Bull. Chem. Soc. Jpn., 70, 1307—1310 (1997)

Superdelocalizabilities of Linear Chains

Makoto Isihara

Department of General Education, Shotoku Gakuen Women's Junior College, Nakauzura 1-38, Gifu 500

(Received November 19, 1996)

This paper analyzes superdelocalizabilities (SDs) of even-membered linear chains in a simple noncomputational way. Two basic relationships connecting SDs to bond orders are prepared in general form and applied to linear chains to provide a plain analytic expression to SDs. Behaviors of SDs in long linear chains are examined by using this analytic expression. Simple estimation formulas are derived for SDs near the end and in the middle of a long linear chain and the distinctive reactivity of end sites is fully verified. Comparisons with behaviors of frontier electron densities and frontier SDs manifest that these approximate indices cannot substitute for SDs in long linear chains.

Superdelocalizability,[1] abbreviated to SD below, is one of the most reliable indices of reactivity in the Hückel model.[2—5] This index has a clear theoretical ground and provides a basis for the use of frontier electron density in molecular orbital theory of reactivity.[2] Nowadays, aside from exceptional cases, we are already in a satisfactory stage as to computation in the Hückel model and have no practical difficulty at all in numerical evaluation of SDs for individual systems. However, it must be still significant and instructive to clarify general characteristics of SDs through noncomputational analysis, particularly for typical systems. Almost no exploration into such subjects has been done so far.[6]

This paper presents a simple noncomputational analysis of SDs of even-membered linear chains. Since no confusion is expected, the word "even-membered" is omitted for brevity in what follows. We have four sections below: The first section prepares two basic relationships connecting SDs to bond orders in general form. The second section provides a plain analytic expression to SDs of linear chains as a result of the relationships. The third section examines behaviors of SDs in long linear chains by using the analytic expression and derives simple estimation formulas for SDs near the end and in the middle of a long linear chain. The last section examines behaviors of approximate indices to SD in long linear chains for the sake of comparison.

1. Preliminaries

Consider a Hückel system specified solely by adjacency of sites. Assume that the system has no nonbonding orbital, namely, that the adjacency matrix A of the system is regular. Denote the j-th column-eigenvector of A by C_j, its eigenvalue by λ_j, and the site-r component of C_j by c_{rj}. We can assume that $\{C_j\}$ will form a complete orthonormal set and, therefore, we have

$$A = \sum_j \lambda_j C_j {}^t C_j. \tag{1}$$

Let us confine ourselves to the state in which each orbital j of positive λ_j is doubly occupied and other orbitals are unoccupied. For this state, the bond order matrix P can be written as[7]

$$P = \sum_j \frac{\lambda_j}{|\lambda_j|} C_j {}^t C_j \tag{2}$$

because of the completeness,

$$\sum_j C_j {}^t C_j = I. \tag{3}$$

The off-diagonal component p_{rs} of P is the bond order between sites r and s, while the diagonal one p_{rr} is the electron density on site r measured from unity. Let us introduce the auxiliary index g taking the values 1, 0, and -1, respectively, for electrophilic, radical, and nucleophilic reactions. Then, we can define the SD matrix $S^{(g)}$ as

$$S^{(g)} = \sum_j \left(\frac{1}{|\lambda_j|} + \frac{g}{\lambda_j} \right) C_j {}^t C_j. \tag{4}$$

The diagonal component $s_{rr}^{(g)}$ is the SD on site r, conventionally denoted by $S_r^{(g)}$. The off-diagonal one $s_{rs}^{(g)}$ is what is called the bond-delocalizability between sites r and s.[8]

From the spectral representations of A, P, and $S^{(g)}$ given above, it follows immediately that

$$AS^{(g)} = P + gI, \tag{5}$$

which can be rewritten, by using the set U_r of sites adjacent to site r, as

$$\sum_{u \in U_r} s_{us}^{(g)} = p_{rs} + g\delta_{rs}. \tag{6}$$

This component expression is ready to lead us to the following two lemmas essential for treating linear chains. Lemma 1 is the case of U_r having only one element s, while Lemma 2 follows from the rq- and rs-component expressions with U_r consisting of q and s.

Lemma 1. Let r be a site connected with only one adjacent site s. Then, $S_s^{(g)}$ satisfies

$$S_s^{(g)} = p_{rs} \tag{7}$$

and does not depend on g.

