Principles of a single-molecule rectifier in electrolytic environment

K. C. Matthew Cheung, Xiaoyu Chen, Tim Albrecht and Alexei A. Kornyshev*

Department of Chemistry, Faculty of Natural Sciences, Imperial College London, SW7 2AZ
*a.kornyshev@imperial.ac.uk

Abstract

The idea of gating the electrical current across a single chain molecule, confined between and linking two electrodes in electrolytic solution, in order to achieve an asymmetric current-voltage plot, was first put forward and substantiated with a detailed theory by Kornyshev and Kuznetsov (ChemPhysChem, 2006), and Kornyshev, Kuznetsov, and Ulstrup (PNAS, 2006). However, not all aspects of that effect have been studied in those papers. Its experimental confirmation, published by Capozzi et al. (Nature Nanotech, 2015), enthused us to revisit that theory, extending it and exploring all the regimes of system operation. In this article we present such comprehensive analysis, which reveals a set of new features. An important finding is that the introduction of more refined models of the electric double layer (beyond linear response) results in stronger rectification effects, already for relatively dilute electrolyte concentrations. The theory equally applies to electrode systems with and without full electrochemical potential control and highlights important differences for these two scenarios, for example with regards to the effect of electrode surface area.

1. Introduction

Electron transport mediated by single molecules has at least 60 year history beginning with the pioneering work of Taube and Myers,¹ for review see.² Bridge-assisted, or the so-called long-distance electron transfer has drawn much attention in the context of the ET mediating function of biological molecules such as proteins or DNA.² New kinds of experimental studies of these processes became possible due to progress in solid-state nanojunctions and electrochemical in situ STM,³,⁴ in the course of which it became clear that the single bridge molecules can be used to provide the transistor functions. First this has been achieved in the presence of redox molecules in the bridge, which was first predicted⁵ and then experimentally confirmed,⁶ giving a bell shape dependence of the current voltage plot with a maximum at the bias voltage corresponding to the energy of the lowest lying unoccupied level of the redox center. That area has been later widely developed⁷,⁸,⁹,¹⁰,¹¹,¹² with an intention to, potentially, exploit the putative or real diode- and transistor-like behavior of such junctions, central to molecular electronics.¹³,¹⁴

However, later another direction has evolved in this research area, boosted by creation of temporary nanoscale junctions,¹⁵,¹⁶ which appeared to be a good and to some extent complementary method for studying single-molecule charge transport in electrochemical environment. In this case, the conductance of the bridge molecules is studied when it is in a temporary contact with the STM tip and the substrate electrode. The signature of such contact, when it gets established, is a distinct modulation in the electrical current across the nanojunctions. For the junction mediated by chain-like alkanedithiols of different (controlled) length, Van Zalinge et al.¹⁶, has even managed to study the temperature dependence of the single molecule conductance. They found Arrhenius-like behavior with an apparent activation energy growing linearly with chain length. Such dependence was rationalized by detailed theory,¹⁷,¹⁸ showing that it is consistent with a tunneling mode of ET along the alkanedithiols. In contrast, the superexchange mechanism was found¹⁹ to give a completely different temperature dependence.
In the next paper,\textsuperscript{20} the current voltage characteristic of such junction embedded into electrolytic solution was studied theoretically, and the effect of rectification for the current as a function of the bias voltage was predicted.\textsuperscript{20} The same type of the current voltage characteristics was shown to follow in a superexchange mechanism of ET, just some parameters of the current voltage plot will be different.\textsuperscript{19} Both refs 19 and 20, have predicted the rectification, as a kind of gating effect, where the potential of each electrode was controlled independently relative to a common reference electrode (bipotentiotiostatic potential control). The effect was shown to increase with the ratio of the molecular bridge length $L$ to the Debye length $L_D$ in a given electrolyte; it vanishes when this ratio becomes smaller than 1.

Those works, however, have exploited linear response approximations for the polarizability of the electrolyte. As we will see in the present paper, account for the nonlinear response of the electrolyte plasma shows that the criterion on $L/L_D$ is stringent only for small electrode polarization. With the latter increasing, what becomes important in that criterion, is not the Debye length but the Gouy length $L_G$ which gets smaller with electrode polarizations. Thus, even for relatively low concentration of electrolyte, i.e. long Debye length (scaling as a square root of salt concentration), the rectification may be strong, because the Gouy length is small.

Developing the findings of refs 19 and 20 and extending the theory now becomes a necessity in view of the first experimental confirmation of the rectification effect in electrolytic environment.\textsuperscript{21} Not aware of refs 19 and 20, Capozzi et al.\textsuperscript{21} seemed to reproduce the main qualitative conclusion of those works: the possibility to achieve rectification in electrolytic environment. In our present paper, we explore all aspects of the rectification effect that have not been studied in the original refs 19 and 20, thus providing a comprehensive extension of that theory. The result of this analysis could navigate future systematic experimental exploration of the capabilities of electrochemical single molecule rectifiers.

The article is structured as follows. In section 2, we distinguish two different configurations of electrochemical single molecule rectifiers, to be considered in the paper. In section 3, we analyze the potential distribution in the nanogap, extending the theory of ref 22, and in section 4 discuss the consequences of these potential profiles for the tunneling barrier. In section 5, we first prepare the basic equations relating the current-voltage characteristic with the electrostatic potential profile in the nanogap, and then proceed to case studies, which highlight the effects of different parameters on the rectification. In addition, we apply the basic formalism to the case of "in situ" bipotentiotiostatic potential control, possible in electrochemical junctions with a common reference electrode. The case of two electrodes without such potential control, namely where only the potential difference between the two electrodes is controlled, is analyzed in section 6. In section 7, we discuss the possible effects of Stern layer on rectification, which may become important at large bias voltages. In section 8, we discuss modifications of the theory when the electron transport is dominated by the superexchange mechanism. Section 9 puts the electrolyte-mediated rectification in a broader perspective of various works on molecular diodes. Section 10 summarizes the main points important for further experimental exploration and design of electrochemically controlled single-molecule rectifiers.

2. Configuration of a single molecule electrochemical rectifier

Similarly to refs 19 and 20, we consider the configuration of nano-scale junction with potentially possible rectifying properties, sketched in Figure 1. It actually shows two variants of configuration, one in which the potentials of each electrode, $\Phi_s$ and $\Phi_t$ are controlled independently with respect to the reference electrode in the bulk of the solution (Figure 1a), and the one in which the bias voltage $V_{bias}$ (defined in eq 15) is directly controlled (Figure 1b). The subscripts s and t stand, respectively, for the
“substrate” and the “tip”, in reminiscence of STM configuration, but more generally could be any two electrodes making up the junction. We will assume that the lateral dimension of the electrodes is much larger than the distance between them. The cell contains electrolytic solution so that the double layers get formed at s- and t-electrodes. In the setup (Figure 1a) each double layer can be charged independently, positively or negatively; in the setup (Figure 1b) the net-charge in the double layer at one electrode must be equal with an opposite sign to the net charge in the other double layer. The gating ability is determined by the potential distribution in the nanogap, and we will thus start with the model analysis of this distribution.

When the molecule chain bridges the electrodes, both electrodes will have potential drops $\Phi_{s,0}$ and $\Phi_{t,0}$ to reach equilibrium in general. To make the algebra simpler, we will consider the two electrodes made of the same material and the two ends of the bridging molecule have the same quality of the interfacial contact with the electrodes, so that $\Phi_{s,0} = \Phi_{t,0}$. The extension on the case of different electrodes is trivial and can follow the recipe described in refs 19 and 20.

![Figure 1](image.png)

**Figure 1.** (a) Schematic representation of an electrochemical junction with full, bipotentiostatic control. Left: Substrate and tip are operated as independent working electrodes (WEs), which are connected via the single-molecule bridge. Their potentials $\Phi_s$ and $\Phi_t$ are defined independently with regards to a common reference electrode RE (in the theory $\Phi_{s,t} = 0$ corresponds to the potential of zero charge, pzc). The current flows between the counter electrode CE (not shown), WE1 and WE2; no (significant) current flows via the RE. The electrolyte solution contains ions that form electric double layers close to each WE. Right: An example of a bipotentiostatic control circuit,\textsuperscript{23} $E_1$ and $E_2$ are potentials, applied relative to ground. Note that the circuit common for WE2 is floating and offset by $E_1 - E_2$ compared to WE1. The output voltages $V_1$ and $V_2$ are controlled in such a way that the potentials of WE1 and WE2, measured via the RE, are equal to the specified values. In this way, both WE potentials can be controlled independently of each other. (b) Equivalent configuration, but now without bipotentiostatic control. $V_{bias}$ is still defined as $\Phi_t - \Phi_s$, but the individual potentials are floating.

