Chemical transferability of functional groups follows from the nearsightedness of electronic matter

Stijn Fias\textsuperscript{a,b,1}, Farnaz Heidar-Zadeh\textsuperscript{b}, Paul Geerlings\textsuperscript{a}, and Paul W. Ayers\textsuperscript{b}

\textsuperscript{a}General Chemistry, Free University of Brussels, B1050 Brussels, Belgium; and \textsuperscript{b}Department of Chemistry & Chemical Biology, McMaster University, Hamilton, ON, Canada L8S 4L8

Edited by John P. Perdew, Temple University, Philadelphia, PA, and approved August 28, 2017 (received for review September 8, 2016)

We establish the physical origins of chemical transferability from the perspective of the nearsightedness of electronic matter. To do this, we explicitly evaluate the response of electron density to a change in the system, at constant chemical potential, by computing the softness kernel, $s(r, r')$. The softness kernel is nearsighted, indicating that under constant-chemical-potential conditions like dilute solutions changing the composition of the molecule at $r$ has only local effects and does not have any significant impact on the reactivity at positions $r'$ far away from point $r$. This locality principle elucidates the transferability of functional groups in chemistry.

nearsightedness | electronic matter | linear response | chemical transferability | density functional theory

One of the fundamental precepts of chemistry is that molecules are built from atoms and functional groups, which largely retain their chemical identity regardless of their molecular environment. This implies that the chemical properties of functional groups are largely transferable between molecules (1). As a corollary, chemical reactivity is nearsighted: Adding or changing a functional group in a molecule rarely affects its reactivity far away from where it was modified. The physical origins of chemical transferability remain obscure but are clearly related to Kohn’s nearsightedness of electronic matter (NEM) principle. NEM has been used to explain the transferability of functional groups along with other concepts like Pauling’s concept of the chemical bond and the “divide and conquer” approach used in the linear scaling strategy to study large molecular systems (2–10). This principle states that for many-electron systems at constant electronic chemical potential the amount of change in electron density at $r_0$, $|\Delta \rho(r_0)|$, induced by a perturbation in the external potential at position $r'$, $\Delta \nu(r')$, with $r'$ outside a sphere with radius $R$ centered on $r_0$, will always be smaller than a maximum value $\Delta \rho(r_0, R)$, no matter how large the perturbation. In other words, the electron density at $r_0$ cannot “see” any perturbation beyond $R$ with an accuracy greater than $\Delta \rho(r_0, R)$. Prodan and Kohn (8) showed that for 1D gapless model systems the decay of $\Delta \rho$ as a function of $R$ (i.e., upon increasing $|r - r_0|$) follows a power law. For “gapped” systems, that is, with a hardness $\eta$ greater than zero, this decay is exponential. This suggests that in the molecular world, where $\eta$ is always greater than zero, the electron density should only be sensitive to “nearby” changes in the external potential. One aim of this paper is to provide the first numerical confirmation of this “nearsightedness principle” for real 3D molecules. In addition, this paper aims at establishing the link between the basic physics encoded in the nearsightedness principle and the empirically observed transferability of functional groups, which is the basis for how chemists conceive organic and inorganic molecules as being built from fragments that largely retain their identity even after being combined to form a larger molecule. If a molecule is only sensitive to nearby changes in external potential, changing a functional group will only affect the reactivity of the molecule in the vicinity of the modification. Conversely, when adding/changing substituents on the molecular backbone, only the changes that are close to the functional group influence the reactivity of the functional group. The plausible link between “physical” and “chemical” nearsightedness was pointed out in 2008 by Richard Bader (10). Cardenas et al. (11) followed Prodan and Kohn stating that in an open-system representation (at constant electronic chemical potential $\mu$) reactivity indicators that depend on more than one point, including the hardness kernel and Kohn–Sham dielectric response function, are nearsighted.

As mentioned, NEM is stated to be the fundamental principle underlying important concepts in (quantum) chemistry like the transferability of functional groups, Pauling’s concept of the chemical bond, linear scaling algorithms (3, 4, 6, 7), and the divide and conquer approach (12), but it can also be understood as the concept underlying the success of machine learning when applied to chemical problems. As such, demonstrating the principle for molecules is an important step in understanding the electronic structure of molecular systems.