Lemma 2. Let r be a site having no bifurcation on it and let q and s be the two sites adjacent to r. Then, the difference between $S_q^{(g)}$ and $S_s^{(g)}$ is independent of g and the relationship

$$S_s^{(g)} - S_q^{(g)} = p_{rs} - p_{rq} \tag{8}$$

holds.

Using the inverse A^{-1} of A, we can write $S^{(g)}$ in an explicit form as

$$S^{(g)} = A^{-1}P + gA^{-1}. \tag{9}$$

Provided that A^{-1} is known, this gives a way of expressing SDs in terms of bond orders in principle. Resulting expressions, however, cannot be expected to be legible generally in themselves. Note that the dependence of $S^{(g)}$ on g is solely ascribed to A^{-1}. When the system is alternant, A^{-1} has non-vanishing components only between starred and unstarred sites. Therefore, $S_r^{(g)}$ does not depend on g, as is well known, and is denoted simply by S_r.

2. Analytic Expression

Consider a linear chain of m sites and number the sites consecutively from 1 to m. Denote the set $\{1, \cdots, m\}$ of natural numbers by M, the odd-number subset of M by M^{o}, and the even-number subset of M by M^{e}. This system is alternant and has a regular adjacency matrix whose eigensolution is known:[9]

$$\lambda_j = 2\cos\left(\frac{j}{m+1}\pi\right) \qquad (j \in M) \tag{10}$$

and

$$c_{rj} = \left(\frac{2}{m+1}\right)^{1/2} \sin\left(\frac{rj}{m+1}\pi\right) \qquad (r,j \in M). \tag{11}$$

Though the substitution of these expressions into

$$S_r = \sum_j^{m/2} \frac{2}{\lambda_j} c_{rj}^2 \tag{12}$$

yields an analytic expression of S_r, the expression is not very transparent. The preceding lemmas can lead us to another analytic expression via the expression of bond orders[10]

$$p_{r-1,r} = y(1) + (-1)^r y(r) \qquad (2 \leqq r \in M) \tag{13}$$

with

$$y(r) = \frac{1}{m+1}\mathrm{cosec}\left(\frac{2r-1}{2m+2}\pi\right). \tag{14}$$

For convenience's sake, let us prepare the symbols

$$Y(r) = \sum_{k=1}^{r} y(k) \qquad (1 \leqq r) \tag{15}$$

and

$$\bar{r} = m - r + 1. \tag{16}$$

From the symmetry of the system, site r and site $\bar{r}$ are equivalent. Note that $\bar{r}$ belongs to M^{e} for $r \in M^{\mathrm{o}}$ and vice versa.

Applying Lemma 2 to this system with the above expression of bond orders, we have

$$S_r - S_{r-2} = (-1)^r (y(r) + y(r-1)) \qquad (3 \leqq r \in M). \tag{17}$$

Since $y(r)$ is positive for all $r \in M$, it is clear that

$$S_r > S_{r-2} \qquad (4 \leqq r \in M^{\mathrm{e}}). \tag{18}$$

The largeness sequence of SDs in a linear chain is thus derivable quite straightforwardly.[6] As an immediate result of Lemma 1 it follows that

$$S_2 = y(1) + y(2). \tag{19}$$

Therefore, successive applications of Eq. 17 to even sites prove S_r to be simply $Y(r)$ for even r, that is, we have[11]

Formula 0.

$$S_r = \frac{1}{m+1}\sum_{k=1}^{r}\mathrm{cosec}\left(\frac{2k-1}{2m+2}\pi\right) \qquad (r \in M^{\mathrm{e}}). \tag{20}$$

This formula enables us to evaluate SDs efficiently unless the chain is infinitely long. Note that the summation over k does not have to be repeated to evaluate each SD. The process of evaluating S_m, the largest SD, produces all.[12]

Naturally, the evaluation of S_r for $r \in M^{\mathrm{o}}$ is replaced by that of $S_{\bar{r}}$ because of the symmetry:

$$S_r = Y(m - r + 1) \qquad (r \in M^{\mathrm{o}}). \tag{21}$$

Still, as a preparation for the next section, we rewrite this as

$$S_r = Y(m+1) - Y(r) \qquad (r \in M^{\mathrm{o}}) \tag{22}$$

to ascribe the r-dependence to $Y(r)$ by using the property

$$y(k) = y(m - k + 2). \tag{23}$$

Furthermore, we can rewrite $Y(m+1)$ as

$$Y(m+1) = 2Y(m/2) + 1/(m+1), \tag{24}$$

which provides the relationship

$$Y(m+1) = 2S_{m/2} + (-1)^{m/2}/(m+1) \tag{25}$$

relating $Y(m+1)$ to the SD on the middle site.

