Relation between transmission rates and the wave functions in carbon nanotube junctions

Ryo Tamura and Masaru Tsukada
Department of Physics, Graduate School of Science, University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113-0033, Japan
(Received 18 February 1999)

Electron transmission and wave functions through junctions with a pair of a pentagonal defect and a heptagonal defect connecting two metallic carbon nanotubes are analyzed by the analytical calculation with the effective-mass equation. The energy region $|E|<E_c$ is considered where the channel number is kept to two. Close relation between the transmission rate and the wave function is found; the transmission rate is given by the inverse squared absolute value of the wave function. The dependence of the transmission rates on the energy and on the size of the junction is clearly explained by the nature of the wave function. Though the wave function and the transmission rate calculated by the tight-binding model agree well with the corresponding analytical results by the effective-mass approximation, the discrepancy becomes considerable when $|E| \simeq E_c$. To study the origin of this discrepancy, an efficient numerical calculation method is developed with a generalized transfer matrix for the tight-binding model. Their numerical results are compared with the corresponding analytical ones and the results show that the origin of the discrepancy comes from the evanescent waves with the longest decay length in the tube parts.

I. INTRODUCTION

Recent experimental development on the carbon nanotubes, especially electronic transport measurements for individual nanotubes,[1] has promoted much interest in the nanotubes as a one-dimensional conductor with nanometer size.[2] Many experimental works concern the theoretical prediction that the nanotube becomes metallic or semiconducting according to its circumference.[3,4] The conductance of the metallic nanotubes with potential energy disorder has been studied.[5,6] The junction structures with a shape of a part of a cone connecting two nanotubes with different radii have also been observed,[7] and studied theoretically.[8–11] They are formed by a pair of a pentagonal defect and a heptagonal defect.[12–14] By composing the nanotube junctions, the electronic circuits with nanometer size might be designed. The atom bond network of the nanotube junctions is uniquely determined by its development map with the vector of the circumference of the connected thicker nanotube $\vec{R}_5$ and that of the connected thinner tube $\vec{R}_7$. We have calculated the conductance with Landauer's formula, i.e., the transmission rate, for the various junctions connecting two metallic nanotubes by the tight binding model.[8,9] Surprisingly, the dependence of the transmission rate on the parameters $\vec{R}_5$, $\vec{R}_7$, and E has a very simple form obeying the scaling law; in the energy region $|E|<E_c$ where the channel number is kept to two in both the tubes, the transmission rate is independent of detailed atomic arrangements as well as the angle between the two tube axes but determined only by the two parameters; one of the two parameters is the ratio of the circumference of the tubes R_7/R_5 and the other is the scaled energy E/E_c. The close relation between the transmission rate and the wave function in the junction part is also found by the tight-binding model.[8,9] When $E=0$, the transmission rate decays with the thickness of the junction as $4/[2+(R_5/R_7)^3+(R_7/R_5)^3]$ while the spatial decay of the corresponding wave function in the junction part obeys the same power law.[8] When $R_7/R_5 \simeq 0.5$, the transmission rate shows a peak structure as a function of the energy, while the corresponding wave function shows a resonant feature, i.e., its amplitude is enhanced in the middle of the junction part.[9]

Recently, Matsumura and Ando have confirmed the power-law decay for $E=0$ by using the effective-mass theory.[15] We generalized their discussion to a more general energy region $|E|<E_c$, and obtained the complete analytical form of the transmission rate with the two parameters R_7/R_5 and $|E|/E_c$.[16] The agreement of the analytical transmission rate with the numerical one is found to be fairly good as long as $|E|$ is not very close to E_c. By our generalization of the effective-mass approximation, the band structures of the periodic multiple nanotube junctions, which are called the ''helically coiled nanotube'' in Ref. 17, can be also obtained as closed analytical forms and their relation to the configuration of the pentagon and the heptagon are clearly explained based on the symmetry features.[18] The effective-mass theory is derived from the tight-binding model and is valid when the energy is near zero. It is suitable to obtain the analytical results and to explain the origin of the scaling law because it treats not the discrete space but the continuous space, i.e., it does not necessitate the detailed atomic structures of the honeycomb lattice.[6,15,16,18–20] Though the analytical transmission rate of the nanotube junction by the effective-mass theory has been studied in detail in this way, the close relation between the transmission rate and the wave function observed in the numerical result is not clarified by the effective-mass theory yet. In this paper, the spatial variations of the wave function are expressed unambiguously by the effective-mass theory and it gives an intuitive explanation for the dependence of the transmission rate on R_7/R_5. Furthermore, it will also be found that its dependence on the energy has a close relation to the other parameter $|E|/E_c$ of the scaling law. These points are discussed in Sec. III. The concise expression of the wave function characterized by only the two parameters will be useful when the scanning tunnel microscope (STM) images of the nanotube junction are analyzed.

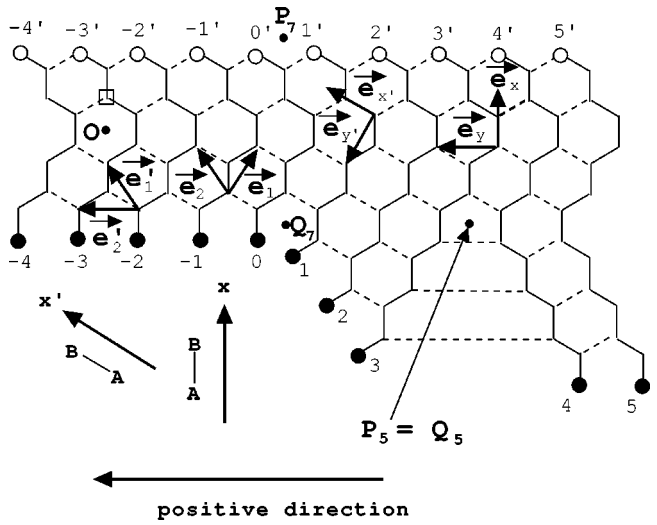

FIG. 1. The development map showing the bond network of the (2,2)-(2,5) junction. The filled circle at the bottom and the open circle at the top in each bar indicate an identical atom. To form the junction, the development map is rolled up so that the filled circle and the open circle coincide with each other in each bar. The solid lines and the dashed lines represent the bonds within the bar and those connecting the neighboring bars, respectively. The points $P_5(=Q_5)$, $P_7(=Q_7)$, and O are the centers of the pentagonal defect, that of the heptagonal defects, and the origin of the coordinate, respectively, which are also shown in Fig. 3. The two sets of the translation vectors, $\{\vec{e}_1,\vec{e}_2\}$ and $\{\vec{e}_1',\vec{e}_2'\}$, and the corresponding sets of orthogonal vectors $\{\vec{e}_x,\vec{e}_y\}$ and $\{\vec{e}_{x'},\vec{e}_{y'}\}$ are shown. Note that definition of sublattices A and B are changed between the two sets.

On the other hand, the discrepancy between the numerical transmission rate with the tight-binding theory and the analytical one by the effective-mass approximation become considerable when $|E|$ is very close to E_c; the numerical transmission rate shows a sharp dip there while the analytical one does not. Since the effective-mass theory is an approximation derived from the tight-binding model, the numerical result by the tight-binding model is considered to be the correct one. To discuss the origin of the sharp dip, the wave functions calculated by the tight-binding model have to be compared with those calculated by the effective-mass theory. To calculate the wave function by the tight-binding model more efficiently, a numerical method is developed in Sec. II. From the comparison of the wave function calculated by the method in Sec. II with the analytical one, we found in Sec. IV that the evanescent waves in the tube parts and the defect levels at the pentagon and the heptagon have important roles in forming the dip structures.

II. THE CONDITIONED TRANSFER MATRIX METHOD

Figure 1 shows the development map of the nanotube junction. It is characterized by bars representing zigzag segments of the C-C bond network in the circumferential direction of the single junction. They are aligned and numbered along the direction $\vec{e}_1-\vec{e}_2$ where $\vec{e}_1$ and $\vec{e}_2$ are the basic translation vectors of the graphite plane. Each bar is connected with adjacent bars by the remaining C-C bonds. The network is rolled up so that the atoms denoted by j in the bottom of the jth bar and the top one of the same bar denoted by j' become the identical one. The j and j' sites are shown by the filled and the open circles in Fig. 1, respectively. When $j\leqslant 0$, the jth bar is formed by the $(m+n)$ unit cells of the two-dimensional (2D) graphite, where the m unit cells are aligned along $\vec{e}_1$ and the others are aligned along $\vec{e}_2$. So the (m,n) tube defined in Ref. 4 is formed there. To connect the (m,n) tube with another (m_2,n_2) tube, the other bars are formed in the following way with integer parameters $l_1 \equiv (m_2+n_2)-(m+n)\geqslant 0$ and $l_2\equiv n_2-n\geqslant 0$. For $1\leqslant j\leqslant l_1$, the jth bar consists of $m+j-1$ unit cells along $\vec{e}_1$ and n unit cells along $\vec{e}_2$, plus an extra atom that is represented in Fig. 1 as the solid jth circle. For $j\geqslant l_1+1$, the jth bar consists of l_2 unit cells along $\vec{e}_2$, followed by $m+l_1-l_2$ unit cells along $\vec{e}_1$, followed by n unit cells along $\vec{e}_2$. The network made in this way represents the junction of the (m,n) tube $(j\leqslant 0)$ and the (m_2,n_2) tube $(j\geqslant l_1+1)$. This junction is called an (m,n)-(m_2,n_2) junction hereafter. Then Fig. 1 corresponds to the (2,2)-(2,5) junction. A seven-membered ring is formed at the bottom of the first bar and a five-membered ring is introduced between the l_1th bar and the (l_1+1)th bar. There are only six-membered rings elsewhere.

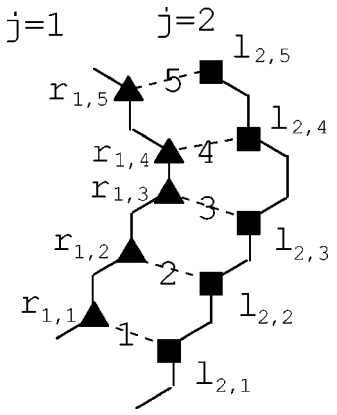

FIG. 2. Notation to represent the amplitudes of the wave function. It is illustrated with the first bar and the second bar in Fig. 1. The site in the jth bar can be classified to the two groups according to whether it connects with the right $(j+1)$th bar or with the left $(j-1)$th bar. The former and the latter are denoted by $r_{j,i}$ (closed triangles) and $l_{j,i}$ (closed squares), respectively, when the corresponding site has the bond i, which connects the neighboring bars and is numbered from the bottom to the top.