The Softness Kernel

The nearsightedness principle is fundamentally a statement about the change in density, $\Delta \rho(r_0)$, in response to a change in external potential, $\Delta \nu(r')$. This suggests that nearsightedness should be related to the linear response function $\chi(r, r') = \left(\frac{\delta \rho(r)}{\delta \nu(r')}\right)_N$ (13, 14). However, the nearsightedness principle only holds under constant-chemical-potential conditions, so the quantity that must be studied is $\left(\frac{\delta \rho(r)}{\delta \nu(r')}\right)_\mu$, which is minus the softness kernel, $-s(r, r')$, and is related to the linear response function by the identity (15)

$$\left(\frac{\delta \rho(r)}{\delta \nu(r')}\right)_N = \left(\frac{\delta \rho(r)}{\delta \nu(r')}\right)_\mu + \left(\frac{\delta \rho(r)}{\delta \mu}\right)_V \left(\frac{\delta \mu}{\delta \nu(r')}\right)_N. \quad [1]$$

Substituting the definitions of linear response function, softness kernel (15), local softness (16), and Fukui function (17) gives

$$\chi(r, r') = -s(r, r') + s(r)f(r'). \quad [2]$$

Significance

The empirically observed transferability of functional groups is shown to follow from Kohn’s nearsightedness principle for electron density. To show this, we analytically calculated and analyzed the softness kernel for molecules. The relevance of this work touches on physics (the nearsightedness principle), molecular chemistry and materials science (chemical transferability and similarity), and computer science (machine-learning approaches for pharmaceutical and materials design).

Author contributions: S.F., P.G., and P.W.A. designed research; S.F. performed research; S.F. and F.H.-Z. contributed new reagents/analytic tools; S.F., F.H.-Z., and P.W.A. analyzed data; and S.F., P.G., and P.W.A. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission. Published under the PNAS license.

1To whom correspondence should be addressed. Email: sfias@vub.ac.be.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1615053114/-/DCSupplemental.
Using the relationship between the local softness and the Fukui function, \( s(r) = S f(r) \), and the relationship between the global softness and the chemical hardness, \( \eta = 1/S \) (16, 18), one obtains a somewhat more traditional form of the Berkowitz–Parr relation,

\[
s(r, r') = -\chi(r, r') + \frac{f(r)f(r')}{\eta}.
\]  

[3]

This is the main working equation in the present paper. In a recent contribution, Yang et al. (19) derived analytical expressions for the linear response function, Fukui function and hardness, based on coupled-perturbed Kohn–Sham theory. These expressions were implemented and combined to calculate the softness kernel using Eq. 3.

Keeping the chemical potential, instead of the number of electrons, constant might seem strange at first, but it is often a more realistic perspective when investigating the reactivity of a molecule. In practice, molecules are not isolated but instead are found in solution, where the chemical potential is fixed by the solvent, allowing (partial) charge transfer to or from the molecule while keeping its chemical potential constant. Moreover, the nearsightedness principle does not hold under constant-N conditions, as is easily seen from a simple counterexample: If one adds a bare helium nucleus, \( \Delta v(r') = -2/r' \), far away from a hydrogen atom placed at \( r_0 \), then the electron density at \( r_0 \) changes dramatically no matter how far away the helium nucleus is.

**Results and Discussion**

**The Atom-condensed Softness Kernel.** The analysis of the six dimensional softness kernel is facilitated by dividing it into atomic contributions (20). To do this, we integrate \( r \) and \( r' \) of the kernel over the domains of atoms \( A \) and \( B \), represented by \( V_A \) and \( V_B \), thereby defining an atom–atom condensed softness kernel:

\[
s_{AB} = \int_{V_A} \int_{V_B} drdr' s(r, r') .
\]  

[4]

In this contribution, the molecule is partitioned into (fuzzy) atoms using the iterative Hirshfeld method (21, 22).

**1,3,5-Hexatriene.** Fig. 1 shows the atom–atom condensed linear response and softness kernel between the carbon atoms of 1,3,5-hexatriene. In the case of polyenes the atom-condensed linear response function is known to show an alternating behavior with maxima between mesomeric active atoms (\( C_1 - C_6 \)) and minima between mesomeric passive atoms (\( \chi_{C_1C_2}, \chi_{C_2C_3}, \) and \( \chi_{C_1C_6} \)) and \( \chi_{C_2C_4} \)).