3. Long Linear Chains

The expressions of SDs obtained in the preceding section suffice to examine behaviors of SDs in long linear chains. It is obvious that the largeness sequence of SDs remains invariant even if linear chains become long. Still, a few remarkable features of SDs in long linear chains can be clarified by quantitative consideration.

First, let us examine SDs in the middle of a long linear chain of m sites, where each site can be written as $m/2+z$ with an integer z satisfying $|z| \ll m$. By rewriting $y(m/2+z)$ as

$$y(m/2+z) = \frac{1}{m+1}\sec\left(\frac{z-1}{m+1}\pi\right) \tag{26}$$

it is readily found that

$$y(m/2+z) \simeq m^{-1} - m^{-2} \qquad (|z| \ll m) \tag{27}$$

and, therefore, that SDs in the middle of the chain become closer to one another in value as m increases. This fact corresponds to the convergence of bond orders between adjacent sites in the middle of the chain to $2/\pi$ in the long chain

limit.[10] What should be noted here is that the SDs themselves diverge logarithmically. Through the power series expansion of cosecant, we can derive the asymptotic estimate $V[m]$ of $Y(m/2)$ to the second order of m^{-1} as

$$V[m] = \frac{1}{\pi}(\ln m + D + Em^{-1} + Fm^{-2}) \tag{28}$$

with

$$D = \ln(8/\pi) + \gamma \approx 1.51193, \tag{29}$$

$$E = -\pi/2 + 1 \approx -0.570796, \tag{30}$$

and

$$F = \pi/2 - 2/3 + \pi^2/72 \approx 1.04121, \tag{31}$$

where γ is Euler's constant. (See Appendix for details of derivation.) With this estimate, we can approximate $S_{m/2+z}$ for even sites as

Formula 1.

$$S_{m/2+z} \simeq V[m] + z(m^{-1} - m^{-2}) \qquad (|z| \ll m,\ m/2+z \in M^e). \tag{32}$$

For odd sites ($m/2+z \in M^o$), we have only to replace z with $1-z$ in the right hand side.

Next, let us consider SDs near the end of a long linear chain. Expanding $Y(r)$ into power series with the assumption $r \ll m$, we obtain

$$Y(r) \simeq \frac{2}{\pi}H_r^o + \frac{\pi}{12}\left(\frac{r}{m}\right)^2 \qquad (r \ll m) \tag{33}$$

with

$$H_r^o = \sum_{k=1}^{r} \frac{1}{2k-1}. \tag{34}$$

Hence, for even sites near site 1, the following formula immediately results.

Formula 2.

$$S_r \simeq \frac{2}{\pi}H_r^o + \frac{\pi}{12}\left(\frac{r}{m}\right)^2 \qquad (r \ll m,\ r \in M^e). \tag{35}$$

If the chain is sufficiently long, these SDs are determined essentially by the simple finite series H_r^o. The smallest SD, S_2, converges to a constant $8/3\pi$ in the long chain limit naturally as p_{12}. In sharp contrast to this, SDs on odd sites near site 1 as well as S_1 diverge logarithmically as m increases. Using $V[m]$ for $Y(m/2)$ in Eq. 24, we obtain the asymptotic estimate $W[m]$ of $Y(m+1)$ to the second order of m^{-1} as

$$W[m] = \frac{2}{\pi}\left(\ln m + D + m^{-1} + Gm^{-2}\right) \tag{36}$$

with

$$G = -2/3 + \pi^2/72 \approx -0.529589. \tag{37}$$

This estimate together with Eq. 33 provides

Formula 3.

$$S_r \simeq W[m] - \frac{2}{\pi}H_r^o - \frac{\pi}{12}\left(\frac{r}{m}\right)^2 \qquad (r \ll m,\ r \in M^o). \tag{38}$$

If the chain is sufficiently long, the difference between S_1 and SDs on odd sites near site 1 depends very little on m, satisfying

$$S_1 - S_r > \frac{2}{\pi}\sum_{k=2}^{r}\frac{1}{2k-1} \qquad (r \ll m,\ 3 \leqq r \in M^o). \tag{39}$$

Thus, the difference between the first and the second largest SDs, $S_1 - S_3$, does not vanish, converging to $16/15\pi$ in the long chain limit. This affirms that end sites remain distinctively reactive even if the chain becomes long.[13]