### 3. Potential distribution in the nanogap

We analyze the potential distribution, using the same approach as ref 22, but explore more of the generally pertinent cases. The geometry of the surfaces confining the gap is assumed to be flat-parallel, with in the first instance the same surface area. With the spatial coordinate $x$ normal to the surface of the
electrodes, the s-electrode will be set at \( x = 0 \) and the t-electrode at \( x = L \). Neglecting edge effects, electrostatic potential will depend only on \( x \), and its distribution \( \Phi(x) \) across the nano-gap will be described by the nonlinear ‘Poisson-Fermi’ equation,

\[
\frac{d^2 \Phi}{dx^2} = \frac{\mu k_B T}{e L_D^2} \frac{\exp(-w) \sinh \left( \frac{e \Phi}{k_B T} \right)}{1 - \gamma + \gamma \exp(-w) \cosh \left( \frac{e \Phi}{k_B T} \right)}
\]

and \( L_D \) is the Debye length in the bulk of the solvent;

\[
\mu = \frac{\varepsilon}{\varepsilon_{\text{eff}}}
\]

where \( \varepsilon \) is the dielectric constant in the bulk of the solvent, \( \varepsilon_{\text{eff}} \) is the effective dielectric constant in the gap;

\[
\gamma = \frac{2c_0}{c_{\text{max}}}
\]

is the compacity parameter, in which \( c_0 \) is the average salt concentration in the bulk of the solvent, \( c_{\text{max}} \) is the maximum possible local concentration of ions near the electrodes (estimated to be 25.96\( M \) using cubic lattice structure as the maximal packing);

\[
w = \frac{W}{k_B T}
\]

where \( W \) is the free energy required to transfer individual ion from the bulk of the solution inside the gap (for simplicity we assume this energy to be the same for the cations and anions of the solution).

The above nonlinear equation (eq 1) can be reduced to the following 1\(^{st}\) order differential equation(s), depending on the signs of the boundary conditions \( \Phi_s \) and \( \Phi_t \). The situation with the opposite sign polarization of the electrodes (eq 5) was solved in ref 22, but the same sign situation was not covered in that paper. So we now present both cases, the same sign solution together with, for completeness, the earlier solved opposite sign one, because we will be showing the rectification results for the both regimes.

**Opposite sign solution** -

\[
\frac{d \Phi}{dx} = \text{sgn}(\Phi_t - \Phi_s) \sqrt{2\mu \left( \frac{k_B T}{e L_D} \right)} A + \ln \left( 1 - \gamma + \gamma \exp(-w) \cosh \left( \frac{e \Phi}{k_B T} \right) \right)
\]

**Same sign solution** -
\[
\frac{d\Phi}{dx} = sgn(\Phi_i) sgn(x - x_0) \sqrt{\frac{2\mu}{\gamma}} \left( k_B T \right) \ln \left( \frac{1 - \gamma + \gamma \exp(-w) \cosh \left( \frac{e\Phi}{k_B T} \right)}{1 - \gamma + \gamma \exp(-w) \cosh \left( \frac{e\Phi_0}{k_B T} \right)} \right)
\]

To avoid repeating lengthy equations in the rest of the paper, we introduce notations for the 'log-square root' parts of in eq 5 and eq 6.

\[
\Lambda(\Phi, A; \gamma, w, T) \equiv \sqrt{A + \ln \left( \frac{1 - \gamma + \gamma \exp(-w) \cosh \left( \frac{e\Phi}{k_B T} \right)}{1 - \gamma + \gamma \exp(-w) \cosh \left( \frac{e\Phi_0}{k_B T} \right)} \right)}
\]

\[
\Gamma(\Phi, \Phi_0; \gamma, w, T) \equiv \ln \sqrt{\frac{1 - \gamma + \gamma \exp(-w) \cosh \left( \frac{e\Phi}{k_B T} \right)}{1 - \gamma + \gamma \exp(-w) \cosh \left( \frac{e\Phi_0}{k_B T} \right)}}
\]

where \( A \) and \( \Phi_0 \) are the 'integration constants' appearing in the 'same sign solution', after taking the first integration of eq 1; for these we will have additional equations (see below): \( \Phi_0 \) is the potential for and \( x_0 \) is the position at which the magnitude of the electric field is zero within the gap. By separating the above variables, integral equations would be obtained. For opposite sign of potentials, two boundary conditions are known: \( \Phi(0) = \Phi_s, \Phi(L) = \Phi_t \). For same sign of potentials, three boundary conditions are known: \( \Phi(0) = \Phi_s, \Phi(L) = \Phi_t \) and \( \Phi(x_0) = \Phi_0 \). Therefore, the complete solution of the above differential equation would be:

**Opposite sign solution** -

\[
x = sgn(\Phi, -\Phi_s) \sqrt{\frac{\gamma}{2\mu}} \left( k_B T \right) \left[ \int_{\Phi_0}^{\Phi} d\Phi' \frac{1}{\Lambda(\Phi', A; \gamma, w, T)} \right]
\]

**Same sign solution** -

For \( 0 < x < x_0 \),

\[
x = -sgn(\Phi, \Phi_s) \sqrt{\frac{\gamma}{2\mu}} \left( k_B T \right) \left[ \int_{\Phi_0}^{\Phi} d\Phi' \frac{1}{\Gamma(\Phi', \Phi_0; \gamma, w, T)} \right]
\]

For \( x_0 < x < L \),

\[
x = sgn(\Phi_i) \left[ 2x_0 + \sqrt{\frac{\gamma}{2\mu}} \left( k_B T \right) \left[ \int_{\Phi_0}^{\Phi} d\Phi' \frac{1}{\Gamma(\Phi', \Phi_0; \gamma, w, T)} \right] \right]
\]

where the constants \( A, \Phi_0 \) and \( x_0 \) are determined by the following equations, which are again different for the opposite sign and the same sign \( \Phi_s \) and \( \Phi_t \):

**Opposite sign solution** -

\[
A(\Phi, \Phi_s; \gamma, w, T) \equiv \left[ \int_{\Phi_0}^{\Phi} d\Phi' \frac{1}{\Lambda(\Phi', A; \gamma, w, T)} \right]
\]

\[
\Phi_t(\Phi, \Phi_s; \gamma, w, T) \equiv \left[ \int_{\Phi_0}^{\Phi} d\Phi' \frac{1}{\Gamma(\Phi', \Phi_0; \gamma, w, T)} \right]
\]
\[ L = \text{sgn}(\Phi_t - \Phi_s) \sqrt{\frac{\gamma}{2\mu}} \left( \frac{eL_{\text{ip}}}{k_B T} \right) \int_{\phi_0}^{\Phi_t} \frac{1}{\Lambda(\Phi, \Lambda; \gamma, w, T)} \, d\Phi \]  

(12)

**Same sign solution -**

\[ L = \text{sgn}(\Phi_s) \sqrt{\frac{\gamma}{2\mu}} \left( \frac{eL_{\text{ip}}}{k_B T} \right) \left[ \int_{\phi_0}^{\Phi_s} \frac{1}{\Gamma(\Phi, \Phi_0; \gamma, w, T)} + \int_{\phi_0}^{\Phi_t} \frac{1}{\Gamma(\Phi, \Phi_0; \gamma, w, T)} \right] \]  

(13)

\[ x_0 = \text{sgn}(\Phi_s) \sqrt{\frac{\gamma}{2\mu}} \left( \frac{eL_{\text{ip}}}{k_B T} \right) \int_{\phi_0}^{\Phi_s} \frac{1}{\Gamma(\Phi, \Phi_0; \gamma, w, T)} \]  

(14)

The above integrals do not have analytical forms, but can be calculated numerically. Thus obtained electrostatic potential distributions are shown for some typical cases in Figure 2.

Important notes on terminology and definitions:

1) Even if we vary the potentials of each electrode independently, we can, conventionally, keep the definition of the bias voltage as follows,

\[ V_{\text{bias}} = \Phi_t - \Phi_s - (\Phi_{t,0} - \Phi_{s,0}) \]  

(15)

For \( \Phi_{s,0} = \Phi_{t,0} \), the expression reduces to \( V_{\text{bias}} = \Phi_t - \Phi_s \); unless specified, the graphs will be plotted for such situation. However as discussed in ref 24, \( \Phi_{s,0} \) and \( \Phi_{t,0} \) may vary with the surface area of the electrodes. Many plots in the rest of the paper, including Figure 2, will be classified subject to the value of \( V_{\text{bias}} \).

2) In the derivations of this section the potential \( \Phi(x) \) has been defined relative to the potential in the bulk of the electrolyte, the values of \( \Phi_s, \Phi_t \) taken as boundary conditions for the potential at the corresponding electrodes. However, in the measurement we determine them relative to the reference electrode. At infinite separation between the two working electrodes, the potential of zero charge of these electrodes (no charge net charge in the corresponding electrical double layers) must correspond to \( \Phi_s = 0 \) and \( \Phi_t = 0 \); but generally pzc-s of the electrodes relative to a given reference electrode are not zero. We will however, not introduce corresponding constant shift in our equations because in the end it will cancel out in eq 15.
4. The effect of potential distribution on the tunneling barrier for electrons

The differences in the form of the potential distribution in the nanogap will affect the tunneling barrier for electrons and thereby the current-voltage characteristic of the bridge-mediated nanogap.