The number of the bonds connecting the $(j-1)$th bar and the jth bar is denoted by b_j as

$$b_j=m+n\ (j\leqslant 0),$$

$$b_j=m+n+j-1\ (1\leqslant j\leqslant l_1),$$

$$b_j=m_2+n_2\ (l_1+1\leqslant j). \tag{1}$$

The amplitudes of the wave function in the jth bar are represented by the vector $\vec{c}_j={}^t(c_{j,1},c_{j,2},\ldots)$. Its component $c_{j,i}$ can be classified to the two groups according to whether the corresponding site connects with the right $(j+1)$th bar or with the left $(j-1)$th bar. These two groups are denoted by r_j and l_j, respectively, as is shown in Fig. 2. Since each site in the jth bar belongs to only and necessarily one of the two

groups, the dimension of $\vec{c}_j$ equals b_j+b_{j+1}, which are denoted by d_j hereafter. The bonds between the $(j-1)$th bar and the jth bar are numbered as Fig. 2 from the bottom to the top. According to this way of numbering, the amplitudes are also numbered by $r_{j-1,i}$ and $l_{j,i}$, which have the bond numbered by i $(i=1,2,\ldots,b_j)$. Then the tight-binding equation becomes

$$Er_{j-1,i}=\gamma\left(l_{j,i}+\sum_{i'} c_{j-1,i'}\right) \tag{2}$$

and

$$El_{j,i}=\gamma\left(r_{j-1,i}+\sum_{i''} c_{j,i''}\right), \tag{3}$$

where the first and the other terms in the right-hand sides represent the bonds between the neighboring bars and those within the bar, respectively. Here we use the tight-binding model including only π orbitals with common hopping integral $\gamma(\simeq -2.7$ eV) and common site energy chosen to be zero. In order to get more accurate results, the effects from the mixing with σ orbitals caused by curvature of the graphitic plane also have to be considered. But we believe that this tight-binding model gives a transparent view about this system and qualitative valid results. Another reason why this model is used is that the purpose of this paper is to focus on the effects from the *connectivity* of the bond network rather than those from the curvature.

The tight-binding equation, Eqs. (2) and (3), can be summarized by the matrix form as

$$A_j\vec{c}_{j-1}+B_j\vec{c}_j=0, \tag{4}$$

where B_j and A_j are a $2b_j\times d_j$ matrix and a $2b_j\times d_{j-1}$ matrix, respectively. In the left (right) tube parts, i.e., when $j\leqslant 0$ $(j\geqslant l_1+2)$, these matrixes become constant matrixes, A_L and B_L (A_R and B_R). The transfer matrixes for the tube parts are obtained as $T_L=-A_L^{-1}B_L$ for the left tube and $T_R=-A_R^{-1}B_R$ for the right tube. The eigenvalues and eigenvectors of T_μ are classified into two groups as $\{\beta_i^\mu,\psi_{+i}^\mu\}$ and $\{1/\beta_i^\mu,\psi_{-i}^\mu\}$. Hereafter, $\mu=R$ and $\mu=L$ represent the left thinner tube and the right thicker tube, respectively. The propagating waves in the former group, which are ψ_{+i} with $|\beta_i|=1$, are assigned to $i=1,\ldots,n_\mu$, where n_μ is called the channel number of the corresponding tube. They carry the probability flow with the positive velocity v_i. The propagating waves in the latter group, $\psi_{-i}=\psi_i^*$, carry that with the negative velocity $-v_i$. From now on, the normalized extended states with unit flow $\tilde{\psi}_i=\psi_i/\sqrt{v_i}$ and $\tilde{\psi}_{-i}=\psi_{-i}/\sqrt{v_i}$ are used instead of ψ_i and ψ_{-i}. The other states for $|i|>n_\mu$ are evanescent waves and the sign of i means the direction along which they decay exponentially, i.e., $|\beta_i^\mu|<1$. The wave function in the tube parts are represented by superposition of these eigenvectors and its coefficients are represented by $\vec{x}$ as

$$\vec{c}_j=U_+^\mu(\Lambda^\mu)^{+j}\vec{x}_+^\mu+U_-^\mu(\Lambda^\mu)^{-j}\vec{x}_-^\mu, \tag{5}$$

where the ith column of $U_\pm^\mu$ is the normalized eigenvector $\tilde{\psi}_\pm^\mu$, and Λ^μ is the diagonal matrix whose (i,i) element is the eigenvalue β_i^μ.

Unlike the tube parts, the ''usual'' square transfer matrix cannot be obtained in the junction part, because the dimensions of $\vec{c}_j$ and that of $\vec{c}_{j+1}$ are different, i.e., the A_j for $1\leqslant j\leqslant l_1$ becomes a rectangle matrix that has no inverse matrix. Nevertheless there is a $d_{j-1}\times 2b_j$ matrix that is similar to the inverse matrix of A_j. It is denoted by $\tilde{A}_j$ and defined by

$$\tilde{A}_j\equiv(A_j^\dagger A_j)^{-1}A_j^\dagger. \tag{6}$$

This matrix satisfies the condition $\tilde{A}_jA_j=1$ but $A_j\tilde{A}_j\neq 1$. By this pseudoinverse matrix, we can define the ''transfer matrix'' $T_j\equiv-\tilde{A}_jB_j$. If Eq. (4) is satisfied, then

$$\vec{c}_{j-1}=T_j\vec{c}_j \tag{7}$$

is satisfied, but its converse does not hold generally. It means that $\vec{c}_j$ in Eq. (7) cannot be chosen to be arbitrary but has to satisfy

$$(B_j+A_jT_j)\vec{c}_j=0. \tag{8}$$

To make Eq. (7) equivalent to Eq. (4), the condition (8) is necessary, so we call T_j a ''conditioned transfer matrix'' hereafter. In our problem, the number of the independent rows of $(B_j+A_jT_j)$ is only one, so Eq. (8) can be written as

$${}^t\vec{s}_j\cdot\vec{c}_j=0, \tag{9}$$

where ${}^t\vec{s}_j$ is one of the nonzero rows of $(B_j+A_jT_j)$. Multiplying T_j generates the coefficient at the bar with the decreased j, so we define the positive direction along which j decreases, i.e., from the right thicker tube to the left thinner tube.

By using the inverse matrix ${}^t({}^tV_+^L,{}^tV_-^L)\equiv(U_+^L,U_-^L)^{-1}$ and $K_j\equiv T_{j+1}T_{j+2}\ldots T_{l_1+1}$, the relation between $\vec{c}_{l_1+1}$ and the right-going waves in the thinner tube $\vec{x}_-^L$ are summarized as

$$\begin{pmatrix} V_-^LK_0 \\ {}^t\vec{s}_{l_1+1} \\ {}^t\vec{s}_{l_1}K_{l_1} \\ {}^t\vec{s}_{l_1-1}K_{l_1-1} \\ \vdots \\ {}^t\vec{s}_2K_2 \end{pmatrix}\vec{c}_{l_1+1}=\begin{pmatrix} \vec{x}_-^L \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{10}$$

The $(m_2+n_2)\times(2m_2+2n_2)$ matrix in the left-hand side of Eq. (10) is represented by Q hereafter.

Now we define all the necessary things to calculate the transmission rate and the wave function. When the electron is incident from the right and transmitted into the thinner tube, i.e., when $\vec{x}_+^R\neq 0$ and $\vec{x}_-^L=0$, the reflected wave $\vec{x}_-^R$ and the transmitted wave $\vec{x}_+^L$ are obtained from

$$\vec{x}_-^R=-(QU_-^R)^{-1}(QU_+^R)\vec{x}_+^R \tag{11}$$

and

$$\vec{x}_+^L = V_+^L K_0 [U_+^R - U_-^R (QU_-^R)^{-1}(QU_+^R)]\vec{x}_+^R , \qquad (12)$$

respectively. The corresponding wave function in the junction part, i.e., $\vec{c}_j$ for $1 \leqslant j \leqslant l_1$, is given by

$$\vec{c}_j = K_j [U_+^R - U_-^R (QU_-^R)^{-1}(QU_+^R)]\vec{x}_+^R . \qquad (13)$$

For the inverse direction of the incident electron, $\vec{x}_+^R = 0$ and $\vec{x}_-^L \neq 0$, the reflected wave $\vec{x}_+^L$ and the transmitted wave $\vec{x}_-^R$ are obtained from

$$\vec{x}_+^L = V_+^L K_0 U_-^R (QU_-^R)^{-1} \begin{pmatrix} \vec{x}_-^L \\ 0 \end{pmatrix} \qquad (14)$$

and

$$\vec{x}_-^R = (QU_-^R)^{-1} \begin{pmatrix} \vec{x}_-^L \\ 0 \end{pmatrix} . \qquad (15)$$

In this case, the wave function in the junction part is given by

$$\vec{c}_j = K_j U_-^R (QU_-^R)^{-1} \begin{pmatrix} \vec{x}_-^L \\ 0 \end{pmatrix} . \qquad (16)$$

The transmission rate is obtained from Eq. (12) or Eq. (15), while the reflection rate is obtained from Eq. (11) or Eq. (14) from the matrix elements corresponding to the open channel, i.e., $x_{\pm i}^{\mu}$ with $i = 1 \sim n_{\mu}$. We have confirmed that unitarity holds very well in the conditioned transfer matrix method. Agreement of their results with those calculated by usual recursive Green's-function methods is also quite good. In both methods, matrix inversions are necessary about l_1 times, but they can be done much faster in the conditioned transfer matrix method than in the recursive Green's-function method, because matrixes that have to be inverted are real sparse symmetric matrixes $A_j^{\dagger} A_j$ in the former method while those in the latter method are complex matrixes with no symmetry. Therefore the conditioned transfer matrix method is much faster than the recursive Green's-function method. This efficiency is expected to be more important when more general related problems are calculated, e.g., those including electron-phonon interaction or electron-electron interaction self-consistently. Another advantage of the conditioned transfer matrix method over the recursive Green's-function method, which is not essential but practically important, is that it is more intuitive and easy to be implemented.

In Sec. IV, the wave functions are calculated numerically by the conditioned transfer matrix method and compared to those calculated analytically by the effective-mass equation.

