2]} \left[ \sqrt{1 + \ln^2[2]/\kappa} - 1 \right] \frac{T_{\text{isl}}}{T_0},
\]

from which we get the optimal efficiency given in the text.