Following ref 20, the overall potential barrier is composed of

$$V(x) = E_F + U - e\Phi(x) + V_{conf} \quad (16)$$

where $E_F$ is the Fermi level of the uncharged electrode, $U$ is the work function, and $V_{conf}$ describes the effect of conformational fluctuations of the molecule chain on the barrier.\textsuperscript{17}
Figure 3. A schematic representation of the effect of the polarization of the electrodes on the overall potential barrier in the nanogap, shown, generally, for different materials of the two working electrodes for (a) the same-sign $\Phi_t, \Phi_s > 0$ (left panel) or $\Phi_t, \Phi_s < 0$ (right panel) and (b) opposite-sign $\Phi_t < 0, \Phi_s > 0$ (left panel) $\Phi_t > 0, \Phi_s < 0$ (right panel). New symbols are introduced for this schematic representation: $U_s$ and $U_t$ represent the work functions of the substrate and the tip respectively, $E_F^s$ and $E_F^t$ represent the Fermi levels of uncharged electrodes, respectively.
At $\Phi_t, \Phi_s > 0$, Fermi levels of both the substrate and the tip will be lowered, while the electrostatic potential distribution $\Phi(x)$ will have an open upward shape. Hence the overall potential barrier $V(x)$ within the gap is generally the highest for electrons to tunnel through, where the barrier attains its maximum in the middle of the gap (Figure 3a (left panel)). If both $\Phi_t, \Phi_s < 0$, Fermi levels of both the substrate and the tip will be raised, while the electrostatic potential distribution $\Phi(x)$ will have an open downward shape. Consequently, the overall potential barrier $V(x)$ has a minimum in the middle of the gap (Figure 3a (right panel)) and is overall lower for electrons to tunnel through (high tunneling probability).

5. Current-voltage characteristic (CV) for the molecule electron transport across the nanogap, mediated by a bridge molecule

5.1. Basic equations

Following ref 20, the current-voltage (I-V) characteristic relationship in the tunneling mode of electron transport can be described by the following equation,

$$I = \frac{4\sigma U}{\beta_0 L} \sinh \left( \frac{e\beta_0 LV_{\text{bias}}}{4U} \right) \exp \left[ \frac{e\beta_0}{2U} \int_0^L \left( \Phi(x) - \frac{\Phi_s + \Phi_t}{2} \right) \right]$$  \hspace{1cm} (17)

$$\beta_0 = \frac{2\sqrt{2mU}}{h}, \quad I_0 = 4\sigma U$$  \hspace{1cm} (18)

where $I$ is the current, $\sigma$ is the temperature dependent conductance of the single molecule chain and $I_0$ is an expression of the dimensional current.

In the nonlinear model of screening, the solution for $\Phi$ cannot be written explicitly in terms of position $x$. However, the derivative of the potential distribution is known. Hence, it is convenient to rewrite the integral in eq 17 in terms of integration over $\Phi$ instead of $x$ (change of variable). Since the expressions for the derivative $\frac{d\Phi}{dx}$ differ in the case of the same sign and opposite sign of $\Phi_s$ and $\Phi_t$, the integral we are after must be written for each case separately. Accordingly, the integrals in the I-V formula (eq 17) are obtained as follows:

**Opposite sign solution** -

$$\int_0^L d\Phi = \text{sgn}(\Phi_s - \Phi_t) \sqrt{\frac{\gamma}{2\mu}} \left( \frac{eL_D}{k_BT} \right)^{\Phi_t}_{\Phi_s} d\Phi \frac{\Phi}{\Lambda(\Phi; \gamma; w, T)}$$  \hspace{1cm} (19)

**Same sign solution** -

$$\int_0^L d\Phi = \text{sgn}(\Phi_s) \sqrt{\frac{\gamma}{2\mu}} \left( \frac{eL_D}{k_BT} \right)^{\Phi_s}_{\Phi_s} d\Phi \frac{\Phi}{\Gamma(\Phi; \Phi_0; \gamma; w, T)} + \int_{\Phi_0}^{\Phi_s} d\Phi \frac{\Phi}{\Gamma(\Phi; \Phi_0; \gamma; w, T)}$$  \hspace{1cm} (20)
Using eqs 12 and 13 that determine the numerical values of $A$ and $\Phi_0$, the latter must be substituted into the above integrals (eqs 19 and 20) to determine the current for given values of $\Phi_s$ and $\Phi_t$.

Recall that the linearized Poisson-Boltzmann model used in refs 20 and 19 gives a very simple potential distribution, good for any sign of

$$
\Phi(x) = A \exp\left(\frac{x}{L_D}\right) + B \exp\left(-\frac{x}{L_D}\right),
$$

$$
A = \frac{\Phi_t - \Phi_s \exp(-L/L_D)}{\exp(L/L_D) - \exp(-L/L_D)}, \quad B = \frac{-\Phi_t + \Phi_s \exp(L/L_D)}{\exp(L/L_D) - \exp(-L/L_D)}
$$

Inserting eq 21 into eq 17 gave an elegant formulae for the I-V curve (similar expression can be found in ref 19 but with different parameters involved, as derived for the superexchange mechanism of electron transport).

$$
I = \frac{\sigma 4U}{\beta_0 L} \sinh\left(\frac{e\beta_0 L V_{bias}}{4U}\right) \exp\left[-\beta_0 L \left(\frac{e(\Phi_s + \Phi_t)}{4U}\right)\right] \left[1 - \frac{\tanh(\kappa L/2)}{\kappa L/2}\right]
$$

### 5.2. The effect of different parameters, approximations, and operation modes

**Different electrolyte concentrations.**

As shown in refs 20 and 19, the concentration of electrolyte is one of the central parameters for rectification. There, its effect was demonstrated through the concentration dependence of the Debye length. This is sufficient for diluted solution, but at high concentrations, it also emerges from the concentration dependence of parameter $\gamma$, as shown in Figure 4.

![Figure 4. Current voltage characteristics shown for different electrolyte concentrations: rectification enhances with concentration increase.](image)

Voltage bias here results from varying $\Phi_t$ at $\Phi_s = 0V$. Other parameters: $L = 2.5nm$, $\epsilon = 80\epsilon_0$, $\mu = 1$, $w = 0$, $U = 5eV$, $T = 298K$, $c_{max} = 25.96M$. The chosen concentrations correspond here to the following values of $\gamma$ (determined by eq 3) and Debye lengths: (blue solid) $\gamma = 2.91 \times 10^{-2}$, $L_D = 0.5nm$; (green dash) $\gamma = 7.27 \times 10^{-3}$, $L_D = 1nm$; (magenta dotted) $\gamma = 1.16 \times 10^{-3}$, $L_D = 2.5nm$; (black solid) $\gamma = 1.16 \times 10^{-7}$, $L_D = 250nm$.

From Figure 4, we see that the increase of concentration systematically enhances rectification, at all other conditions unchanged. The physics of it is basically the same as described in ref 19. For high concentration of ions, the Debye screening length is short, and the Gouy length (the actual double layer extension) is also small. Thus, the potential barrier is more strongly deformed. But under positive bias, electrons will have to tunnel through the high and convex potential barrier, while under negative bias,
they tunnel through a lower and concave potential barrier (its concave character gets sharper for shorter Debye lengths).

Importantly, for very low concentration of ions, the Debye length will at some point get longer than the gap width. In such limit, the electrostatic potential profile becomes more like a linear function with a negligible screening effect, hence the tunneling electrons experience practically no difference between positive and negative bias voltage. This explains why the rectification disappears at very low salt concentration, like in the case of the \( \mu M \) one in Figure 4. In the same graph, the \( \pm 100 \text{mV} \) stretch of almost linear I-V for a centimolar concentration is a manifestation of the same effect. It is apparently linear not because the I-V itself should be linear, but because it results from a competition: at moderate voltages the counterions have not been drawn into the gap in sufficiently large quantities and the I-V curve would have shown a positive second derivative, similarly to the I-V of an ion-free dielectric gap, but this trend is compensated by incoming counterions, which tend to change the second derivative negative. As a result the curve looks straight, until with further increase of voltage the gap will be sufficiently enriched with counterions, and we will start seeing rectification.

Different length of the bridge molecule/gap size.

![Figure 5](image)

**Figure 5.** The effect of molecular length on I-V curves: longer chain molecules provide stronger rectification. The curves are labeled with the corresponding values of \( L \), in nm. Other parameters: \( \varepsilon = 80\varepsilon_0 \), \( L_D = 1 \text{nm} \) (\( \Rightarrow c_0 = 0.0941M, \gamma = 0.00727 \)), \( \mu = 1, \ w = 0, \ U = 5eV, \ T = 298K \). The bias voltage is varied by changing \( \Phi_t \) at \( \Phi_s = 0V \).

This is demonstrated in Figure 5: longer chain of molecules provide stronger rectification. The electric field is mostly screened after a distance of the Debye length (or even shorter at high voltages it becomes a Gouy length) and the potential barrier has a more extended flat region in between the electrodes. Hence the difference between the positive and negative polarization will be more pronounced, because in latter case the flat region is at the bottom of the concave barrier, and in the former case it is at the top of the convex barrier. This explains why longer chain molecules experiences stronger rectification. This has been already explained in refs 20 and 19.