III. EFFECTIVE-MASS THEORY AND ITS APPLICATION TO THE SINGLE NANOTUBE JUNCTION

Figure 3 shows a development map of the nanotube junction neglecting the atomic sites. The vectors $\vec{R}_5$ and $\vec{R}_7$ represent the circumferences of the thicker tube and the thinner tube, respectively. Lines $E_7P_7P_5E_5$ and lines $F_7Q_7Q_5F_5$ are stuck to each other so that the points connected by $\vec{R}_j$ become identical ($j = 5,7$). Then $P_5(=Q_5)$ and $P_7(=Q_7)$ turn out to be the centers of a heptagonal defect and a pentagonal defect, respectively. These points are marked also in Fig. 1. Thus, from now on, the indices ''7'' and ''5'' are used to represent the thinner and the thicker tube, respectively.

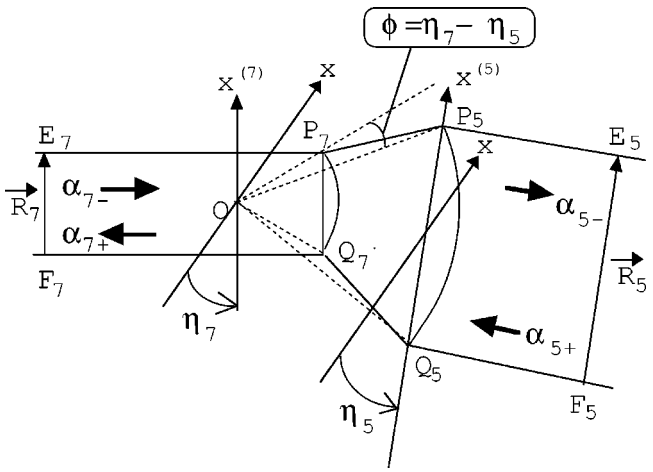

FIG. 3. Development map of the nanotube junctions. The lines $E_7P_7P_5E_5$ are connected and become identical with the lines $F_7Q_7Q_5F_5$, respectively. The rectangles $E_7P_7Q_7F_7$ and $P_5E_5F_5Q_5$ form the thinner tube and the thicker tube, respectively. The triangle OP_7P_5 is the same as the OQ_7Q_5 rotated by 60°. The quadrilateral $P_7P_5Q_5Q_7$ forms a junction part with a shape of a part of a cone. A heptagonal defect and a pentagonal defect are introduced at $P_7(=Q_7)$ and $P_5(=Q_5)$, respectively. The direction of the circumferences of the tubes in the development map is represented by their angles η_5 and η_7 measured counterclockwise with respect to the x axis, which is defined in Fig. 1. The angle between the axes of the two tubes ϕ is defined as $\phi = \eta_7 - \eta_5$.

Then the position of a general site $\vec{r}$ can be labeled by (s,q_1,q_2), where q_1 and q_2 are integer components representing the position of the unit cell $\vec{q} \equiv q_1\vec{e}_1 + q_2\vec{e}_2$ and s represents the sublattice $s = A,B$. The relation between the position vector of the site $\vec{r}$ and its label (s,q_1,q_2) is written by

$$\vec{r}(A,q_1,q_2) = \vec{q} + (\vec{e}_1 + \vec{e}_2)/3 = \frac{\sqrt{3}}{2}\left(q_1 + q_2 + \frac{2}{3}\right)\vec{e}_x + \frac{1}{2}(q_2 - q_1)\vec{e}_y \qquad (17)$$

and

$$\vec{r}(B,q_1,q_2)=\vec{q}+2(\vec{e}_1+\vec{e}_2)/3 = \frac{\sqrt{3}}{2}\left(q_1+q_2+\frac{4}{3}\right)\vec{e}_x+\frac{1}{2}(q_2-q_1)\vec{e}_y. \tag{18}$$

In the above, $\vec{e}_x\equiv(\vec{e}_1+\vec{e}_2)/\sqrt{3}$ and $\vec{e}_y\equiv\vec{e}_2-\vec{e}_1$, which define the coordinate system (x,y) shown in Fig. 1. The amplitude of the wave function at these sites is denoted by $\psi(s,q_1,q_2)$. In the effective-mass theory, the wave function is represented by

$$\psi(s,q_1,q_2)=F_s^K(\vec{r})w^{(q_1-q_2)}+F_s^{K'}(\vec{r})w^{(q_2-q_1)} \quad (s=A,B). \tag{19}$$

Here $w\equiv\exp(i2\pi/3)$ and $\vec{r}$ is defined by Eqs. (17) and (18) as a function of the label (s,q_1,q_2). In Eq. (19), $F_{A,B}^{K,K'}$, $w^{(q_1-q_2)}$, and $w^{(q_2-q_1)}$ are the envelope wave functions and the wave function of the Bloch states at the K and the K' point, respectively. This wave function is expressed by a vector $\vec{\psi}=(F_A^K(\vec{r}),F_B^K(\vec{r}+\vec{e}_x/\sqrt{3}),F_A^{K'}(\vec{r}),F_B^{K'}(\vec{r}+\vec{e}_x/\sqrt{3}))$ hereafter. This definition of F's is different from other references[6,15,16,19] by certain factors. The reason why this definition is used is that the representation of the time reversal operation I, $I\psi=\psi^*$, becomes simpler as

$$I(F_A^K,F_B^K,F_A^{K'},F_B^{K'})=((F_A^{K'})^*,(F_B^{K'})^*,(F_A^K)^*,(F_B^K)^*). \tag{20}$$

By using Eq. (19), the effective-mass equation is derived from the tight-binding equation used in the preceding section. When energy E is zero, $w^{q_1-q_2}$ and $w^{q_2-q_1}$ are solutions of the tight-binding model, so that the corresponding solutions of the envelope functions F are constant. Thus when E is not zero but close to zero, spatial variation of the envelope functions is slow compared to the lattice constant, $|\vec{e}_x|=|\vec{e}_y|\equiv a\simeq 0.25$ nm. In this case, it is a good approximation to take only the first-order term in the Taylor expansion of the envelope function as $F(\vec{r}+\vec{e}_x)\simeq(1+a\partial_x)F(\vec{r})$. From this approximation, one obtains[6,19,20]

$$(-i\partial_y+\partial_x)F_B^K(\vec{r}+\vec{b})=\epsilon F_A^K(\vec{r}), \tag{21}$$

$$(-i\partial_y-\partial_x)F_A^K(\vec{r})=\epsilon F_B^K(\vec{r}+\vec{b}), \tag{22}$$

$$(i\partial_y+\partial_x)F_B^{K'}(\vec{r}+\vec{b})=\epsilon F_A^{K'}(\vec{r}), \tag{23}$$

$$(i\partial_y-\partial_x)F_A^{K'}(\vec{r})=\epsilon F_B^{K'}(\vec{r}+\vec{b}), \tag{24}$$

where $\epsilon=2E/(\sqrt{3}\gamma a)$ and $\vec{b}\equiv\vec{e}_x/\sqrt{3}$. When the plane-wave solution $F_A^K,F_B^K\propto\exp(\pm i\vec{k}\cdot\vec{r})$ is used in Eqs. (21) and (22), one can get the linear isotropic dispersion relation,

$$k\equiv|\vec{k}|=|\epsilon|, \tag{25}$$

and $F_B^K(\vec{r}+\vec{b})/F_A^K(\vec{r})=\pm\exp(i\eta)$, where $(k_x,k_y)=(-\epsilon\sin\eta,\epsilon\cos\eta)$, i.e., η is the angle of $\vec{k}$ with respect to the y axis measured counterclockwise. It follows that the corresponding wave function, $\vec{\psi}_{K\pm}$ is written as

$$\vec{\psi}_{K\pm}=(e^{-i\eta/2},\pm e^{i\eta/2},0,0)e^{\pm i(\vec{k}\cdot\vec{r})}. \tag{26}$$

The wave number must satisfy the boundary condition in the metallic nanotube parts as

$$\vec{k}\cdot\vec{R}_j=2\pi l_j \quad (j=5,7), \tag{27}$$

where l_j is an integer representing the number of oscillations of the envelope function around the circumferences. As for the semiconducting tubes, the boundary condition becomes different from Eq. (27),[19] but we concentrate our discussion on the metallic nanotubes in this paper. The upper sign and the lower sign in Eq. (26) represent the direction of the propagating waves. When E is close to zero, i.e., $k\simeq 0$, the only possible number of l_j is zero. It means that $\vec{k}$ is perpendicular to $\vec{R}_j$ so that η is the angle of $\vec{R}_j$ with respect to the x axis. Within the effective-mass theory scheme, the possible maximum value of l_5 (l_7) is the channel number of the thicker tube (the thinner tube). Therefore, the range of the energy where the channel number is kept to two in the thicker tube (the thinner tube) is $|E/\gamma|<\sqrt{3}\pi a/R_5$ ($|E/\gamma|<\sqrt{3}\pi a/R_7$). Discussions in this paper are concentrated on the energy region where the channel number is kept to two in both tubes. It is represented by $|E|<(\sqrt{3}\pi|\gamma|)(a/R_5)\equiv E_c$ and E_c is called a threshold energy hereafter. From the propagating waves near the K point, $\vec{\psi}_{K\pm}$, the other propagating waves $\psi_{K'\pm}$ are obtained by the time reversal operation (20) as

$$\vec{\psi}_{K'\pm}=I\vec{\psi}_{K\mp}=(0,0,e^{i\eta/2},\mp e^{-i\eta/2})e^{\pm i(\vec{k}\cdot\vec{r})}. \tag{28}$$

Note that the direction of the propagation is reversed by the time reversal operation I. In order to discuss the wave function in the junction part, the polar coordinate (r,θ) is useful. Its relation to the coordinate (x,y) is the usual one, i.e., $r=\sqrt{x^2+y^2}$,$\tan\theta=y/x$. Then the wave function satisfies the wave equation $r^2(\partial_x^2+\partial_y^2+k^2)F=(z^2\partial_z^2+z\partial_z+\partial_\theta^2+z^2)F=0$, where $z=kr$. The solution is represented by Bessel functions J_m and Neumann functions N_m as

$$F=\sum_{m=-\infty}^{\infty}e^{im\theta}[c_mJ_{|m|}(z)+d_mN_{|m|}(z)]. \tag{29}$$