Estimations by Formulas 2 and 3 are practically exact if the condition $r \ll m$ is fulfilled. Even if the condition is not strictly fulfilled, the estimations are surprizingly excellent. Furthermore, the formulas seem to withstand extended use without serious sacrifice of accuracy. Though the first and the zeroth order approximations are sufficient as the case may be, their validity is remarkably restricted. As far as individual numerical estimations are concerned, we do not have to rely on them at all because of the simplicity of the second order correction terms. For SDs in the middle of long linear chains, estimation by Formula 1 works quite satisfactorily. When the condition $z \ll m$ is not strictly fulfilled, the second order terms become rather superfluous here. Though these estimation formulas originally concern SDs near the end or in the middle of a long linear chain, they enable us to evaluate any SD in long linear chains within the worst error of about 0.1 percent. For moderately long chains ($m \approx 10^2$), it will be appropriate to use Formulas 2 and 3 for $r/m < 0.38$, Formula 1 for $0.45 \leqq r/m \leqq 0.5$, and Formula 1 to the first order for $0.38 \leqq r/m < 0.45$. We have only to replace the borderline value 0.45 with 0.48 for longer chains ($m \approx 10^3$).

4. Approximate Indices

Frontier superdelocalizability,[14] abbreviated to frontier SD below, and frontier electron density can be regarded as practical approximations to SD.[2,3] Using the highest occupied orbital h, we can write the frontier electron density f_r and the frontier SD S'_r on site r as

$$f_r = 2c_{rh}^2 \tag{40}$$

and

$$S'_r = 2c_{rh}^2/\lambda_h (= f_r/\lambda_h) \tag{41}$$

for alternants having no nonbonding orbital without regard to types of reactions. For a linear chain of m sites, we have

$$f_r = \frac{2}{m+1}\left(1 - (-1)^r \cos\left(\frac{r}{m+1}\pi\right)\right) \qquad (r \in M) \tag{42}$$

and

$$S'_r = y(1)\left(1 - (-1)^r \cos\left(\frac{r}{m+1}\pi\right)\right) \qquad (r \in M) \tag{43}$$

using the eigensolution for the highest occupied orbital $m/2$.

By checking the sign of $f_r - f_{r-2}$, it is readily found that the largeness sequence of frontier electron densities[15] coincides with that of SDs in linear chains. It is obvious, however, that all the frontier electron densities tend to zero, becoming closer to one another in value, as m increases. Thus, frontier electron density cannot substitute for SD regarding behaviors in long linear chains.

The largeness sequence of frontier SDs is naturally identical with that of frontier electron densities and, therefore, with that of SDs in linear chains. However, it follows from Eq. 43 that

$$S'_r \simeq \frac{2}{\pi}(1-(-1)^r)+O((r/m)^2) \qquad (r \ll m) \tag{44}$$

and

$$S'_{m/2+z} \simeq \frac{2}{\pi}+O(z/m) \qquad (|z| \ll m). \tag{45}$$

Namely, in the long chain limit, the frontier SDs converge to zero on even sites near site 1, to $4/\pi$ on odd sites near site 1 as well as on site 1, and to $2/\pi$ in the middle of the chain. Here, neither the difference between S_1 and S'_1 nor that between S_2 and S'_2 is significant. It is crucial that the difference between the largest frontier SDs, S'_1 and S'_3, becomes smaller, tending to zero as m increases. Thus, frontier SD also fails to indicate the distinctive reactivity of end sites in long linear chains.

The above-mentioned defect of frontier electron density is primarily due to the omission of the denominator λ_h which tends to zero as m increases. The defect is, therefore, removed in frontier SD to some extent. However, a long linear chain has many orbitals of small $|\lambda_j|$ near λ_h besides the highest occupied orbital, but the contributions from these orbitals are not taken into account even in frontier SD. Similar situations can occur in other "large" systems having many sites.

Appendix

Using Bernoulli's numbers $\{B_n\}$, we can write the power series expansion of cosecant as

$$\operatorname{cosec} x = \frac{1}{x}+\sum_{n=1}^{\infty} b_n x^{2n-1} \qquad (0<x<\pi) \tag{46}$$

with

$$b_n = (2^{2n}-2)B_n/(2n)!. \tag{47}$$

This expansion rewrites $Y(m/2)$ as the sum of

$$Q = \frac{2}{\pi}\sum_{k=1}^{m/2}\frac{1}{2k-1} \tag{48}$$

and

$$R = \sum_{n=1}^{\infty} b_n\left(\frac{\pi}{2}\right)^{2n-1}\left(\frac{1}{m+1}\right)^{2n}\sum_{k=1}^{m/2}(2k-1)^{2n-1}. \tag{49}$$