![Fig. 1. Atom-condensed linear response and softness kernel (\( s_{AB} \)) of 1,3,5-hexatriene. The curves of the softness kernel using \( f^+ \) and \( f^- \) are overlapping.](image)

In the following examples, these alchemical changes have been computed to probe the effect of replacing one functional group by another while keeping the electronic chemical potential constant. The chemical potential was chosen as the average of the energies of the highest-occupied molecular orbital (HOMO) and lowest-unoccupied molecular orbital (LUMO) (corresponding to \( f^0 \)), to keep the chemical potential at the Fermi level of the molecule. We verified that nearly identical results are obtained when the chemical potential is set to the HOMO or LUMO energy (corresponding to \( f^+ \) or \( f^- \), respectively).

**Functionalization of 1,3,5-Hexatriene.** To study the effect of changing functional groups, the alchemical change in density of 1,3,5-heptatriene to 1,3,5-heptatrien-1-amine, 1,3,5-heptatrien-1-ol, and 1-fluoro-1,3,5-heptatriene at constant chemical potential was calculated using Eq. 5. The \( CH_3 \) group was annihilated and replaced by a \( NH_3 \), \( OH \), or \( F \) group, respectively. The changes in the electron population of the carbon atoms in hexatriene, evaluated using the iterative Hirshfeld partitioning, are shown in Fig. 3 (corresponding to minus the change in the charge of the atoms). When changing the methyl group to an amino group, only the charge on the first carbon atom (the carbon connected
Figs. S1 r 5 and vol. 114 11635

Fig. 4 shows the change in density when sub-
in this saturated system. 

Alchemical change in electron population using the linear response

\[ \chi(\mathbf{r}, \mathbf{r}') \]

October 31, 2017

S2

Atom-condensed linear response and softness kernel (\( f^* \) and 1.14 electrons with \( f^0 \)) is located on these carbon atoms.

Changing the methyl group to a hydroxyl group or fluorine atom at constant number of electrons withdraws electrons from hexatriene, making C6 more positively charged. At constant chemical potential, the molecule accepts charge from the electron reservoir to compensate for this electron withdrawal (0.12 electrons for OH and 0.23 electrons for F). These results prove that changing the functional group at constant chemical potential only affects the atoms in the vicinity of the functional group. This observation, ubiquitous in organic chemistry, is the basis for the concept of the transferability of functional groups.

**Heptatrienyl Cation.** Fig. 4 shows the change in density when substituting one of the terminal carbon atoms in the heptatrienyl cation with a nitrogen atom to form 1,3,5-hexatrien-1-amine. (In the linear response case, this is an especially extreme change, resulting in the 1,3,5-hexatrien-1-amine dication.) The figure shows that, at constant number of electrons, the nitrogen withdraws electrons from the whole molecule, with an important depletion of the density at the opposite side of the molecule on carbon atoms C5 and C6 when using the softness kernel, even though the value of \( s(\mathbf{r}, \mathbf{r}') \) is essentially zero when \( r \) is located on the methyl group and \( r' \) is located on these carbon atoms.

Changing the methyl group to a hydroxyl group or fluorine atom at constant number of electrons withdraws electrons from hexatriene, making C6 more positively charged. At constant chemical potential, the molecule accepts charge from the electron reservoir to compensate for this electron withdrawal (0.12 electrons for OH and 0.23 electrons for F). These results prove that changing the functional group at constant chemical potential only affects the atoms in the vicinity of the functional group. This observation, ubiquitous in organic chemistry, is the basis for the concept of the transferability of functional groups.

**Heptane.** In saturated systems, \( \chi(\mathbf{r}, \mathbf{r}') \) is already nearsighted due to the lack of a conjugated \( r \) system (14, 23). For the softness kernel to be likewise nearsighted, the second term in Eq. 3 should not counteract the nearsightedness of \( \chi(\mathbf{r}, \mathbf{r}') \). To test this, we studied the alchemical change of the terminal methyl group to a fluorine atom in heptane, forming 1-fluorohexane. As shown in Fig. 5, the change in density evaluated by the linear response and the softness kernel are almost identical, with the latter being slightly more localized. This establishes that the second term of \( s(\mathbf{r}, \mathbf{r}') \) in Eq. 3 does not counteract the nearsightedness of \( \chi(\mathbf{r}, \mathbf{r}') \) in this saturated system.