**The nonlinearity of screening: how does it affect rectification?**

Here we compare the results with those of the linear response theory. Curves in Figure 6 show the results for the current-voltage plots, obtained in the linear model and the nonlinear theory. Since the linear model follows from the nonlinear one when \( e\Phi(z) \ll k_B T \), assumed that the electrical potential is relatively smaller than the thermal energy, hence the two approaches give the results in the narrow proximity \( V_{bias} = 0V \) (as the bias voltage has been varied by changing \( \Phi_s \) at \( \Phi_s = 0V \)). In the high
concentration limit, corresponding to short Debye length (Figure 4a), both approaches give rise to strong rectification. For the Debye length longer than the length of the molecule, the linear model gives no rectification as the potential distribution across the gap becomes linear. However, in the nonlinear model what matters is the Gouy length which for large polarizations of the electrode becomes substantially shorter than the Debye length, and hence rectification survives even at relatively larger concentrations. The long linear stretch has the same origin as discussed in the context of Figure 4.

(a) 
(b)

Figure 6. Nonlinear screening amplifies rectification. Comparison of I-V curves calculated via the linearized model (green/dashed curves) and the nonlinear model (blue/solid curves). Parameters: $L = 2.5nm$, $\varepsilon = 80\varepsilon_0$, $\mu = 1$, $w = 0$, $T = 298K$; (a) $L_D = 0.5nm$, (b) $L_D = 10nm$. The bias voltage is varied by changing $\Phi_t$ at $\Phi_s = 0V$. The linearized model is independent of $\gamma$, unlike the nonlinear model, for which the calculated value of $\gamma$ is shown. Remarkably, with account for nonlinear screening, rectification takes place even at Debye length four times larger than the length of the molecule!

The influence of the special properties of the nanogap.

In the curves plotted so far, the properties of the electrolyte in the gap were taken to be the same as in the bulk. However, generally speaking, this may require correction. Below we will take this into account through the parameters $\mu \neq 1$ and $w \neq 0$. Varying $\mu > 1$, we assume that the dielectric polarizability of the solution in the gap may be ‘frozen’, delivering a lower effective dielectric constant. Varying $w$, we cover the cases when it takes additional free energy for an ion to get into the gap ($w > 0$) or alternatively its free energy gets lower when it gets there. The net value of $w$ is determined by the competition between the entropic disfavor for an ion to get into a tighter confinement of the gap and image force attraction to metallic electrodes, and generally we do not know the result of this competition. We first, formally, vary $\mu$ at $w = 0$, and $w$ at $\mu = 1$. 
Figure 7. Reducing the dielectric constant in the gap alone, at an unchanged propensity of ions to get there, would have increased rectification. I-V curves plotted for indicated values of $\mu$. Other parameters: $L = 2.5\,nm$, $\varepsilon = 80\varepsilon_0$, $L_D = 1\,nm$ ($\Rightarrow c_0 = 0.0941M$, $\gamma = 0.00727$), $w = 0$, $U = 5eV$, $T = 298K$. The bias voltage is varied by changing $\Phi_t$ at $\Phi_s = 0V$.

For $\mu > 1$, which means $\varepsilon > \varepsilon_{eff}$, i.e. dielectric constant in the bulk liquid is larger than that in the gap, the Debye length in the bulk liquid is longer than in the gap. Hence for larger $\mu$, the ions will form more densely packed double layers near the electrodes. This results in stronger rectification for larger $\mu$ (at a fixed $w$), as shown in Figure 7.

Figure 8. Increasing the propensity of ions to get into the gap (at unchanged dielectric properties of the gap) enhances rectification. I-V curves plotted for indicated values of $w$. Other parameters: $L = 2.5\,nm$, $\varepsilon = 80\varepsilon_0$, $L_D = 1\,nm$ ($\Rightarrow c_0 = 0.0941M$, $\gamma = 0.00727$), $\mu = 1$, $U = 5eV$, $T = 298K$. The bias voltage is varied by changing $\Phi_t$ at $\Phi_s = 0V$.

The trend in Figure 8 is easy to explain. Negative values of $w$ act like an extra ‘push’ to move ions into the gap, like an attractive potential, while larger/positive values of $w$ act like a repulsive potential. With $w < 0$, that extra push helps counterions to build more compact double layers: negative values of $w$ counterbalance the repulsion between the counterions. To remind us once more, to which effect this ‘counterion condensation’ leads: for such negative $w$, electrons have to tunnel through a higher potential barrier under a positive bias voltage at a fixed $\Phi_s$ ($\Phi_s = 0V$ in the considered example), while under negative voltage that barrier is smaller. This explains why the rectification is stronger if $w$ is negative. Positive $w$ tend to exclude ions from the gap, making the screening length smaller and reducing rectification.

Independent variation $\mu$ and $w$ in Figure 7 and Figure 8 is, however, a fiction, which helps to reveal the role of each of these model parameters. Indeed, both $\mu$ and $w$ would vary simultaneously with the variation of the properties of the gap and, in the first place, the gap width. To parameterize the dependence of $\mu$ and $w$ on , let us use an interpolation formulae, proposed in ref 22.
\[ \mu = 1 + (v - 1) \exp \left( -\frac{L}{\lambda} \right) \]  

(23)

where \( v = \epsilon / \epsilon^{*} \) (assumed to be 10) and \( \epsilon^{*} \) is the short range dielectric constant in the narrow gap, and \( \lambda \) is the correlation length.

\[ w = -\zeta z^{2} \frac{L_{p}}{L} \left[ 1 - 2 \exp \left( -\frac{\epsilon_{0}}{\epsilon_{\text{eff}}} \frac{L}{d^{*}} \right) \right] + \zeta^{2} \frac{L_{p}}{2a} \left( \frac{\epsilon}{\epsilon_{\text{eff}}} - 1 \right) \]  

(24)

where \( \epsilon_{0} \) is the vacuum permittivity, \( d^{*} \) is the penetration depth into the metal (assumed to be 0.5\( \text{Å} \)), \( \zeta \) is the gap-shape factor (assumed to be 1 for flat gap), \( a \) is the radius of the ion (assumed to be 2\( \text{Å} \)), and \( z \) is the valence number of the ion (assumed to be 1). Note that eq 24 unambiguously describes the increase of \( w(L) \), it does not cover a possibility of an opposite behaviour.

The effect of variation of the thickness of the gap which affects both \( \mu \) and \( w \) is shown in Figure 9.

**Figure 9.** The combined effect of \( \mu \) and \( w \) using the new approximation formulae: the effect of \( w \) overcomes the effect of \( \mu \) when the width of the nanogap is small. I-V curves plotted for indicated values of \( \mu \) and \( w \). Other parameters: \( \epsilon = 80 \epsilon_{0}, L_{D} = 1\, \text{nm} \) \( \Rightarrow c_{0} = 0.0941\, \text{M}, \gamma = 0.00727 \), \( U = 5\, eV, T = 298K \); (a) \( L = 3.5\, \text{nm} \) (b) \( L = 1\, \text{nm} \). The bias voltage is varied by changing \( \Phi_{t} \) at \( \Phi_{s} = 0\, V \).

When the width of the gap is large, the variation of both \( \epsilon_{\text{eff}} \) (eq 23) and \( w \) (eq 24) is negligible and has no effect on the I-V curve. As a result, Figure 9a shows no new features when approximations eqs 23 and 24 are involved, as compared to the case of constant \( \mu \) and \( w \). However, when the width of the gap is small, \( \epsilon_{\text{eff}} \) (eq 23) will get smaller and \( w \) (eq 24) will get large (more positive) enough to have effect on the system. From the analysis given by Figure 7 and Figure 8, we have learned that the rectification is stronger for larger \( \mu \) (smaller \( \epsilon_{\text{eff}} \)), and weaker for larger/more positive \( w \), hence the two factors (\( \epsilon_{\text{eff}} \) and \( w \)) work against each other in Figure 9b. From Figure 9b, we see that the rectification is indeed weaker if the approximation (eqs 23 and 24) is taken into account, because here the effect of larger \( w \) overcomes
the effect of larger $\mu$. This result warns us against trying to build a rectifier with very narrow nanogaps, where such trend may occur.

The modes of electrode polarization matter

We now explore which effect different modes of control of $\Phi_s$ and $\Phi_t$ has on rectification. Figure 10 shows I-V curves plotted at different fixed values of $\Phi_s$ and $\Phi_t$ respectively. For the chosen set of parameters, the two graphs are in principle equivalent to each other, as one can be obtained from the other by the corresponding inversion - the positive bias region in Figure 10a corresponds to the negative bias region in Figure 10b, and the negative bias region in Figure 10a corresponds to the positive bias region in Figure 10b. As can be seen from eq 15, the visual difference is due to the sign convention and the current direction; there is no difference in the absolute magnitude of current. For example, $\Phi_s = 0.3V$, $\Phi_t = -0.2V$ and $\Phi_s = -0.2V$, $\Phi_t = 0.3V$, the only difference would be the direction of current.