In Fig. 1, site i is identical with site i' for $i=1,\ldots l_1$, while corresponding labels (s,q_1,q_2) are different between i and i'; that for i is $(B,1,-4-i)$ and that for i' is $(A,3+i,-2-i)$. In this way, the wave function in the junction part must satisfy the condition

$$\psi(A,q_1,q_2)=\psi(B,q_1+q_2,-q_1-1). \tag{30}$$

Here transformation of the label $(A,q_1,q_2)\rightarrow(B,q_1+q_2,-q_1-1)$ is equivalent to clockwise rotation by $\pi/3$ with respect to the origin O. From Eqs. (30) and (19), the boundary conditions in the junction part are derived as

$$F_A^{K'}(r,\theta+\pi/3)=wF_B^K(r,\theta) \tag{31}$$

and

$$F_A^K(r,\theta+\pi/3)=\frac{1}{w}F_B^{K'}(r,\theta). \tag{32}$$

The other boundary conditions

$$F_B^K(r,\theta+\pi/3)=wF_A^{K'}(r,\theta) \tag{33}$$

and

$$F_B^{K'}(r,\theta+\pi/3)=\frac{1}{w}F_A^K(r,\theta) \tag{34}$$

are also obtained in the same way. The same boundary conditions are first discussed by Matsumura and Ando.[15] The difference of Eqs. (31)–(34) from those of Matsumura and Ando by certain factors is due to the difference of the definition of $F_{A,B}^{K,K'}$. From Eqs. (31) and (32), the terms in Eq. (29) for $F_A^{K'}$ and F_B^K are not zero only when $m=3p+2$ (p = integer). Because the open channel $l_j=0$ in the tube parts is spatially uniform along the circumference, it is better fitted to the components with smaller $|m|$ in Eq. (29) than to those with larger $|m|$. So we assume that one can neglect all the terms except those with $p=0$ and $p=-1$ in Eq. (29) (assumption I). Then the wave functions can be written as

$$F_A^{K'}=e^{2i\theta}f_2(z)+e^{-i\theta}f_1(z) \tag{35}$$

and

$$F_B^K=e^{2i\theta}f_2(z)-e^{-i\theta}f_1(z), \tag{36}$$

where

$$f_m(z)=c_mJ_m(z)+d_mN_m(z) \quad (m=1,2). \tag{37}$$

From Eqs. (21) and (24), the other two wave functions $F_B^{K'}$ and F_A^K can be derived from $F_A^{K'}$ and F_B^K as

$$F_B^{K'}=\frac{\epsilon}{|\epsilon|}[-e^{i\theta}\tilde{f}_2(z)+e^{-i2\theta}\tilde{f}_1(z)], \tag{38}$$

$$F_A^K=\frac{\epsilon}{|\epsilon|}[e^{i\theta}\tilde{f}_2(z)+e^{-i2\theta}\tilde{f}_1(z)], \tag{39}$$

where

$$\tilde{f}_1(z)=c_1J_2(z)+d_1N_2(z),$$

$$\tilde{f}_2(z)=c_2J_1(z)+d_2N_1(z). \tag{40}$$

In the above we used the recursion formula of the Bessel functions and Neumann functions.[21] It is easily confirmed that Eqs. (38) and (39) satisfy the boundary conditions Eqs. (33) and (34). The amplitude of the open channel in the tube, which is denoted by α, is obtained from Eq. (26) as

$$\alpha_{j\pm}^K=\frac{1}{2\sqrt{R_j}}\int_{Q_j}^{P_j}dx^{(j)}(e^{i\eta_j/2}F_A^K\pm e^{-i\eta_j/2}F_B^K) \quad (j=5,7) \tag{41}$$

for the K point. The value of $|\alpha|^2$ equals the probability flow as shown in the Appendix. The indices $+$ and $-$ mean the direction along which the electron waves propagate. The path of the integral of Eq. (41) is the straight line P_jQ_j, and the angle with respect to the x axis is denoted by η_j. Equations for $\alpha_{j\pm}^{K'}$ are obtained from Eq. (41) by replacing $\pm$, η_j, and K in the right-hand side with $\mp$, $-\eta_j$, and K', respectively. To simplify the calculation, the integrations in the above are transformed as

$$\int_{Q_j}^{P_j}dx^{(j)}\rightarrow R_j\int_{-(2/3)\pi+\eta_j}^{-\pi/3+\eta_j}d\theta. \tag{42}$$

If variation of the wave function along the radial directions is slow near $r=R_j$, this replacement can be allowed (assumption II). The relation between the amplitudes of the open channel in each tube, $\vec{\alpha}_j={}^t(\alpha_{j+}^K,\alpha_{j+}^{K'},\alpha_{j-}^{K'},\alpha_{j-}^K)$, and the coefficients representing the wave functions in the junction part, $\vec{g}={}^t(c_2,d_2,c_1,d_1)$, are summarized in the following:

$$\vec{\alpha}_j=\frac{1}{2}\sqrt{R_j}P(\eta_j)M\Lambda(\eta_j)L(kR_j)\vec{g}\equiv Y_j\vec{g}, \tag{43}$$

where M is a constant matrix given by[21]

$$M=\begin{pmatrix} -i\frac{\epsilon}{|\epsilon|} & 0 & 0 & -\frac{\sqrt{3}}{2}\frac{\epsilon}{|\epsilon|} \\ 0 & -\frac{\sqrt{3}}{2} & -i & 0 \\ 0 & -\frac{\sqrt{3}}{2} & i & 0 \\ i\frac{\epsilon}{|\epsilon|} & 0 & 0 & -\frac{\sqrt{3}}{2}\frac{\epsilon}{|\epsilon|} \end{pmatrix}. \tag{44}$$

$\Lambda(\eta)$ is a diagonal matrix, where $\Lambda_{1,1}=\Lambda_{3,3}^*=e^{i\eta}$ and $\Lambda_{2,2}=\Lambda_{4,4}^*=e^{2i\eta}$. $P(\eta)$ is defined by Eq. (41) as

$$P(\eta)=\begin{pmatrix} e^{i\eta/2} & e^{-i\eta/2} & 0 & 0 \\ 0 & 0 & e^{-i\eta/2} & -e^{i\eta/2} \\ 0 & 0 & e^{-i\eta/2} & e^{i\eta/2} \\ e^{i\eta/2} & -e^{-i\eta/2} & 0 & 0 \end{pmatrix}. \tag{45}$$

The matrix elements of $L(z)$ are $L_{11}=L_{33}=J_1(z)$, $L_{12}=L_{34}=N_1(z)$, $L_{21}=L_{43}=J_2(z)$, and $L_{22}=L_{44}=N_2(z)$. The other matrix elements of $L(z)$ are zero. From Eq. (43), the relation between $\vec{\alpha}_7$ and $\vec{\alpha}_5$ is given by $\vec{\alpha}_7=Y_7Y_5^{-1}\vec{\alpha}_5$ with the three parameters, kR_7, kR_5, and $\phi\equiv\eta_7-\eta_5$ (angle between $\vec{R}_5$ and $\vec{R}_7$ in the development map). This reads as

$$\begin{pmatrix} \vec{\alpha}_{7+} \\ \vec{\alpha}_{7-} \end{pmatrix}=\begin{pmatrix} t_1,t_2^* \\ t_2,t_1^* \end{pmatrix}\begin{pmatrix} \vec{\alpha}_{5+} \\ \vec{\alpha}_{5-} \end{pmatrix} \tag{46}$$

where

$$t_1 = h_+ \begin{pmatrix} \cos\left(\frac{3}{2}\phi\right), & i\sin\left(\frac{3}{2}\phi\right) \\ i\sin\left(\frac{3}{2}\phi\right), & \cos\left(\frac{3}{2}\phi\right) \end{pmatrix}, \tag{47}$$

$$t_2 = h_- \begin{pmatrix} -\cos\left(\frac{3}{2}\phi\right), & -i\sin\left(\frac{3}{2}\phi\right) \\ i\sin\left(\frac{3}{2}\phi\right), & \cos\left(\frac{3}{2}\phi\right) \end{pmatrix}. \tag{48}$$

The factors h_+ and h_- in Eqs. (47) and (48) are represented by

$$h_\pm = -\tfrac{1}{4}[X_{12}(kR_7,kR_5) \mp X_{21}(kR_7,kR_5)] + \frac{i}{2\sqrt{3}}[X_{11}(kR_7,kR_5) \pm \tfrac{3}{4}X_{22}(kR_7,kR_5)], \tag{49}$$

where

$$X_{i,j}(z_1,z_2) \equiv \sqrt{z_1 z_2}\,\pi\{J_i(z_1)N_j(z_2) - N_i(z_1)J_j(z_2)\}. \tag{50}$$

The considered energy region $|E|<E_c$ corresponds to the region $0<kR_5<2\pi$. It can be easily confirmed analytically that Eq. (46) satisfies the time reversal symmetry and unitarity. The transmission rate per channel denoted by T is calculated from Eq. (46) as

$$T = 1/|h_+|^2 = \frac{2}{\left\{(1/6)\sum_{i=1}^{2}\sum_{j=1}^{2}(3/4)^{i+j-2}X_{i,j}^2(kR_5,kR_7)\right\}+1}, \tag{51}$$

and it gives the conductance σ as $\sigma = 2T$ by Landauer's formula. In Fig. 4, the two solid lines represent the values of T calculated by Eq. (51) for two values of $R_7/R_5 = 0.516, 0.844$. The other lines represent those calculated by the tight-binding model for the (17,17)-(18,21) junction, the (23,8)-(16,22) junction, the (14,5)-(16,22) junction, and the (10,10)-(18,21) junction. The values of $(\phi, R_7/R_5)$ of these four junctions are $(0.014\pi, 0.871)$, $(0.116\pi, 0.844)$, $(0.114\pi, 0.516)$, and $(0.014\pi, 0.512)$, respectively. Since the (18,21) tube and the (16,22) tube have similar radii, their threshold energies are almost the same, i.e., $E_c \simeq 0.16|\gamma|$. It can be seen that the transmission rate by the tight-binding model with fixed R_7/R_5 is almost independent of the angle ϕ and its agreement with Eq. (51) is fairly good.[16] When $E=0$, T becomes consistent with Ref. 15 as

$$T = \frac{4}{(R_5/R_7)^3 + (R_7/R_5)^3 + 2}. \tag{52}$$

Equation (52) reproduces well the numerical results in Ref. 8.