**Three-Dimensional Systems.** The examples considered until now are all linear hydrocarbons. To examine the nearsightedness in three dimensional structures, the alchemical change of methyl-cubane to fluoro-cubane was studied. The results are shown in Fig. 6. The changes in density evaluated by the linear response and the softness kernel are nearly identical and resemble the changes in density obtained for heptane. Numerically the change in density decays slightly faster at constant chemical potential than at constant number of electrons. These data, along with data for functionalized neopentane (where similar trends were observed), are included as Figs. S1 and S2. Our conclusions
Fig. 4. Alchemical change in density using the linear response (Top) and softness kernel (Bottom) for the heptatrienyl cation to 1,3,5-hexatrien-1-amine. The left carbon atom is changed. Yellow (purple) indicates regions of density accumulation; red (light blue) indicates regions of density depletion.

about the NEM and chemical transferability also hold, therefore, for 3D systems.

Metals. For gapless systems, $\Delta \rho$ has an asymptotic decay as a function of $R$ that is inverse-polynomial, in contrast to gapped systems, which have an exponential decay (8). To assess this computationally, we considered (equidistant) linear hydrogen chains, ($H_2)_n$, ($n = 5, 6, \ldots, 10$), as a model for this type of system, since they approach a conductor in the limit of infinitely large chains. For these chains, the data of which are included in Alchemical Changes, the condensed softness kernel is much more localized than the condensed linear response kernel, demonstrating that the NEM principle still holds when going toward a zero band gap.

The Applicability of Softness Kernel to Alchemical Changes. In the previous section, the change in the electron density upon replacing one functional group by another was studied at constant number of electrons and constant chemical potential using the linear response function and softness kernel, respectively. Since the linear response function and softness kernel only provide first-order estimates of the change in the electron density when changing the external potential, this raises the question of whether higher-order contributions to the density are likewise nearsighted. For these chains, the data of which are included in Alchemical Changes, the condensed softness kernel is much more localized than the condensed linear response kernel, demonstrating that the NEM principle still holds when going toward a zero band gap.

Implications. The NEM was introduced by W. Kohn as the theoretical basis underlying the linear-scaling algorithms proposed in density functional theory. In these algorithms, the molecular system is divided into overlapping subsystems comprised of a group of atoms for which the Kohn–Sham equations are solved separately. Since each region is treated independently, this approach scales linearly with system size and can be easily parallelized. This insight justifies algorithms such as the fast multipole method (3), density-matrix-based linear scaling methods (4), the natural linear scaling coupled-cluster method (6, 7), and the divide-and-conquer method (12). The latter explicitly uses a common

Fig. 5. Alchemical change in density using the linear response (Top) and softness kernel (Bottom) for heptane to 1-fluorohexane. The left carbon atom is changed (the annihilated H atoms are shown as reference). Yellow (purple) indicates regions of density accumulation; red (light blue) indicates regions of density depletion.

with an important depletion of density on the carbon atom far away from where the substitution took place. It seems that the first-order estimate by the linear response provides a plausible qualitative picture of the change in the density upon alchemical substitution. To determine the change in density at constant chemical potential, the number of electrons in 1,3,5-hexatrien-1-amine was increased in steps of 0.05 until the orbital energy of the partially filled orbital exceeded that of the LUMO in the heptatrienyl cation. Interpolation led to an estimate of 1.02 electrons being transferred from the reservoir to the molecule. This slightly differs from the value of 1.12 electrons found using the softness kernel. The difference of 0.1 electrons originates from the difference between the first-order estimate and the full self-consistent field calculation, showing that the first-order estimate recovers 90.2% of the change in number of electrons. When comparing the difference in electron density between the heptatrienyl cation and the 1,3,5-hexatrien-1-amine “+0.98” cation a picture similar to that of the softness kernel is obtained (cf. Fig. 7). Virtually no change in density can be observed on C6 and C7, illustrating again the nearsightedness of the change in density at constant chemical potential.

Implications. The NEM was introduced by W. Kohn as the theoretical basis underlying the linear-scaling algorithms proposed in density functional theory. In these algorithms, the molecular system is divided into overlapping subsystems comprised of a group of atoms for which the Kohn–Sham equations are solved separately. Since each region is treated independently, this approach scales linearly with system size and can be easily parallelized. This insight justifies algorithms such as the fast multipole method (3), density-matrix-based linear scaling methods (4), the natural linear scaling coupled-cluster method (6, 7), and the divide-and-conquer method (12). The latter explicitly uses a common
CHEMISTRY

Fig. 6. Alchemical change in density using the linear response (Top) and softness kernel (Bottom) for methylcubane to fluorocubane (the annihilated H atoms are shown as reference). Yellow (purple) indicates regions of density accumulation; red (light blue) indicates regions of density depletion.

chemical potential to control the electron population and to allow fractional subsystem charges. The results presented in this paper provide numerical corroboration that the constant chemical potential in the divide-and-conquer method has, as a desirable side effect, the removal of possible nonlocal interactions between subsystems. This ensures that the molecular systems can be divided into locally interacting subgroups, allowing one to achieve linear computational scaling.