![Figure 10. Fixing different values of $\Phi_s$ and $\Phi_t$ respectively: higher magnitude of tunneling current if both $\Phi_s$ and $\Phi_t$ are negative, while smaller magnitude of tunneling current when both $\Phi_s$ and $\Phi_t$ are positive. I-V curves plotted for indicated values of (a) $\Phi_s$ and (b) $\Phi_t$. Other parameters: $L = 2.5nm$, $\varepsilon = 80\varepsilon_0$, $L_D = 1nm$ ($\Rightarrow c_0 = 0.0941M, y = 0.00727$), $\mu = 1$, $w = 0$, $U = 5eV$, $T = 298K$.](image)

As there is no physical difference between Figure 10a and Figure 10b above, we focus to the features displayed in Figure 10a. If $\Phi_s$ is fixed positive and $V_{bias}$ is positive, then $\Phi_t$ is also positive and larger than $\Phi_s$. If $\Phi_s$ is fixed negative, in order to maintain negative bias, $\Phi_t$ has got to be also negative and larger in absolute value than $\Phi_s$. From the energy diagram (Figure 3), it is clear that electrons have a higher potential barrier to tunnel through if both $\Phi_s$ and $\Phi_t$ are positive, while electrons have a lower barrier to tunnel through if both $\Phi_s$ and $\Phi_t$ are negative. These expectations agree with Figure 10a that the magenta line ($\Phi_s = 0.3V$) has negligible current under positive bias, while the blue line ($\Phi_s = -0.3V$) has the highest magnitude of current under negative bias.

It is important to make clear the mode of polarization of the electrodes, $\Phi_s$ and $\Phi_t$, when rectification does not take place. Indeed in all the above considered examples, the potential barrier for the tunneling
electrons under positive bias is different from that under negative bias, and this causes the rectification. When the barrier for tunneling has the same shape for positive and negative bias, there will be no rectification. This will happen if we apply $\Phi_s = -\Phi_t$. In Figure 11a, the potential distributions of $V_{bias} = 0.4V$ and $V_{bias} = -0.4V$ are shown respectively, for which the resulting potential barriers have the shape, that can be obtained by inversion one from each other. Therefore, the corresponding I-V curve (Figure 11b) shows no rectification.

![Figure 11a](image1)

![Figure 11b](image2)

Figure 11. When rectification disappears: antisymmetric polarization of electrodes: $\Phi_s = -\Phi_t$. (a) Potential distributions plotted for indicated values of $\Phi_s$ and $\Phi_t$. (b) I-V curve for $L = 2.5nm$, $\varepsilon = 80\varepsilon_0$, $L_D = 1nm$ ($\Rightarrow c_0 = 0.0941M, \gamma = 0.00727$), $\mu = 1$, $w = 0$, $U = 5eV$, $T = 298K$. The bias voltage is varied by changing $\Phi_t$ at $\Phi_s = 0V$.

6. The two electrode geometry without bipotentiostatic control

Having considered the electrodes the potentials of which can be controlled by independent reference electrodes (c.f. Figure 1a), we now focus on a more specific set-up, in which the two electrodes is connected with a battery (c.f. Figure 1b). In this setup, the double layer at each electrode must have the same absolute values of the net charge, of opposite signs, of course. Hence, if we allow the surface areas of the two electrodes, $S_s$ and $S_t$, be different, we get the condition on surface charge densities,

$$|\sigma_s| = \frac{S_s}{S_t} |\sigma_t| \quad (25)$$

where $\sigma_s$ and $\sigma_t$ are the surface charge densities of the substrate and tip, respectively, $S_s$ and $S_t$ are the surface areas of the two electrodes, respectively.

By considering the boundary conditions on the surfaces of the two electrodes, the surface charge density can be related to the potential of each electrode,

$$\frac{|\sigma_t|}{\varepsilon_{eff}} = \sqrt{\frac{2\mu}{\gamma} \left( \frac{k_B T}{eL_D} \right)} A(\Phi_s, A; \gamma, w, T) \quad (26)$$
\[
\left| \sigma_s \right| = \sqrt{\frac{2\mu}{\varepsilon \gamma}} \left( \frac{k_B T}{e L_D} \right) \Lambda(\Phi_s; A; \gamma, w, T) \tag{27}
\]

where \( A \) is the integration constant, as in eq 5, and the system is expected to have opposite sign of potentials due to the conservation of charge. Given the value of \( |\sigma_s| \) and the ratio of the two areas (eq 25), both \( \Phi_s \) and \( \Phi_t \) can be determined from eqs 26 and 27 numerically. Using again the above equations (eqs 12-20), the tunneling current can then be determined. In a real experiment, one will vary \( V_{bias} \) (battery) as independent variable. However, using this formula, it will technically be easier to run the values of \( |\sigma_s| \) with a fixed area ratio, calculating \( V_{bias} \) via eqs 26 and 27 numerically, and then to plot the I-V curve parametrically in terms of \( |\sigma_s| \) (as hidden running variable).

Note that from eqs 26 and 27, the values of \( \Phi_s \) and \( \Phi_t \) can only be calculated numerically, except for the case when the two areas are the same. If the two electrodes have same area, eqs 26 and 27 can be solved analytically as follows,

\[
\Phi_s = -\Phi_t \tag{28}
\]

Hence, according to the previous analysis, rectification effect will not be observed in this case (Figure 11b). If the areas are different, the rectification effect emerges, as the system loses symmetry.

Before moving on to a more quantitative discussion of this scenario, a particular subtlety is worth noting. Namely, if the nanogap electrodes have different electroactive areas, their open-circuit potentials \( \Phi_{s,0} \) and \( \Phi_{t,0} \), respectively, will be different too (say, measured relative to a common reference electrode).\(^{24}\) The effect can be quite significant, i.e. on the order of 100 mV, depending on the electrode area. The effect arises because the open-circuit voltage between two electrodes is defined by the net current being zero, which in turn is the sum of the anodic and cathodic microscopic currents on each electrode. The latter are affected by the electrode area and hence also the net current. However if the current is plotted as a function \( V_{bias} \) defined as in eq 15, the result formally would not be changed. Had we plotted the current as a function of \( \Phi_t - \Phi_s \) at \( \Phi_{s,0} \) and \( \Phi_{t,0} \) not cancelling each other, this would have led to a horizontal shift of the I-V curve, thus also affecting the apparent rectification ratio.

In principle, the area dependence of the open-circuit potential can be calculated\(^{24}\) or measured. Its value will depend, \textit{i.e.} on the temperature, the electrode material used, the surface properties of the electrode and its geometry, and the composition of the solution. Due to this strong dependence on the experimental circumstances, we thus suggest to measure the open-circuit potential of each nanogap electrode at the beginning of an experiment and then rescale the experimental results for comparison with the theory.

Figure 12b clearly demonstrates that rectification is observed only when the two electrodes are not of the same size. At the same bias voltage \( V_{bias} \) shown in Figure 12a, the potential drop on each side of the electrode is the same if the areas are the same, while if the surface charge density on one side of electrode is higher, then the potential drop on that side is larger. If the charge density on one side is much larger than the other side (magenta curves in Figure 12), the overall potential barrier is higher for electrons to tunnel through under positive bias (due to the flatter shape near the large electrode), while the overall potential barrier is lower for electrons to tunnel through under negative bias. This explains why the rectification effect is stronger when the area ratio is higher (Figure 12b).
Figure 12. Larger difference in surface areas of the two electrodes connected via a battery produces stronger rectification. Both graphs are plotted for indicated values of the ratio $S_s/S_t$. (a) Potential Distribution under same bias voltage (b) I-V curves with different ratios of surface area. Other parameters: $L = 2.5\text{nm}$, $\varepsilon = 80\varepsilon_0$, $L_D = 0.5\text{nm}$ ($\Rightarrow c_0 = 0.3764\text{M}$, $\gamma = 0.02911$), $\mu = 1$, $w = 0$, $U = 5eV$, $T = 298K$. The bias voltage is varied by changing $\Phi_t$ at $\Phi_s = 0V$.

It can thus be concluded that for the two-electrode setup, rectification is not expected if the two electrodes have same area, while the rectification effect will likely get stronger if the area ratio gets larger.

7. Going to higher electrode polarization: the role of the compact layer correction

As is well known in electrochemistry, describing the potential distribution with just the diffuse double layer theory stops working at high voltages and high electrolyte concentrations. When the Gouy length is long, the major part of the potential drop will fall across the diffuse layer. Then the description of the potential distribution in the gap that we described above is sufficient. However the Gouy length gets shorter with the increase of the voltage drop across the double layer (and the shorter the higher the electrolyte concentration). In that case, introduction of the compact layer will be needed to explain the experimental data for the measured capacitance and the current voltage characteristics of electrode reactions. We refer a reader less familiar with electrochemistry to textbooks or more up-to-date reviews.

The simplest form of that correction is known to be a Stern layer correction, which adds the so called compact layer capacitance in series to the diffuse layer one. The compact layer contribution takes into account that ions cannot come closer to the surface than to the distance of a radius of hydrated ion. Whereas the parameter $\gamma$ takes into account the excluded volume effect in the packing of ions relative to each other, the compact layer correction accounts for the limited ability of ions to come too close to the surface. There were many extensions of that model and the ways to go beyond it. But in this article we focus on qualitative effects and do not need to go into this complexity. We extend our theory of the
potential distribution in the nanogap with account of the classical description of compact layer contribution, appending it in order not to overload the main text with additional algebra and information.

The upshot of that analysis is, however, twofold: (1) the introduction of the compact layer makes the potential distribution even more inhomogeneous, which amplifies the rectification; (2) there is very little or no difference between the linear and nonlinear theory in the case of high electrolyte concentration, but there is a difference when the concentration is low. The latter point is obvious, because at high electrolyte concentration the potential drop across the diffuse part of the double layer is small and it does not matter much with which theory one describes it. When the concentration is low, the diffuse layer part dominates in the overall potential drop and the decrease of the diffuse layer portion with increasing voltage will be stronger in the nonlinear theory, hence stronger is the rectification effect. All these trends are clearly demonstrated in Figure A.2 and Figure A.3 of the Appendix A.