We are now ready to consider the analytical calculation of the wave functions by the effective-mass theory and its relation to the transmission rate. Let us consider the case where the two waves $\vec{\alpha}_{5+}^{(i)}$ $(i=1,2)$ are incident from the thicker tube, and there is no incidence from the thinner tube, i.e., $\vec{\alpha}_{7-} = 0$. As in Sec. II, it is called the positive incidence direction. Whenever the two incident waves are ''orthogonal'' and have the same flow, i.e., $|\vec{\alpha}_{5+}^{(1)}| = |\vec{\alpha}_{5+}^{(2)}|$ and ${}^t\vec{\alpha}_{5+}^{(1)*} \cdot \vec{\alpha}_{5+}^{(2)} = 0$, the corresponding two transmitted waves are also orthogonal and have the same flow, because $t_1^\dagger t_1 = 1/T$. Note that it does not hold generally and is characteristic of the nanotube junction.[22,23]

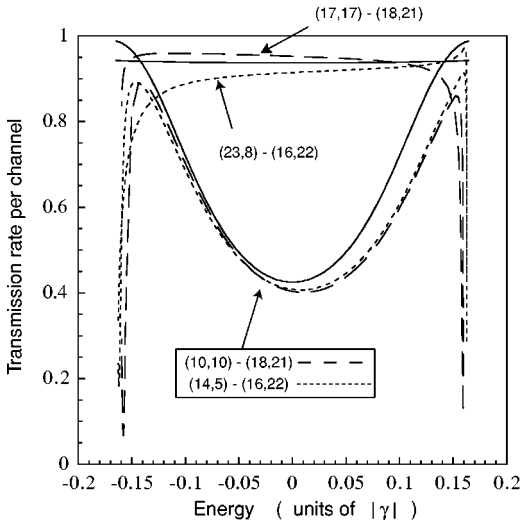

FIG. 4. Transmission rates per channel T as a function of the energy. The solid lines represent those calculated by Eq. (51) for two values of $R_7/R_5 = 0.516, 0.844$. The other lines represent those calculated by the tight-binding model for the (17,17)-(18,21) junction, the (23,8)-(16,22) junction, the (14,5)-(16,22) junction, and the (10,10)-(18,21) junction. The values of $(\phi, R_7/R_5)$ of these four junctions are $(0.014\pi, 0.871)$, $(0.116\pi, 0.844)$, $(0.114\pi, 0.516)$, and $(0.014\pi, 0.512)$, respectively. Since the (18,21) tube and the (16,22) tube have similar radii, their threshold energies are almost the same, i.e., $E_c \simeq 0.16|\gamma|$.

The two orthogonal transmitted waves are represented by $\vec{\alpha}_{7+}^{(1)} = {}^t(a_1, a_2)$ and $\vec{\alpha}_{7+}^{(2)} = {}^t(-a_2^*, a_1^*)$. In the analytical calculation, we use different normalization from that of Sec. II; we take the transmitted waves with unit flow represented by $|\vec{\alpha}_{7+}^{(i)}| = 1$ in this section while the incident waves have unit flows $|\vec{\alpha}_{5+}^{(i)}| = 1$ in Sec. II. Corresponding wave functions denoted by $\psi^{(i)}$ are obtained from $\vec{g} = Y_7^{-1}\vec{\alpha}_7^{(i)}$ and Eqs. (35)–(40). They depend on the phases of a_1 and a_2, but the sum of the squared absolute values of them, $\Psi(s,q_1,q_2) \equiv |\psi^{(1)}(s,q_1,q_2)|^2 + |\psi^{(2)}(s,q_1,q_2)|^2$, does not depend on them and can be determined uniquely. Therefore, we discuss $\Psi = |\psi^{(1)}|^2 + |\psi^{(2)}|^2$ rather than $|\psi^{(1)}|^2$ and $|\psi^{(2)}|^2$ hereafter. Similar discussions are also possible for the negative incidence direction. The $\Psi(s,q_1,q_2)$ in the junction part is derived as

$$\Psi(A,q_1,q_2) = \frac{1}{3r}\sum_{i=1}^{2}(3/4)^{i-1}\left[X_{i,1}^2 + X_{i,2}^2 + 2\frac{\epsilon}{|\epsilon|} \times \cos\left(\theta + \frac{2\pi}{3}(q_1 - q_2)\right)X_{i,1}X_{i,2}\right] \tag{53}$$

and

$$\Psi(B,q_1,q_2)=\frac{1}{3r}\sum_{i=1}^{2}(3/4)^{i-1}\left[X_{i,1}^2+X_{i,2}^2-2\frac{\epsilon}{|\epsilon|}\times\cos\left(\theta-\frac{2\pi}{3}(q_1-q_2)\right)X_{i,1}X_{i,2}\right], \quad (54)$$

where $X_{i,j}=X_{i,j}(kR_t,kr)$ is defined by Eq. (50), with the circumference R_t of the tube into which the electron is transmitted. That is to say, $R_t=R_7$ and $kR_t<kr<kR_5<2\pi$ for the positive incidence direction, while $R_t=R_5$ and $kR_7<kr<kR_t<2\pi$ for the negative incidence direction. Note that the squared wave function Ψ in Eqs. (53) and (54) does depend on the sign of energy unlike the conductance in Eq. (51). Note also that the sign of E is the inverse of that of ϵ, $E/|E|=-\epsilon/|\epsilon|=\pm 1$, because the transfer integral γ is negative. In the right-hand sides of Eqs. (53) and (54), only the third term depends on the angle θ, and it becomes zero as E approaches zero, because $X_{i,i}\to 0$. Therefore the spatial oscillation of Ψ along the θ direction decreases with decreasing $|E|$.

There are several possible choices of the basic translation vectors $\vec{e}_1$ and $\vec{e}_2$. The wave function represented by Eqs. (53) and (54) should be invariant under the change of the choice. When $\vec{e}_1'=\vec{e}_2$ and $\vec{e}_2'=\vec{e}_2-\vec{e}_1$ are used instead of $(\vec{e}_1,\vec{e}_2)$, the corresponding x',y' axes are rotated by $\pi/3$ with respect to original x,y axes as shown in Fig. 1. Correspondingly, the polar coordinates and the labels are transformed as $(r,\theta)\to(r,\theta-\pi/3)$, $(A,q_1,q_2)\to(B,q_1+q_2,-q_1-1)$ and $(B,q_1,q_2)\to(A,q_1+q_2+1,-q_1-1)$. For example, the site indicated by the square symbol near the origin O in Fig. 1 has two different labels, which are $(A,0,0)$ defined by $(\vec{e}_1,\vec{e}_2)$ and $(B,0,-1)$ defined by $(\vec{e}_1',\vec{e}_2')$. Under this transformation of (r,θ) and (s,q_1,q_2), the values of Eqs. (53) and (54) are invariant, that is to say, the results are independent of the way of the labeling, or choice of the vectors, $\vec{e}_1$ and $\vec{e}_2$.

In order to relate the wave function to the transmission rate, the integral of the squared wave function with a fixed r defined below should be introduced:

$$\Phi(kR_t,kr)\equiv\frac{9r}{14\pi}\int d\theta[\Psi(A;r,\theta)+\Psi(B;r,\theta)]=\frac{1}{7}\sum_{i=1}^{2}\sum_{j=1}^{2}(3/4)^{i-1}X_{i,j}^2(kR_t,kr). \quad (55)$$

The third terms in the right-hand side of Eqs. (53) and (54) do not contribute to the integral because q_1-q_2 varies rapidly as a function of θ. We call Φ defined by Eq. (55) a radial norm, hereafter. The radial norm is normalized so that it equals unity at $r=R_t$, i.e., at the exit of the transmitted wave from the junction part. It corresponds to the unit flow of the transmitted wave. Thus at the entrance of the incident wave into the junction part, the radial norm Φ becomes the sum of the amplitude of incident wave $1/T$ and that of the reflected wave $R/T=(1-T)/T$. Accordingly, the transmission rate T can be given by $T=2/(\Phi+1)$, where $\Phi=\Phi(kR_7,kR_5)$ for the positive incidence direction and $\Phi=\Phi(kR_5,kR_7)$ for the negative incidence direction. Figure 5 is the schematic development maps representing this relation

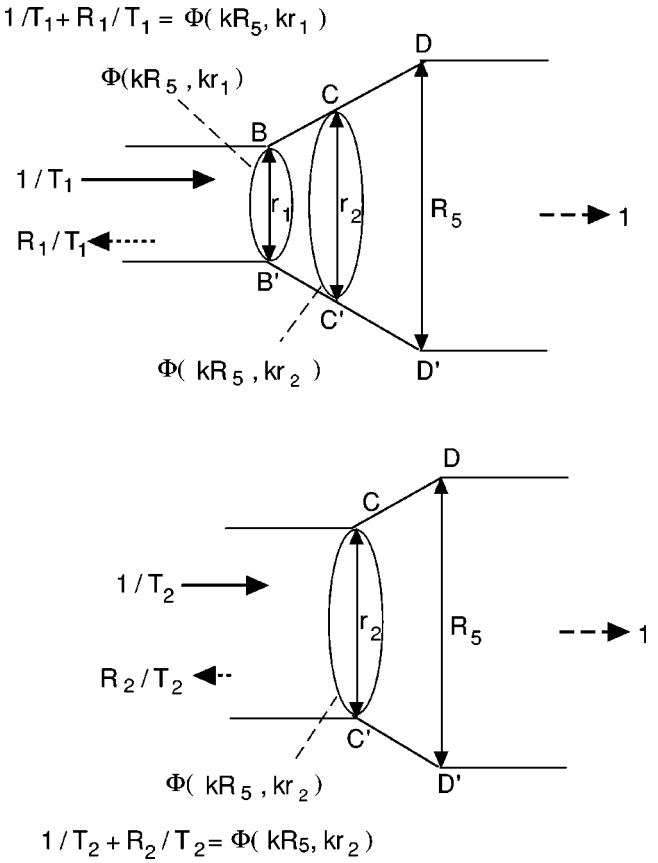

FIG. 5. Schematic development maps representing the relation between the radial norm Φ and the transmission rate T. The thicker tube and the region $CC'D'D$ are common in both the development maps.

