Supplementary Figure 1: Extinction cross section for a nanowire dimer with a sub-nanometer gap taking the nonlocal hydrodynamic-diffusion response into account. The radius of the sodium wires is $R = 25 \text{ nm}$ with the gap $g$ varying from $-19 \text{ Å}$ to $+19 \text{ Å}$. Calculations are for the same parameters as in Figure 3 in the main text. For comparison, the solid black lines show the nonlocal response in the absence of diffusion ($\beta \neq 0$ and $D = 0$).
Supplementary Figure 2: Circuit model for diffusion-tunneling dynamics in the dimer gap region.
Supplementary Table 1: Table of central length scales and parameters for Au, Ag, and Na. The entries for different metals are based on Fermi wavelengths ($\lambda_F$), Fermi velocities ($v_F$), and plasma frequencies ($\omega_p$) taken from Ref. 1, while the values for $\tau$ originate from various references as indicated in the right-most column of the table.
Supplementary Note 1

In this section we offer the detailed derivation of the hydrodynamic-diffusion problem of an electron responding to an external electric field. We start from the linearized hydrodynamic equation-of-motion for an electron in an electric field:

\[ \partial_t \mathbf{v} = -\gamma \mathbf{v} + \frac{(-e)}{m} \mathbf{E} - \frac{\beta^2}{n_0} \nabla n_1. \]  

(1)

Here, \( \mathbf{v} \) is the velocity correction to the static sea of electrons and for the density \( n(r, t) = n_0 + n_1(r, t) \) the latter term is likewise the small \( (n_1/n_0 \ll 1) \) induced density variation associated with the driving \( \mathbf{E} \) field. Now, rather than the usual simple form for the continuity equation we instead have the linearized convection-diffusion equation:

\[ \partial_t (-e)n_1 = D \nabla^2 (-e)n_1 - \nabla \cdot \{ (-e)n_0 \mathbf{v} \} = -\nabla \cdot \mathbf{J} \]  

(2)

where the current density is then given by Fick’s law:

\[ \mathbf{J} = (-e)n_0 \mathbf{v} - D \nabla(-e)n_1. \]  

(3)

Multiplying the equation-of-motion, Supplementary Equation (1), by the equilibrium density \( n_0 \), taking the time-derivative \( \partial_t \) and re-arranging we get:

\[ (\partial_t + \gamma) \partial_t \{ (-e)n_0 \mathbf{v} \} = \frac{n_0 e^2}{m} \partial_t \mathbf{E} - \beta^2 \nabla \{ \partial_t (-e)n_1 \}. \]  

(4)

Using Fick’s law, Supplementary Equation (3), we then find:

\[ (\partial_t + \gamma) [\partial_t \mathbf{J} + D \nabla (\partial_t (-e)n_1)] = \frac{n_0 e^2}{m} \partial_t \mathbf{E} - \beta^2 \nabla \{ \partial_t (-e)n_1 \}. \]  

(5)
From the diffusion-convection equation, Supplementary Equation (2), we consequently get

\[(\partial_t + \gamma) [\partial_t \mathbf{J} - D \nabla (\nabla \cdot \mathbf{J})] = \frac{n_0 e^2}{m} \partial_t \mathbf{E} + \beta^2 \nabla (\nabla \cdot \mathbf{J}). \tag{6}\]

Finally, Fourier transforming with respect to time and re-arranging the terms we get

\[
\left\{ \frac{\beta^2}{\omega (\omega + i\gamma)} + \frac{D}{i\omega} \right\} \nabla (\nabla \cdot \mathbf{J}) + \mathbf{J} = \frac{\varepsilon^2 n_0}{m \sigma_D} \mathbf{E}. \tag{7}\]

This is the generalized constitutive equation with a complex nonlocal correction to the local-response Ohm’s law. According to our previous work\(^7\), the Maxwell’s equation can then be re-written in the form of the equations in the main text, i.e. Equation (2) with \(\xi^2\) given by Equation (6).

In our use of Supplementary Equation (7) we neglect electron spill-out and the associated boundary conditions used in prior nonlocal hydrodynamics\(^6\)–\(^9\) remain unchanged in the presence of diffusion. This means that \(n \cdot J = 0\) on the metal surfaces, implying that no electrons escape the metal volumes. This is an appropriate description of noble metals commonly employed in plasmonics while spill-out effects are important in less common metals like sodium\(^3\),\(^10\),\(^11\).