8. Superexchange Model

Mechanisms of electron transport through chain molecules are usually considered within three categories: (1) consecutive site-to-site hopping, (2) direct tunneling, and (3) super-exchange. In consecutive hopping, after each elementary hopping act, the environment equilibrates. In superexchange, it does not. The electrons occupy only virtually the LUMO levels of the molecules in the course of their transport across the molecule. These states serve as bridge states and they lie high enough, relative to the energy of the tunneling electrons.

The work in ref 20 was motivated by measurements of Van Zalinge et al. on the current through alkanedithiols, which has revealed an Arrhenius dependence of the current on temperature and an almost linear dependence of the measured activation energy on the chain length. It was possible to explain that dependence on the basis of tunneling coupled to conformational fluctuations. In our present paper, we so far focused our analysis on the direct tunneling mechanism based on the Gamow formula, which seems to be quite appropriate for alkanedithiols.

However, for some other bridge molecules the contribution of virtual intermediate states may be significant, and the superexchange mechanism may dominate the electron transport. That mechanism has been explored in ref 19 based on the Breit-Wigner scattering formula. It was found that although the expressions for the current and its temperature dependence are very different for the direct tunneling and superexchange mechanisms, the current voltage part of this dependence, determined by the electrostatic potential distribution in the nanogap, is in the end governed by the same equations, resulting in the same form of $j(V_{bias})$.

For both mechanisms (tunneling and superexchange), the current-voltage dependence have been studied within the linearized version of the ‘Poisson-Fermi’ equation. In the present paper, we have extended the nonlinear theory of the potential distribution in the nanogap, and found an economic way how to integrate these results into the expression for the current-voltage characteristic. This was done within the direct tunneling model, but we can easily use the same approach for the superexchange model. The recipe for doing this, together with more details on the physics and the basic equations of the two mechanisms, is presented in the Appendix B.
9. Discussion

In this article we have focused on a specific, but rather universal mechanism of the rectification of electrical current across molecular junctions imbedded into electrolytic solutions. We may call this mechanism ‘electrostatic’ as it is related with the electrical double layer effects that modify the barrier for electron tunneling. We have revealed the situations when in such systems there should be rectification and when it should disappear. This is not the mainstream research in molecular diodes, as only several cited papers on this mechanisms and electrolytic junctions have been published. It is, therefore, expedient to briefly review other approaches to molecular diodes and put the option of electrostatic rectification into a wider perspective. Note that there are many papers on this subject, but we will pick up just a few, most pertinent for this discussion.

The idea of single molecule rectification is commonly attributed to the pioneering theoretical analysis by Aviram and Ratner. In its essence, this early work suggested to utilize different thresholds for the electron injection at positive and negative bias. In one direction, the threshold is low and electronic current could flow already at small voltages; in the opposite direction, it will be absent below a much large absolute value of the voltage. To realise such system, they suggested to use molecules containing electron donor centre on one side of the junction and an acceptor centre on the other side. The system therefore will be intrinsically asymmetric. More specifically, there must be a substantial difference in the energy of the highest occupied molecular orbital (HOMO) of the donor unit and the lowest unoccupied molecular orbital (LUMO) of the acceptor: the current resulting from resonant electron transfer from the LUMO of the acceptor to the HOMO of the donor would be greater for the opposite bias which results from the nonresonant transfer. The system asymmetry can be further boosted by different electronic coupling between the electrodes and the bridge molecule or (but not necessarily) by different nature/electronic structure of the two electrodes.

Quite a number of papers were devoted to realization of this principle or to reaching rectification by any means (see a review in ref 33). In many cases the effect has been observed, but the interpretation was not unambiguous. In ref 34, a molecular rod have been designed by synthesizing a molecule with two weakly coupled electronic $\pi$-systems with mutually shifted energy levels; this ‘rod’ bridges two gold electrodes. The introduced asymmetry provides rectification, whereas two alternative (control) bridges which do not possess such asymmetry do not.

A number of experiments have been focussed on reaching rectification in metal-insulator-metal (MIM) junctions, in which the role of the insulator was played by two self-assembled monolayers, each covering each electrode, which are then put in contact. Majda’s group worked systematically on measuring currents on two mercury drop electrodes functionalized with alkane thiolute layers, put in dense contact with each other (the use of mercury electrodes for such purpose was pioneered by Mann and Kuhn, who have built variable length metal-Langmuir-Blodgett-film-metal tunneling junctions). Although this system suffers from pure stability, as due to van der Waals attraction the mercury droplets - even when covered by SAMs - tend to fuse, particularly, in the case of short alkanethiolates, it is still an interesting model system for the study of electron transfer through molecular junctions. The SAMs here seem to be
too dense to expect to see any of the effects described in our work, had this systems been imbedded in electrolytic solution in one of the discussed configurations. No visible rectification has been reported for that system. However, in a similar kind of system, in which one of the electrode was mercury and the other one was silver or gold, but in which SAMs were very different [it was formed from dialkyl disulphide with a covalently linked tetracyanoquinodimethane group support by silver (or gold) and SAM formed of an alkanethiolate supported by mercury] a dramatic rectification has been observed.\textsuperscript{37} Such asymmetric system provided one order-of-magnitude strong rectification effect (at $\pm 1\text{V}$). Since the rectification was shown not to be affected by replacement of gold with silver, it was concluded that these were the properties of the molecule that determined the effect. Here local electron energy levels of the functional groups are clearly distinct from those of the methylene; furthermore the functional groups are asymmetrically coupled to the electrodes – they are covalently linked to Ag or Au electrode through the alkane chain, whereas they are in van der Waals contact with the SAM on Hg. Interesting observation is that the longer the alkane chain, the weaker the rectification! This is opposite to the effect of the molecular length in the electrostatic mechanism, as described in the previous sections. But there is no controversy here, because the mechanisms of rectification are drastically different: in ref 37 crucial for the rectification was the presence of the TCNQ group with low lying vacant molecular orbital. The dissimilarity of the molecular structure of the SAMs on the two electrodes alone does not result in rectification.

As pointed out in ref 38, the rectification effects in bridge junctions observed so far are mainly due to four factors: (i) the geometric contact between the molecular bridge and the electrodes; (ii) the coupling between the electrodes and the molecule; (iii) the position of the Fermi level; (iv) the electrostatic profile within the nanogap. In some reported studies on rectification,\textsuperscript{39,40,41} symmetric I-V curves are obtained with the use of symmetric bridging molecules and symmetric contact with electrodes, which is in agreement with ref 38 that rectification will not be observed if symmetry is obeyed in the system. Meanwhile with the use of asymmetric molecules,\textsuperscript{39,40,41} rectifications were then observed. In addition, Datta’s group\textsuperscript{42} has simulated rectification effects based on the asymmetric coupling between the electrodes and the molecule.

Throughout the present paper, we assumed the use of symmetric molecules and symmetric contacts with the electrodes. The rectification effect is resulted from different deformation of the potential barrier at positive and negative bias voltages. By fixing $\Phi_s$ and varying $\Phi_t$ (in most cases unless specified), the overall potential barrier experienced by the tunneling electrons under positive bias is different under the negative bias, than at positive bias, which comes out due to the Debye screening effect, and is thus an essential feature of the electrochemical (electrolytic) configuration of the system. This is an emerging direction of research in single molecule diodes, exploiting in particular the electrochemical junction with full, bipotentiostatic control (as shown in Figure 1) to achieve rectification by manipulating $\Phi_s$ and $\Phi_t$ independently.

As a highlight to section 6, a recent study by Capozzi et al.\textsuperscript{21} on rectification has been conducted under electrolytic environment. The group used electrodes made of same material but of different areas to provide the asymmetry for rectification. In section 6, we used a simplified model to investigate the I-V characteristics by varying the surface area ratio. Due to the difference of surface charge density residing
on the electrodes, rectification effects are observed for those with different areas in Figure 12. Therefore experiments can also be carried out, by having different sizes of electrodes, to achieve rectification even without the bipotentiostatic control.

We raised the issues of alternative mechanisms, to stress that the electrostatic mechanism of rectification is best to study in the systems that in quasi-MIM (or SAM-to-SAM) configurations will not exhibit any rectification! Otherwise other effects may complicate the results, such as e.g. a competing trend in the length dependence. Obviously, dense coverage of electrodes by the ‘bridge’ molecules will be ‘killing’ for the electrostatic mechanism, as in this case the electrolyte will not be able to screen electric field and modify the shape of the tunneling barrier.

Few other points to conclude the discussion.