between the radial norm and the transmission rate. It shows two junctions with a common thicker tube where the incidence direction is negative. Though they have different transmission rates, T_1 and T_2, their wave functions in the region $CC'D'D$ and the transmitted wave are common. From now on, we call $2/(\Phi+1)$ an ''inverse'' of the radial norm for simplicity. The inverse of the radial norm, $2/[\Phi(kR_7,kR_5)+1]$, $2/[\Phi(kR_5,kR_7)+1]$ and the transmission rate T of Eq. (51) are shown in Fig. 6 as a function of E/E_c by the dashed lines, the dotted lines, and the solid lines, respectively. The curves are shown for the cases $R_7/R_5=0.1$, 0.3, 0.5, 0.7, and 0.9. For each value of R_7/R_5, the inverse of the radial norm is close to the corresponding transmission rate. Therefore, we can discuss the dependence of the transmission rate on R_7/R_5 and $|E|/E_c$ by the radial norm. For this purpose, we show in Fig. 7 the radial norm for the negative incidence directions $\Phi(kR_5,kr)$ as a function of r/R_5 at six energies $E/E_c=kR_5/(2\pi)=0.1$, 0.3, 0.5, 0.65, 0.8, and 1.0. There are four regions of the parameter space that have different characters; they are region I ($1>r/R_5>0.9$; $1>|E|/E_c\geq 0$), region II ($0.9>r/R_5>0.1$; $0.7>|E|/E_c\geq 0$), region III ($0.9>r/R_5>0.1$; $1>|E|/E_c>0.7$), and region IV ($0.1>r/R_5>0$; $1>|E|/E_c\geq 0$). In region I, all the curves in Fig. 7 are very close to unity. It indicates that the corresponding transmission rate is near unity and independent of $|E|/E_c$ as is shown by the lines of $R_7/R_5=0.9$ in Fig. 6. In region II, the radial norm shows monotonic decrease as either r/R_5 or $|E|/E_c$ increases. The decrease with respect to r/R_5 becomes steep as $|E|/E_c$ de-

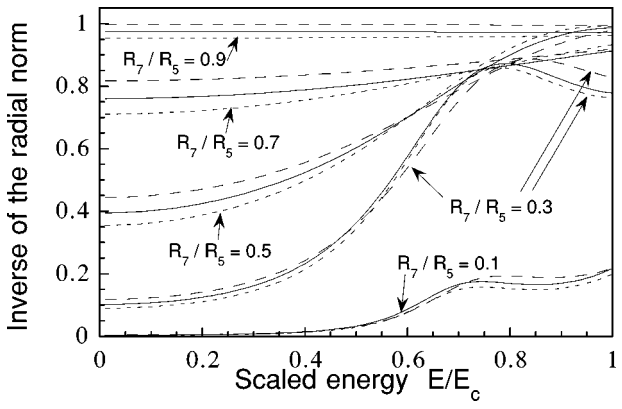

FIG. 6. Inverse of the radial norm for the positive incidence direction defined by $2/[\Phi(kR_7,kR_5)+1]$, and that for the negative incidence direction $2/[\Phi(kR_5,kR_7)+1]$ are shown by the dashed lines and the dotted lines, respectively. Here the radial norm Φ is defined in Eq. (55). Values of R_7/R_5 are 0.1, 0.3, 0.5, 0.7, and 0.9, which are attached to the corresponding lines. The transmission rate per channel calculated by Eq. (51) is shown by the solid lines. It can be seen that the radial norm gives a good estimated value of the transmission rate. The horizontal axis is the energy normalized by E_c. Here $|E|<E_c$ represents the energy region where the channel number is kept to two in both the tube parts.

creases and is almost proportional to $(R_5/r)^3$ when $|E|/E_c$ <0.1. In the region III, however, the radial norm oscillates as r/R_5 increases. When $|E|/E_c=1$ the radial norm has a minimum value at around $r/R_5=0.5$. Correspondingly, the curves of $R_7/R_5=0.5$ are larger than those of $R_7/R_5=0.3$ and 0.7 at $|E|/E_c=1$ in Fig. 6. It oscillates also with increasing $|E|/E_c$ when $0.4>R_5>0.2$. The lines of $R_5/R_7=0.3$ in Fig. 6 show the corresponding oscillations. Around the boundary between region II and III, the decrease of the radial norm with increasing r/R_5 becomes critical as is indicated by the plateau of the curve corresponding to $|E|/E_c=0.65$ in Fig. 7. Correspondingly, all the curves for $R_7/R_5=0.7$, 0.5, and 0.3 in Fig. 6 cross each other around the boundary energy $E/E_c=0.7$. In region IV, the radial norm is nearly proportional to $(R_5/r)^3$ so that the transmission rate is almost proportional to $(R_7/R_5)^3$ at each value of $|E|/E_c$.

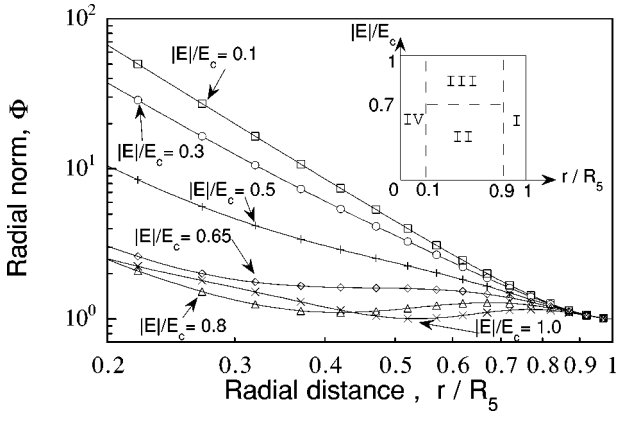

FIG. 7. Radial norm for the negative incidence direction, i.e., $\Phi(kR_5,kr)$ is shown as a function of r/R_5 for six energies, $kR_5/(2\pi)=E/E_c=0.1$, 0.3, 0.5, 0.65, 0.8, and 1.0. Here Φ is defined by Eq. (55). The values of E/E_c are attached to the corresponding lines. Both the axes are represented with a logarithmic scale. The inset shows the regions of the parameter space I, II, III, and IV, which are used in the discussion in the text.

The oscillating behavior in region III can be interpreted as a resonant effect in the following discussion. The wave function propagates both in the radial and the angular direction. Its wavelength along the radial direction is assumed to be close to that of the plane wave obtained from the dispersion relation (25); the wavelength is given by $2\pi/k=R_5|E_c/E|$. When the length of the junction part along the radial direction, i.e., R_5-R_7, coincides with half of the wavelength, the resonance occurs. From this discussion, the condition of the resonance can be easily obtained as

$$R_7/R_5=1-|E_c/2E|. \tag{56}$$

The minimum points of the curves in Fig. 7 in region III correspond to this resonance following approximately condition (56); it comes around $r/R_5=0.5$ when $|E|/E_c=1$ and moves towards the left as $|E|/E_c$ decreases. Around the resonance point, the transmission is almost perfect. Such a resonance, however, does not occur in regions II and IV, since the radial norm becomes divergent when kr approaches zero, i.e., when either $|E|/E_c=kR_5/(2\pi)$ or r/R_5 approaches zero. This divergence comes from the terms of $krN_2^2(kr)$ in Eq. (55), which is almost proportional to $1/(kr)^3$ for small kr values. Accordingly, the transmission rate decays with the power law in proportion to $(R_7/R_5)^3$ in regions II and IV.

IV. COMPARISON BETWEEN THE ANALYTICAL RESULTS AND THE NUMERICAL RESULTS

Since the effective-mass theory is an approximation of the tight-binding model as is explained in Sec. III, its results have to be confirmed by comparing them to the corresponding numerical results from the tight-binding model. In this section, we shall compare the wave function calculated numerically by the tight-binding model $\vec{c}_j$ with those determined analytically by Eqs. (53) and (54), Ψ. For the numerical calculation with the tight-binding model, the conditioned transfer matrix method explained in Sec. II is used. Figure 8 shows the ratio of the corresponding quantities, $(|c_{j,i}^{(1)}|^2+|c_{j,i}^{(2)}|^2)/(T\Psi)$ as a function of radial distance r, for the energy $E=-0.05|\gamma|$. The values are plotted for the (10,10)-(18,21) junction in Fig. 8(a) and for the (17,17)-(18,21) junction in Fig. 8(b). Here the incidence direction is positive; $c_{j,i}^{(k)}$ is caused by the incident waves $\tilde{\psi}_{+k}^R$. For the opposite incidence direction, similar results are also obtained. Here the factor $1/T=|h_+|^2$ defined by Eq. (49) is necessary because the transmitted waves have unit flow in Eqs. (53) and (54), while the incident waves have unit flow in Eq. (13). One can see that the ratio is close to unit, i.e., the coincidence is fairly good. Correspondingly, the transmission rate by the conditioned transfer matrix and that by Eq. (51) coincide well with each other when $|E|<0.14|\gamma|$, as shown in Fig. 4. The deviation from unity becomes smaller as $|E|$ approaches zero, because all the assumptions postulated in deriving Eqs. (53) and (54) are more valid for smaller $|E|$. The deviations, however, become larger when $r\simeq R_5$ in Fig. 8(a), while such a feature is not found in Fig. 8(b). The most natural explanation for this difference is that the deviations are caused mainly by the evanescent waves in the tube parts, and their effects become more significant as their decay lengths along