Rewriting Q as

$$Q = \frac{2}{\pi}\left(H_m-\frac{1}{2}H_{m/2}\right) \tag{50}$$

with Harmonic numbers $\{H_n\}$ and applying the well-known asymptotic expansion

$$H_n \simeq \ln n+\gamma+\frac{1}{2}n^{-1}-\frac{1}{12}n^{-2} \qquad (1 \ll n) \tag{51}$$

to H_m and $H_{m/2}$, we obtain

$$Q \simeq \frac{1}{\pi}\left(\ln m+\ln 2+\gamma+\frac{1}{6}m^{-2}\right). \tag{52}$$

On the other hand, for large m, since

$$\left(\frac{1}{m+1}\right)^{2n} \simeq m^{-2n}(1-2nm^{-1}+n(2n+1)m^{-2}) \tag{53}$$

and

$$\sum_{k=1}^{m/2}(2k-1)^{2n-1} \simeq m^{2n}\left(\frac{1}{4n}-\frac{2n-1-\delta_{n1}}{6}m^{-2}\right), \tag{54}$$

we have

$$R \simeq \frac{1}{\pi}L_{-1}-\frac{1}{2}L_0 m^{-1}+\left(\frac{\pi}{24}L_1+\frac{5}{12}L_0+\frac{\pi}{72}\right)m^{-2} \tag{55}$$

with

$$L_\alpha = \sum_{n=1}^{\infty}(2n)^\alpha b_n\left(\frac{\pi}{2}\right)^{2n-1-\alpha} \qquad (\alpha = 0, \pm 1). \tag{56}$$

The constants L_0, L_{-1}, and L_1 can be evaluated by using Eq. 46 and its integrated and differentiated forms:

$$L_0 = \operatorname{cosec} x - 1/x\,|_{\pi/2} = 1-2/\pi, \tag{57}$$

$$L_{-1} = \ln\left(\frac{\tan(x/2)}{x/2}\right)\Big|_{\pi/2} = \ln(4/\pi), \tag{58}$$

and

$$L_1 = (1/x-\cot x)\operatorname{cosec} x\,|_{\pi/2} = 2/\pi. \tag{59}$$

References

1) K. Fukui, T. Yonezawa, and C. Nagata, *Bull. Chem. Soc. Jpn.*, **27**, 423 (1954).

2) K. Fukui, T. Yonezawa, and C. Nagata, *J. Chem. Phys.*, **26**, 831 (1957).

3) K. Fukui, T. Yonezawa, and C. Nagata, *J. Chem. Phys.*, **27**, 1247 (1957).

4) K. Fukui, T. Yonezawa, and C. Nagata, *Bull. Chem. Soc. Jpn.*, **34**, 37 (1961).

5) K. Fukui, "Molecular Orbitals in Chemistry, Physics, and Biology," ed by P.-O. Löwdin and B. Pullman, Academic Press, New York (1964), pp. 513—537.

6) The largeness sequence of SDs in linear chains has been proved essentially in a footnote on p. 834 in Ref. 2.

7) This bond order matrix differs from the p-density matrix (bond order-electron density matrix) by the unit matrix.

8) K. Fukui, A. Imamura, T. Yonezawa, and C. Nagata, *Bull. Chem. Soc. Jpn.*, **34**, 1076 (1961).

9) For example: C. A. Coulson and H. C. Longuet-Higgins, *Proc. R. Soc. London, Ser. A*, **192**, 16 (1947).

10) C. A. Coulson, *Proc. R. Soc. London, Ser. A*, **169**, 413 (1939).

11) Unfortunately, it seems to be impossible to put the finite series of Formula 0 into a direct function of r.

12) While frontier electron density can be used only as a relative index of reactivity in a system, SD serves as an absolute index of reactivity. Hence, it will never be futile to have simple formulas evaluating SDs.

13) The author knows no direct experimental data on site reactivities of "isolated" long linear chains, though his inquiry into the literature is far from exhaustive.

14) This quantity was introduced by Fukui and his co-workers, referred to as one-term approximation of superdelocalizability (Ref. 3). The concise name "frontier superdelocalizability" seems to have originated with A. J. Wohl: "Molecular Orbital Studies in Chemical Pharmacology," ed by L. B. Kier, Springer Verlag, Berlin (1970), p. 105.

15) This largeness sequence has been mentioned in a footnote on p. 835 in Ref. 2 as an evident fact.