This paper also provides direct computational evidence for the transferability of functional groups: Under the constant-chemical-potential conditions that are typical for molecules in dilute solutions or on molecular surfaces, the effect of changing a molecular substituent decreases rapidly with distance and is essentially negligible for atoms that are separated by more than three chemical bonds from the site of functionalization. This quick decrease of the influence of functionalizing the molecular backbone by modifying a substituent-functional group supports chemists’ intuition that functional groups “operate” locally: They largely retain their identity and therefore act similarly independent of the backbone (i.e., functional groups are transferable entities). Although this paper studies the effect on the backbone when changing a functional group, the nearsightedness principle also holds in the reverse sense. Changes to the molecular backbone will only affect the functional group in a nonnegligible way if they occur in the vicinity of the group. This effect is well known in organic chemistry. Consider, for example, the evolution in the pK$_a$ on 2-, 3-, and 4-chlorobutanoic acid where, compared with the unsubstituted acid, the pK$_a$ decreases by 2.0, 0.7, and 0.3 units, respectively. The relevant functional group property (acidity) is retained throughout, but the strength of the effect, however, diminishes exponentially after the nearsighted region. The nearsightedness implies that molecular systems can be divided in subgroups, whose reactivity is affected only mildly by the other subgroups of the system. This “chemical” nearsightedness principle enables the chemical divide-and-conquer perspective of large molecules as assemblages of chemically transferable functional groups. This ensures that similar molecular subgroups have similar functionality and justifies the use of molecular similarity measures. This not only provides the theoretical basis for applying modern machine-learning approaches for chemical property prediction but also for traditional quantitative structure/activity/property relationships (27–29). For example, it is popular to define a “distance” between molecules solely based on the identity and position of their constituent atoms [e.g., the Coulomb matrix measure (30)]. One can then use a machine-learning method like kernel-ridge regression to make a prediction for the properties of an unknown molecule, based on the idea that its properties will most closely resemble those of the structurally similar molecules.

Conclusions

The sensitivity of the electron density to changes in molecular structure, at constant chemical potential, was studied using the softness kernel. To do this, we first developed an analytical expression for the softness kernel based on the Berkowitz–Parr relation and then computed this quantity for a representative sequence of saturated and unsaturated organic molecules. The atom–atom condensed softness kernel shows that the effects of changing the external potential is negligible after about three bonds. The corresponding change in electron density at constant number of electrons is defined by the linear response function, where the effects of changing the external potential decay slowly and can be important even six or seven bonds away. By considering alchemical changes—where one functional group was changed to another at fixed chemical potential—the nearsightedness of the softness kernel has been confirmed numerically on molecules. This provides computational evidence for the...
transferability of functional groups: The molecular systems can be divided into locally interacting subgroups which retain a similar functionality and reactivity, which can only be influenced by changes in the direct environment of the functional group. This justifies what is colloquially known as the methyl-ethyl-propyl-buty-futile rule: Changing a molecular substituent more than four carbons away does not lead to significant chemical effects.

Materials and Methods
The softness kernel was calculated using Eq. 3 together with the equations given in ref. 19. The geometries were optimized at PBE/cc-pVDZ using Gaussian 09 (31). The Kohn–Sham orbitals were obtained from the same software package at the same level of theory, and the values of the linear response function, Fukui function, and hardness were analytically calculated using our in-house program developed for this purpose. The necessary exchange-correlation functionals were also obtained from the Libxc library (32). Both the atoms-in-molecules partitioning and the calculations with fractional number of electrons were performed using the open source HORTON package (33). When performing alchemical changes, the functional groups where “grown” at the position of their corresponding optimized coordinates.

ACKNOWLEDGMENTS. S.F. thanks the Research Foundation Flanders for financially supporting his postdoctoral research at the General Chemistry (ALGC) group. This work was supported by grants from the Natural Sciences and Engineering Research Council and Compute Canada (to P.W.A. and F.H.-Z.) and the Free University of Brussels through a Strategic Research Program grant to the ALGC group (to P.G.).