The first one is the practicality of the electrolytic contacts in molecular electronics, for using these junctions as diodes. The tunneling time of the electron across the barrier, as the time interval between the injection and the release of the electron to the opposite electrode may depend on a number of factors, like the thickness and height of the barrier, the energy of tunneling electrode and the driving force – the electric field (for comprehensive analysis see, e.g. Landauer and Buttiker, see also a discussion in ref 33). But most typical estimates gives a short time of the order of $10^{-14}$s, so this stage is not determining for the operation of the electrostatic (electrolytic) molecular diodes. The bottleneck will be the switching time related to the speed of changing the barrier when switching the potential. It is generally determined by the time to recharge the electrical double layers via ion migration between them (relaxation of the double layers when the charge is delivered, i.e. accommodation of the extra charge, is a much faster process). Such migration time, also known as RC-time of charging a capacitor trough a resistor, can be estimated as

$$\tau_m \approx \left( L \cdot L_G \right)/D,$$

where $D$ is the diffusion coefficient of ions, $L_G$ is the characteristic thickness of the double layer – the Gouy length, and $L$ is the distance between the electrodes in the junction. Taking $L_G \sim 1nm$ and $L$ as large as $5nm$, and $D \sim 5 \times 10^{-6} cm^2/s$, we obtain $\tau_m \approx 10^{-8}$s. In bipotensiostatic mode, the estimate may be more complicated as it may depend on specific configuration of the electrodes; but experiments often show the time constant in the microsecond range $L$ should rather be the distance between the electrodes and the reference electrode. Further building a microfluidic chip based on the above discussed principles might allow to further shorten the time constant.

Not unrelated with the previous discussions, we should mention which other effects have been left behind our consideration in the main text of these article. If the ‘contact’ time is slow, one may need to incorporate the effects of inelastic tunneling due to interactions of the phonon modes in the bridge, electrodes, or the solvent, as studied by many authors – we note the systematic studies of Troisi in this direction, also papers by Medvedev and Datta’s group.

Somehow, the predictions on ‘electrostatic rectification’ effects that we analysed above should be tested with great care in terms of the choice of bridge molecules, types of contacts and electrodes, in order not to compromise them by competing effects. The goal of this article was to provide a road map for such investigations.
10. Conclusion

On top of the predictions of refs 20 and 19, this comprehensive analysis, which has covered a wider range of situations and parameters, brings several important conclusions.

- The account for nonlinear screening suggests that the rectification of the current across a single molecule in electrochemical environment may be seen more moderate (even small) electrolyte concentrations than it has been previously expected, based on the linear response theory.

- The rectification is based on the asymmetry of polarization of electrodes and can be seen in otherwise perfectly symmetric, in material terms, systems. However antisymmetric polarization of electrodes destroys rectification.

- In a two electrode system (symmetric in material terms), with electrodes connected through a battery, rectification can only be achieved if the two electrodes have substantially different surface areas.

The theory of the current-voltage curves discussed above all rests on the potential distribution in the nanogap, calculated so far without the account of the presence of the bridge molecule. We expect it not to be substantial to change the main qualitative predictions of the theory. However, this needs to be verified by molecular dynamics simulations. Still the primary task now will be to test the listed set of predictions experimentally, navigated by and equipped with all the recipes discussed above.
Appendix A: Potential distribution in the nanogap with account for compact layers and their effect on the current-voltage characteristic

To describe the potential distribution in the main text, we have used a simple version of the diffuse double layer theory corrected with account of the excluded volume of ions. However, for larger electrode polarizations, when the diffuse layer becomes too compact, it may need a well-known correction – introduction of the compact layer. Not exploring this question in all detail, we briefly consider the consequences that such extension of the theory may have on the rectification.

![Figure A.1. The geometry and nomenclature of a nano-gap with compact layers.](image)

The potential distribution \( \Phi(x) \) across the nanogap \( (d_{c,s} \leq x \leq L_{tot} - d_{c,t}) \) will follow the same set of equations (eqs 9-14 of the main text), but depend on \( \Phi_{c,s} \) and \( \Phi_{c,t} \) instead of \( \Phi_s \) and \( \Phi_t \). Meanwhile the potential distribution within the two compact layers (as in dielectric films) can easily be written down as -

within \( 0 \leq x \leq d_{c,s} \),

\[ \Phi(x) = \begin{cases} \Phi_{c,s} & \text{for } 0 \leq x \leq d_{c,s} \\ \varepsilon_{eff} & \text{for } d_{c,s} \leq x \leq L_{tot} - d_{c,t} \\ \Phi_{c,t} & \text{for } L_{tot} - d_{c,t} \leq x \leq L_{tot} \\ \end{cases} \]
\[ E_{\text{field}} = \frac{\Phi_s - \Phi_{c,s}}{d_{c,s}} \quad (A.1) \]
\[ \Phi(x) = \Phi_s - \frac{\Phi_s - \Phi_{c,s}}{d_{c,s}} x \quad (A.2) \]

within \( L_{tot} - d_{c,t} \leq x \leq L_{tot} \),

\[ E_{\text{field}} = \frac{\Phi_{c,t} - \Phi_t}{d_{c,t}} \quad (A.3) \]
\[ \Phi(x) = \Phi_t - \frac{\Phi_{c,t} - \Phi_t}{d_{c,t}} (x - L_{tot}) \quad (A.4) \]

As the Poisson Equation is a second order differential equation, hence only two boundary conditions are needed to define the problem. The values of \( \Phi_{c,s} \) and \( \Phi_{c,t} \) are not given, but are to be found via the boundary conditions at the surfaces of the electrodes,

\[ \varepsilon_{c,s} \frac{\Phi_s - \Phi_{c,s}}{d_{c,s}} = -\varepsilon_{eff} \frac{d\Phi}{dx} \bigg|_{x=d_{c,s}} \quad (A.5) \]
\[ \varepsilon_{c,t} \frac{\Phi_t - \Phi_{c,t}}{d_{c,t}} = -\varepsilon_{eff} \frac{d\Phi}{dx} \bigg|_{x=L_{tot}-d_{c,t}} \quad (A.6) \]
\[ L_{tot} = L + d_{c,t} + d_{c,t} \quad (A.7) \]

The derivatives in both eqs A.5 and A.6 depend on the signs of \( \Phi_{c,s} \) and \( \Phi_{c,t} \) as described in eqs 5 and 6 of the main text. However, these equations are not analytical, and could only be solved numerically to find the right values of \( \Phi_{c,s} \) and \( \Phi_{c,t} \).

After introduction of the compact layers, eq 17 of the main text must be rewritten as

\[ I = \sigma \frac{4U}{\beta_0 L_{tot}} \sinh \left( \frac{e\beta_0 L_{tot} (\Phi_s - \Phi_t)}{4U} \right) \exp \left[ \frac{e\beta_0 L_{tot}}{2U} \int_0^{L_{tot}} dx \left( \Phi - \frac{\Phi_s + \Phi_t}{2} \right) \right] \quad (A.8) \]
\[ \int_0^{L_{tot}} dx \Phi = \int_0^{d_{c,s}} dx \Phi + \int_{d_{c,s}}^{L_{tot}-d_{c,t}} dx \Phi + \int_{L_{tot}-d_{c,t}}^{L_{tot}} dx \Phi \quad (A.9) \]

The second integral on the R.H.S. of eq A.9 can be evaluated using eq 19 or eq 20 (of the main text) with the found \( \Phi_{c,s} \) and \( \Phi_{c,t} \) instead of \( \Phi_s \) and \( \Phi_t \). The first integral and the third integral on the R.H.S. of eq A.9, can be integrated using the corresponding potential functions,

\[ \int_0^{d_{c,s}} dx \Phi = \frac{d_{c,s}}{2} \left[ \Phi_s + \Phi_{c,s} \right] \quad \int_{L_{tot}-d_{c,t}}^{L_{tot}} dx \Phi = \frac{d_{c,t}}{2} \left[ \Phi_t + \Phi_{c,t} \right] \quad (A.10) \]
Adding up the three integrals together, defines the ‘algorithm’ for the calculation of the current-voltage characteristic.

In Figure A.2 & Figure A.3, we demonstrate the consequences of the introduction of the compact layer. In the considered examples, the dielectric constant of the compact layer was taken to be \((5\varepsilon_0)\) and is relatively smaller than in the middle of the gap which was taken to be \((80\varepsilon_0)\). It is hence obvious that the potential drop (and the electric field) in the diffuse layer part has been suppressed significantly both in Figure A.2 and Figure A.3. Had we taken the dielectric constant in the gap be comparable to that in the compact layer, the suppression of the potential drop across (and the electric field in) the double would be less strong. As shown in Figure 6, the nonlinear screening makes a difference for rectification when the concentration of ions is low. The same effect is seen here for high concentration of ions (Figure A.2), both the potential distribution and the I-V curve exhibit negligible difference between the linear model and the nonlinear model, while for low concentration of ions (Figure A.3) such difference is substantial.

![Figure A.2. Compact layer correction at a high concentration: nonlinear model coincides with linear model.](image)

Calculations via linear and nonlinear models for (a) potential distributions (b) I-V curves. Parameters: \(L_{tot} = 3\text{nm}, \varepsilon = 80\varepsilon_0, c_0 = 0.0941\text{M} \implies L_D = 1\text{nm}, \gamma = 0.00727\), \(\mu = 1, w = 0, U = 5eV, T = 298K, \varepsilon_{c,s} = 5\varepsilon_0, \varepsilon_{c,t} = 5\varepsilon_0, d_{c,s} = 0.5\text{nm}, d_{c,t} = 0.5\text{nm}\). The bias voltage is varied by changing \(\Phi_s\) at \(\Phi_s = 0V\).
Appendix B: Superexchange Model

Let us compare general expressions for current-voltage plots for the direct tunneling mechanism\textsuperscript{20} and the superexchange mechanism,\textsuperscript{19} to specify where the main differences lie. As everywhere in this paper, to make algebra simpler, we consider the two electrodes made of the same material.