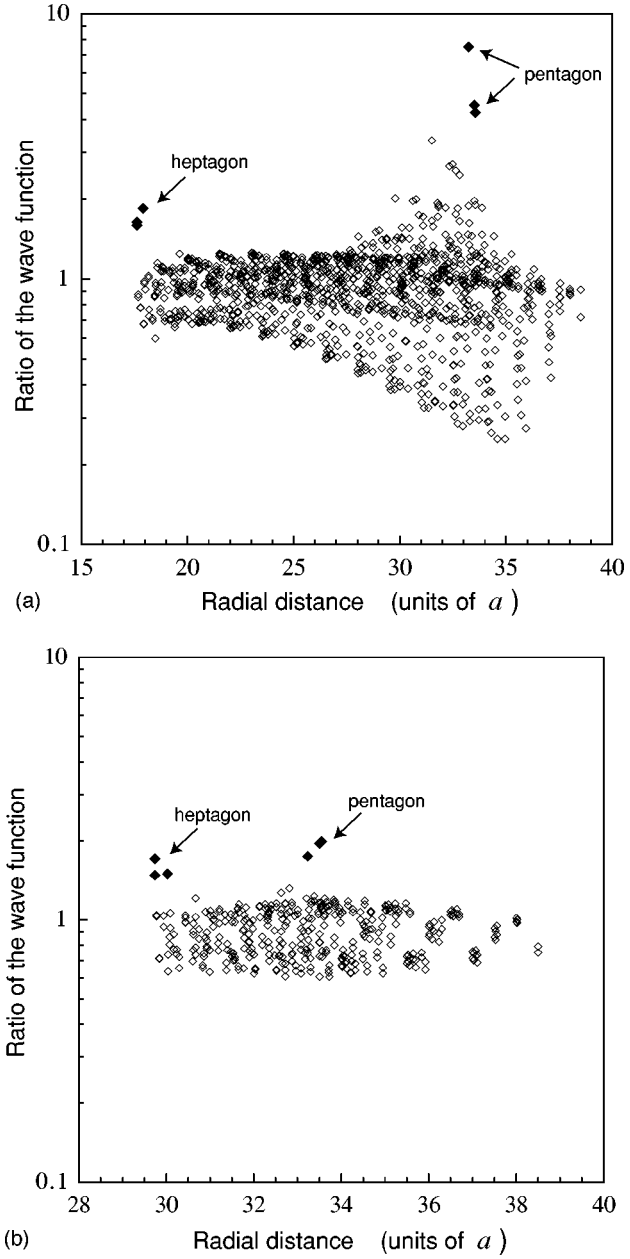

FIG. 8. Comparison between the wave functions $|c_{j,i}^{(k)}|^2$ calculated by the tight-binding model, and Ψ defined by Eqs. (53) and (54); it shows $(|c_{j,i}^{(1)}|^2+|c_{j,i}^{(2)}|^2)/(T\Psi)$ as a function of radial distance, r, (a) for the (10,10)-(18,21) junction and (b) for the (17,17)-(18,21) junction. They are shown for $j=1,2,\ldots,l_1$. The energy is $E=-0.05|\gamma|$. Here $|c_{j,i}^{(k)}|^2$ is caused by the incident waves with positive direction $\tilde{\psi}_{k+}^R$. Here $1/T=|h_+|^2$, defined by Eq. (49), is used for the normalization. The vertical axis is represented with a logarithmic scale. Data at the pentagon and at the heptagon are shown by the closed diamonds indicated by the arrows.

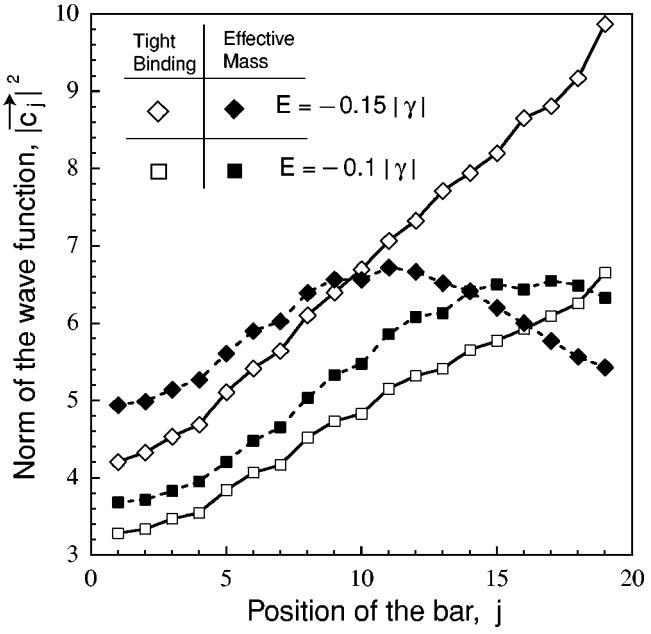

FIG. 9. Norm of the wave function of each bar $|\vec{c}_j|^2$ of the (14,5)-(16,22) junction for $E=-0.1|\gamma|$ and for $E=-0.15|\gamma|$. The horizontal axis represents the position of the bar j shown in Fig. 1. The incidence direction is positive, i.e., from right to left in this figure. The dotted lines with closed symbols and the solid lines with open symbols correspond to those calculated by Eqs. (53) and (54), and those calculated by Eq. (13), respectively. The diamonds and the squares correspond to $E=-0.15|\gamma|$ and $E=-0.1|\gamma|$, respectively.

the tube axis become larger, i.e., as they approach the extended states. The decay length is determined by $l_{5,7}$ in Eq. (27) as $a/\sqrt{(2\pi l_{5,7}a/R_{5,7})^2-\epsilon^2}$. Thus the discussion below is concentrated on the most extended evanescent waves corresponding to $l_{5,7}=\pm 1$. When $R_5 \gg R_7$, as in the case of Fig. 8(a), the evanescent waves in the thicker tube side are much more extended and have larger effects than those in the thinner tube side. As a result, the deviation near the thicker tube side is enhanced compared to that in the other place. On the other hand, when $R_5 \simeq R_7$, as the case of Fig. 8(b), the evanescent waves have similar decay lengths both in the thicker and in the thinner tube side, so that the difference of the deviation does not depend on the sides.

To see the spatial variation of the wave function, we show $|\vec{c}_j|^2$ of the (14,5)-(16,22) junction calculated by the conditioned transfer matrix method and the corresponding quantities determined analytically by Eqs. (53) and (54). The latter is shown by the closed symbols joined by the dotted lines and the former is shown by the open symbols joined by the solid lines for $E=-0.1|\gamma|$ and $E=-0.15|\gamma|$ in Fig. 9. The horizontal axis is the number of the bar j, which is almost proportional to the radial distance r. The diamonds and the squares correspond to $E=-0.15|\gamma|$ and $E=-0.1|\gamma|$, respectively. The squares show that the results by the two methods agree quite well. In contrast to this good agreement, the deviation becomes quite large for the diamonds when $j>10$, i.e., near $r=R_5$; the wave function by the conditioned transfer matrix grows rapidly while the analytical one decays while approaching the thicker tube side. When $|E|\simeq E_c$, the decay lengths of the evanescent waves are very large, and cause the large discrepancies in this way. Corresponding to it, large discrepancies occur also in the transmission rate; only the data by the tight-binding model show the sharp dips near $E=\pm 0.16\gamma$ in Fig. 4. The sharp dip appears approximately when $0.9 \leqslant |E/E_c| \leqslant 1$, i.e., when the decay length of the evanescent waves in the thicker tube are larger than the diameter of the thicker tube. Evanescent waves with their decay length larger than the diameter of the corresponding tube are called ''quasiextended'' evanescent waves hereafter.

In Fig. 8, data corresponding to sites belonging to the pentagonal defect or the heptagonal defect are shown by closed diamonds and indicated by arrows. They indicate that the numerical wave function calculated by the tight-binding model is more localized at the defects than the analytical wave function. Hereafter a localization strength for the θ direction at the defect is defined as the ratio between the numerical norm per site at the defect and that of the bars the defect belongs to. It is calculated by the tight-binding model and represented by $(\Sigma_{i,k\in C_7}|c_{0,i}|^2+|c_{1,k}|^2)(d_0+d_1)/(7|\vec{c}_0|^2+7|\vec{c}_1|^2)$ at the heptagon and $(\Sigma_{i,k\in C_5}|c_{l_1,i}|^2+|c_{l_1+1,k}|^2)(d_{l_1}+d_{l_1+1})/(5|\vec{c}_{l_1}|^2+5|\vec{c}_{l_1+1}|^2)$ at the pentagon with the notation of Sec. II. Here $i\in C_n$ means that site i belongs to the n-membered ring defect. Since the radial distance r in the bar is almost constant, the analytical quantity corresponding to the localization strength has a maximum represented by $\Sigma_{i=1}^2(|X_{i,1}|+|X_{i,2}|)^2/\Sigma_{i=1}^2(|X_{i,1}|^2+|X_{i,2}|^2)\leqslant 2$ that is shown from Eqs. (53) and (54). Therefore, we can say that the wave function is ''quasilocalized'' at the defect when the localization strength is larger than two. The localization strength at the pentagon and that at the heptagon are shown as a function of the energy in Fig. 10(a) and Fig. 10(b), respectively. The calculated junctions are common with those in Fig. 4. The localization strength of the (17,17)-(18,21) junction shows the quasilocalization at both the defects when $|E|\simeq E_c$. It also shows asymmetry with respect to the energy axis; the localization strength at the pentagon (at the heptagon) around $E=-E_c$ ($E=+E_c$) is larger than that around the energy with the opposite sign. These results indicate that the quasilocalization occurs due to the mixing between the quasiextended evanescent waves and the ''defect levels.'' Here the defect levels are defined as n discrete energy levels $-2|\gamma|\cos(2\pi l/n)$ ($l=1,2,\ldots,n$) of an isolated n-membered ring calculated by the tight-binding model.[13] Since the energy region around zero is considered now, the discrete level closest to $E=0$, which is at $E\simeq -0.618|\gamma|$ for the pentagon, and at $E\simeq +0.445|\gamma|$ for the heptagon, is the most important. The closer distance in the energy between the two states enhances their mixing so that it causes the asymmetry with respect to the energy. Since the defect levels are caused by the discreteness of the lattice, these results cannot be reproduced by the effective-mass equation.

As R_7/R_5 decreases, however, the localization strength at the heptagon decreases as is shown in Fig. 10(b); the localization strength at the heptagon of the (10,10)-(18,21) junction is less than that of the (17,17)-(18,21) junction. It can be explained as follows. The energy where the quasiextended evanescent waves in the thinner tube side appear is around $E=\pm E_cR_5/R_7$. As R_7/R_5 decreases, it becomes distant from the considered energy region $|E|<E_c$ so that there is no quasiextended evanescent waves in the thinner tube side. Though there are the quasiextended evanescent waves in the thicker tube side, the defect level of the heptagon cannot be mixed with them due to the large spatial distance between the heptagon and the thicker tube.