The local-response approximation (LRA) neglects the Laplacian term in Supplementary Equation (7) and in problems using such a constitutive equation the induced charge \(\Delta n\) becomes a delta function at the surface of the metal. The nonlocal correction will smear this density profile due to both convection and diffusion. Related phenomena and formalisms are known from semiconductor drift-diffusion theory\(^12\) as well as from fluid mechanics and chemical engineer-
ing, where convection-diffusion systems may exhibit an increased effective diffusivity of chemical species known as Taylor dispersion\textsuperscript{13,14}.

In Supplementary Table 1 we summarize characteristic length scales and parameters for noble metals commonly considered in experimental realizations as well as in recent theoretical works.

**Supplementary Note 2**

For the dimer we here address the relative importance of diffusive damping (characterized by a resistance $R_{\text{dif}}$ in a circuited model) and the damping associated with the relaxation of a possible quantum tunneling current (characterized by $R_{\text{tun}}$) short-circuiting the classically forbidden capacitive gap (characterized by a capacitance $C$). In a simple picture, $C$ and $R_{\text{tun}}$ constitute a parallel circuit\textsuperscript{15} connected in series with $R_{\text{dif}}$, see Supplementary Figure 2. The circuit impedance is then

$$Z = R_{\text{dif}} + \frac{R_{\text{tun}}}{1 + i\omega\tau_{\text{tun}}} = R_{\text{dif}} - \frac{i}{\omega C} + O[1/(\omega\tau_{\text{tun}})^2]$$

(8)

where $\tau_{\text{tun}} = R_{\text{tun}}C$ is the tunneling RC time\textsuperscript{15}. The tunneling dynamics simplifies in the slow adiabatic-following regime and the limit of fast external driving\textsuperscript{16}. This analysis suggests that the high-frequency dimer dynamics can become entirely dominated by the diffusive broadening and the junction capacitance. Tunneling dynamics has been explored in the context of the mesoscopic capacitance\textsuperscript{17} and ultra-fast tunneling experiments have reported tunneling RC times in the picosecond range\textsuperscript{15,18}. Thus, at optical frequencies the plasmon response is expected to be fast on the scale of the characteristic RC time and from the experiments it seems reasonable to assume that $\omega\tau_{\text{tun}} \gg 1$. Consequently, the relaxation is dominated by diffusive broadening rather than the
short-circuiting tunneling current.

We note that the above circuit analysis is completely independent on whether tunneling relaxation occurs within the gap (as it is the case in the quantum-corrected model\textsuperscript{19}) or if it takes place inside the metal surfaces (in agreement with the common understanding of relaxation within mesoscopic quantum electron transport).

In Supplementary Figure 1 we show the extinction for dimers of $R = 25\, \text{nm}$ wires and with a gap $g$ varying from $-19\, \text{Å}$ to $+19\, \text{Å}$. While the wires themselves have a too large radius ($R \gg |\xi|$) to support nonlocal effects, the small gap ($g \sim |\xi|$) causes a strong broadening associated with the complex-valued nonlocal response.

**Supplementary Note 3**

We consider a metal sphere of radius $R$ embedded in a homogeneous background dielectric environment with permittivity $\varepsilon_b$. In the quasi-static limit, the optical response of the sphere is described by the dipole polarizability $\alpha$ that features a nonlocal generalization of the Clausius–Mossotti factor, and is given as\textsuperscript{20,21}

$$\alpha = 4\pi R^3 \frac{\varepsilon_D - \varepsilon_b (1 + \delta_{nl})}{\varepsilon_D + 2\varepsilon_b (1 + \delta_{nl})}, \quad \delta_{nl} = \frac{\varepsilon_D - \varepsilon_{\infty}}{\varepsilon_{\infty}} \frac{j_1(k_{nl}R)}{k_{nl}R j_1'(k_{nl}R)},$$

where $\varepsilon_D = \varepsilon_{\infty} - \omega_p^2/(\omega^2 + i\gamma\omega)$ is the Drude dielectric function, $k_{nl}^2 = (\omega^2 + i\omega\gamma - \omega_p^2)/(\beta^2 + D\gamma - iD\omega)$ is the wave vector of the longitudinal wave, and $j_1$ is the spherical Bessel function of first order.
Supplementary References