First of all, we note that the theory for both direct and superexchange mechanisms of tunneling through bridge mediated junctions, as derived in refs \textsuperscript{20} and \textsuperscript{19}, was based on the idea of independent conformational fluctuations of the elementary units of the chain molecules, coupled with the tunneling process.

In direct tunneling, the mutual arrangement of the molecule was assumed to be not optional in the ground state, and the conformational fluctuations, each costing activation energy, were bringing that arrangement closer to or farther from the minimal barrier. Fluctuations to the optimal state where utilized by electrons for the most efficient tunneling. Each group was assumed to fluctuate independently, in a Gaussian manner, so that the overall probability of tunneling was proportional to the product of probabilities of conformational fluctuations of each group. As a result the overall tunneling probability had an Arrhenius shape, with an activation energy equal to the sum of activation energies of fluctuations of each individual group.

In superexchange, each group in the chain was assumed to provide a LUMO level that can be virtually exploited by the tunneling electron, but the position of that level ‘breaths’ with the conformational fluctuations of the groups; again, each of them was assumed to breathe independently, Gaussian-wise. Statistical averaging over such fluctuations leads to a completely different, non-Arrhenius temperature dependence of the current.

For direct tunneling, slightly rearranged expression derived in ref \textsuperscript{20} reads
\[ I = \frac{4\sigma U}{\beta_0 L} \sinh \left( \frac{e\beta_0 LV_{\text{bias}}}{4U} \right) \exp \left( -e\beta_0 L \frac{1}{4U}[\Phi_s + \Phi_t] \right) \exp \left( \frac{1}{2} \beta_0 L \frac{1}{UL} \int_0^L dx e^{\Phi(x)} \right) \]  \hspace{1cm} (B.1)

\[ \sigma \approx \sigma_0 \exp \left\{ -\beta_0 L - \frac{\beta_0 L}{U} \frac{\alpha}{1 + \beta_0 L} \frac{k_B T}{\hbar \omega B \left( \hbar \omega / k_B T \right)} \frac{Q^2}{2} \right\} \]  \hspace{1cm} (B.2)

Here some new notations appeared which we will need to define. \( \alpha \) is the coupling constant, of the dimensionality of energy, which characterizes the effect of harmonic conformational fluctuations on the tunneling barrier, \( \omega \) is the characteristic frequency of those fluctuations, \( B(y) = \frac{\tanh(y)}{y} \), and \( Q \) is the value of the dimensional harmonic coordinate which provides the minimum barrier for tunneling. \( L \) is the length of each independently fluctuating group of the chain. In the absence of coupling of tunneling with the conformational fluctuations, the tunneling factor would have been just \( \sigma \approx \sigma_0 \exp(-\beta L) \).

The I-V formula for the superexchange mechanism, as derived in ref 19, reads (some symbols here are renamed to tag along with the present paper),

\[ I = I_0 \sinh \left( \frac{LeV_{\text{bias}}}{\bar{U}} \right) \exp \left( -\frac{Le(\Phi_s + \Phi_t)}{\bar{U}} \right) \exp \left( \frac{2L}{\tilde{U}} \frac{1}{\tilde{U}L} \int_0^L dx e^{\Phi(x)} \right) \]  \hspace{1cm} (B.3)

\[ I_0 = \frac{2e}{\pi \hbar} \exp \left[ 2N \frac{E_r k_B T}{\bar{U}^2} \left( B \left( \hbar \omega / 2k_B T \right) \right)^{-1} \right] \Delta_s \Delta_t 2^{(N-1)} \frac{\Delta_s \Delta_t 2^{(N-1)}}{NU^{2N-1}} \]  \hspace{1cm} (B.4)

We will define here only those symbols which have not been defined in the main text. \( \bar{U} \) has a different meaning than \( U \), it is defined as the difference between the basic energy of the electronic state in the bridge molecule (the one not affected by fluctuations) and the Fermi-levels of the electrodes in the absence of potential bias. \( \Delta_s \) and \( \Delta_t \) are the coupling constants of the molecules with the substrate and the tip respectively; \( \Delta \) is the hopping overlap integral between the neighboring groups. \( E_r = \frac{1}{2} \hbar \omega g^2 \), where \( g \) is the coupling constant corresponding to the vibrational mode of frequency \( \omega \); \( N = \frac{L}{\ell} \) is the number of the individually fluctuating groups in the bridge molecules.

In the two mechanisms, the hyperbolic-sine term and the two following exponential terms have the same structure, but the parameters are different. In the superexchange formula (eq B.3), \( \frac{Le}{\bar{U}L} \) replaces \( \frac{\beta_0 L}{4U} \) of the direct tunneling formula (eq B.1). The main difference lies in the form of the electrostatic-potential independent factors, \( I_0 \) and \( \sigma \), which have entirely different dependence on temperature. Therefore, temperature is the key parameter to differentiate which mechanism dominates the transport of electrons in a particular experiment.
Biographical details

**K. C. Matthew Cheung** is an undergraduate student at the Department of Physics of Imperial College of Science, Technology and Medicine, London. He performed this research at the Department of Chemistry of Imperial College within the framework of the Undergraduate Research Opportunities Programme (UROP), supported by the Centre for Doctoral Training on Theory and Simulation of Materials.

**Xiaoyu Chen** received her Bsc in Chemistry from Imperial College London in 2015. She did her Bsc project on current rectification in single molecule electrochemical junctions with Professor Alexei Kornyshev. She is now doing her Msc in Molecular Modelling at University College London.
Tim Albrecht received his PhD in Chemistry from the Max-Planck Institute of Bioinorganic Chemistry and the Technical University of Berlin 2003 and then moved to Copenhagen to work as Postdoctoral Researcher (Research Assistant Professor) with Prof. Jens Ulstrup on single-molecule charge transport in electrochemical environments. He accepted a Lecturer position in Interfacial Sciences at Imperial College London in 2006, was promoted to Senior Lecturer in 2011 and then to Reader in 2014. He has published 70 research articles, co-authored several book chapters and edited two special issues in the J. Phys. Cond. Mat. on single-molecule electronics and nanopore sensing, respectively. He is Chair of the RSC Electrochemistry Group, their representative at the International Society of Electrochemistry (ISE), as well as the UK Regional Representative to the ISE. He has strong interactions with industry, for example with Agilent (now Keysight), Nanomeasurements Division, Rio Tinto, Hitachi and Siemens. His main research interests focus on nanoscale charge transport and charged interfaces, including nanopore sensing and nanopore/electrode structures as well as single-molecule electronics. In 2014 he received the ISE Tajima Prize.

Alexei Kornyshev is an expert in condensed matter theoretical chemical physics and its applications to biophysics, electrochemistry, nanoscience, and energy generation and storage. Through the years his particular areas of research were electron and proton transfer reactions and processes, hydration, structure and electrical properties of metal/electrolyte and liquid-liquid interfaces, transport phenomena in solid electrolytes and polymer electrolyte membranes, aggregation and recognition of helical biopolymers and the structure of chiral liquid crystals, theory of modern fuel cells and super-capacitors, electrowetting. He has published > 220 original papers and >30 review articles, and several patents, edited 5 multi-author books and a number of special journal issues. He is on Editorial Boards of IOP’s J. Phys. Cond. Matter, Wiley’s ChemElectroChem, and NPG’s Scientific Reports. His current research interests focus on (i) DNA biophysics, (ii) optofluidics and nanophotonics, and (iii) ionic liquids and their applications in energy research and nanomechanics, (iv) molecular electronics and machines.

He received his Master’s degree from Moscow Physical Engineering Institute in theoretical nuclear physics in 1970. He did his Ph.D. in 1974 in physical and mathematical sciences (with R. R. Dogonadze), and then worked as a staff member of the Frumkin Institute of Electrochemistry (Acad. Sci.), Moscow (1973-1991); in 1986 he received D.Sc. degree in Chemical Sciences. In 1991-1992 as a Humboldt Prize Awardee he spend a year on leave at TU Munich, and was then invited to the Research Center Juelich, to lead a theory group in a new Institute of Materials and Processes in Energy Systems, focusing on fundamentals of fuel cells. In 1997 he was promoted to Head of Division of Theoretical Physical Chemistry of that Institute, receiving later a joint appointment with Heinrich-Heine University of Dusseldorf as Professor of Theoretical Physics. In 2002 he joined Imperial College London as a Chair of Chemical Physics supported by 2001 Royal Society Wolfson Research Merit Award, where he works since then. He is a fellow of IUPAC, Institute of Physics, International Society of Electrochemistry, Royal Society of Chemistry, Royal Society of Biology, and a foreign member of Royal Danish Academy of Science. His was a recipient of 2003 Ch.-F.Schoenbein Contribution-to-Science Medal of the European Fuel Cell Forum, 2006 Geoffrey Barker Electrochemistry Medal and 2010 RSC Interdisciplinary Prize, Medal and Lectureship of the Royal Society of Chemistry.