By comparing the (23,8)-(16,22) junction with the (17,17)-(18,21) junction or comparing the (14,5)-(16,22) junction with the (10,10)-(18,21) junction in Fig. 10, one can see that increase of ϕ from 0.01π to 0.1π with a fixed

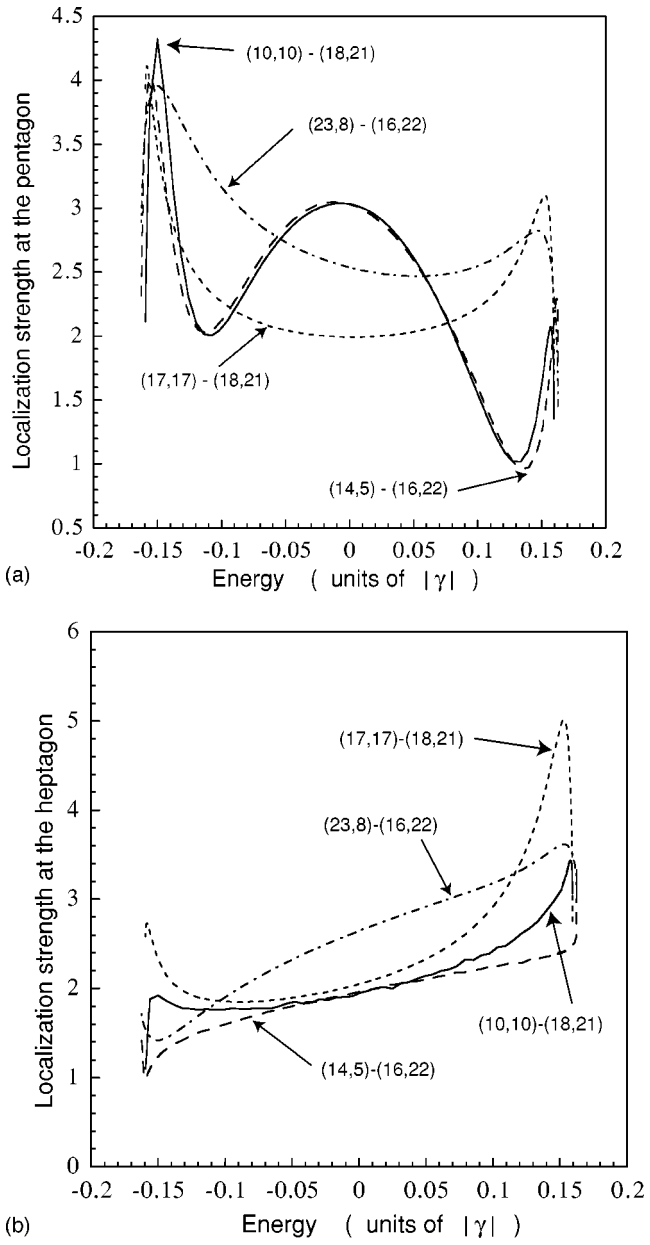

FIG. 10. The localization strength for the θ direction (a) at the pentagon and (b) at the heptagon as a function of the energy. The incidence direction is positive. The solid lines, the dotted lines, the dashed lines, and the dot-dashed lines correspond to the (10,10)-(18,21) junction, the (17,17)-(18,21) junction, the (14,5)-(16,22) junction, and the (23,8)-(16,22) junction, respectively. The calculated junctions are common in this figure and Fig. 4.

R_7/R_5 causes the decrease of the localization strength at the heptagon when $|E|\simeq E_c$. Though the reason for this decrease is not clear yet, the dependence of the localization strength on R_7/R_5 with a fixed ϕ is the same as that explained in the preceding paragraph.

In summary, we can say that the defect level is mixed with the quasiextended evanescent waves only when they are close to each other both in the energy and in the space. The analytical results by the effective-mass approximation are not appropriate when the evanescent waves become quasiextended. The width of this energy region, however, is only about $0.1E_c$. Except for these narrow energy regions, the analytical results are appropriate enough.

V. SUMMARY AND CONCLUSION

In this paper, the junctions connecting the two metallic nanotubes with different circumferences by a pair of the pentagonal and the heptagonal defect are investigated. Both the wave function and the transmission rate are analytically obtained by the effective-mass equation, and a close relation between them is found. To discuss the close relation, the junction part is divided into pieces according to the radial distance r from the thinner tube side and the radial norm of the wave function is evaluated in each piece. The radial norm is determined by the two parameters $|E|/E_c$ and r/R_5, where E_c and R_5 are the threshold energy and the circumference of the thicker tube, respectively. The transmission rate approximates the inverse of the radial norm where r is substituted with the circumference of the thinner tube R_7. From the dependence of radial norm on the two parameters, the parameter space is roughly classified into four regions; region I ($1>r/R_5>0.9$; $1>|E|/E_c\geqslant 0$), region II ($0.9>r/R_5>0.1$; $0.7>|E|/E_c\geqslant 0$), regions III ($0.9>r/R_5>0.1$; $1>|E|/E_c>0.7$), and region IV ($0.1>r/R_5>0$; $1>|E|/E_c\geqslant 0$). In region I, the radial norm is close to unity so that the corresponding transmission is almost perfect, independent of $|E|/E_c$. In region III, the radial norm oscillates near unity as the function of either of the two parameters. The period of the oscillation with respect to r is approximately the same as half wavelength. Here the wavelength is that of the plane wave, i.e., it is obtained from the linear dispersion relation of the monolayer graphite, Eq. (25). Almost perfect transmission due to the resonance occurs in region III when the radial length of the junction part R_5-R_7 coincides with this period, because the corresponding radial norm has a minimum value there. As either $|E|/E_c=kR_5/(2\pi)$ or r/R_5 decreases, however, the radial norm becomes divergent owing to the term $krN_2^2(kr)$ so that the corresponding transmission rate approaches zero. It leads to the power-law decay of the transmission rate proportional to $(R_7/R_5)^3$ in the regions II and IV instead of the resonant feature.[8,9,16]

The nanotube junction can be considered as a combined system, that is the quasi-two-dimensional structure (the junction part) connecting the two quasi-one-dimensional structures (the tube parts). The wave functions in the former part become the 2D waves, whose radial parts are given by the sum of $J_i(kr)$ and $N_i(kr)$ $(i=1,2)$, while they are the 1D plane waves in the tube parts. Nevertheless, the condition of the resonance in region III, Eq. (56), is obtained from the wavelength of the 1D plane wave. It indicates that the four components of the wave function in the junction part, $J_i(kr)$ and $N_i(kr)$ $(i=1,2)$, are combined appropriately to be fitted well with the 1D plane waves in the tube parts. On the other hand, the wave function shows the power-law decay along the radial direction in regions II and IV, because only the component $N_2(kr)$ becomes dominant. In other words, the wave function in the junction part has the two-dimensional character in regions II and IV so that its matching with the plane wave in the tube parts becomes worse than in region III. The wave function in the junction part has different dimensionalities in this way.

The wave functions analytically obtained by the effective-mass equation are compared with that numerically obtained by the tight-binding model. When $|E|$ is very close to E_c, the discrepancies between them becomes considerable. From the dependence of the discrepancies on the radial distance, we speculate that they are mainly due to the evanescent waves with their decay length larger than the diameter of the corresponding tube and they are called quasiextended evanescent waves. The quasiextended evanescent waves cause the sharp dip of the transmission rate, which is absent in the analytical result. When the quasiextended evanescent waves and the energy levels of the n-membered ring defect ($n=5,7$) are close both in the energy and in the space, mixing between them appears as the enhancement of the wave function at the corresponding defect. Nevertheless, coincidence between the analytical wave functions and the numerical ones is fairly good except for the narrow energy region $0.9<|E/E_c|<1$.

The nature of the wave function discussed in this paper can be observed by STM, because STM images reflect the local density of states, which are proportional to the sum of the squared wave functions, Eqs. (53) and (54), in the nanotube junction part. The third terms in Eqs. (53) and (54) give rise to the $\sqrt{3}\times\sqrt{3}$ pattern or the oscillation along the θ direction in the STM images. They fade away as $|E|$ approaches zero. It is also expected that the images depend on the direction of the electronic current ''along the tube axes,'' which should not be confused with that of the tunneling current ''from the STM tip to the sample.'' When $R_5/R_7\gg 1$ and $E\simeq 0$, for example, the squared absolute value of the wave function per atom in the junction part is almost proportional to r^2 when the current flows from the thicker tube to the thinner tube and to $1/r^4$ when it flows along the inverse direction. We expect that these results will promote the investigation of the nanostructures composed by the nanotubes including the nanotube junctions.

ACKNOWLEDGMENTS

We would like to thank H. Matsumura and Professor T. Ando for their useful suggestions. This work was supported in part by the Grant-in-Aid for Scientific Research on the Priority Area ''Fullerenes and Nanotubes'' by the Ministry of Education, Science, and Culture of Japan.

APPENDIX

In the tight-binding model, the probability flow from site i to site j is represented by $\mathrm{Im}(a_i^* H_{i,j} a_j)$, where a_i is the amplitude of the wave function at site i and $H_{i,j}$ is the hopping integral connecting site i and j. Conservation of the flow is guaranteed by the tight-binding equation $Ea_i=\Sigma_j H_{i,j}a_j$. In this paper, $H_{i,j}$ is nonzero only when i and j are nearest neighbors. Consider bars for $j\leqslant 0$ forming to the (m,n) nanotube part in Fig. 1. The amplitudes of wave functions are represented by $r_{j,i}$ and $l_{j,i}$ in the same way as Fig. 2. Then the flow between the $j-1$th bar and the jth bar is represented by $\gamma\Sigma_{i=1}^{m+n}\mathrm{Im}(r_{j-1,i}^* l_{j,i})$. Hereafter the common hopping integral between the nearest neighbors γ and the lattice constant a is chosen to be units. The flow J corresponding to $\vec{\psi}_{K+}$ in Eq. (26) is $J=m\sin(\eta-k_2+2\pi/3)+n\sin(-\eta+k_1+2\pi/3)$, where the first term comes from the m bonds, $1\leqslant i\leqslant m$, and the second term comes from the other n bonds. Here $k_i=\vec{k}\cdot\vec{e}_i$ and η is the angle of $\vec{R}_7$ with

respect to x axis, i.e., $(\cos(\eta),\sin(\eta))=(\sqrt{3}(m+n)/(2R_7),(n-m)/(2R_7))$ and $R_7=\sqrt{m^2+n^2+mn}$. By developing about k_i to first order and using the boundary condition $mk_1+nk_2=0$, the flow J is represented as

$$J=R_7-\frac{\sqrt{3}k_1}{2R_7}m(n+m)+O(k_1^2)+O(k_2^2). \qquad \text{(A1)}$$

The absolute value of the second term in Eq. (A1) is less than $\sqrt{3}\pi R_7/R_5$, since the energy region where the channel number is kept to two, i.e., $kR_5 \leqslant 2\pi$, is considered now. It can be deduced from it that the flow is the almost constant value R_7 when $R_5 \gg \sqrt{3}\pi$. On the other hand, the value of $|\vec{\alpha}_{7+}|^2$ obtained by substituting F's in Eq. (41) with those of Eq. (26) for $\vec{\psi}_{K+}$ is also R_7. Therefore, the normalization factor in Eq. (41) is correct. Difference of the normalization factor does not change the transmission rate, but when the ratio between the wave function in the tube parts and that in the junction part is considered, the normalization factor in Eq. (41) has to be used.

[1] M. Bockrath, D. H. Cobden, P. L. McEuen, N. G. Chopra, A. Zettl, A. Thess, and R. E. Smalley, Science **275**, 1922 (1997); T. W. Ebbesen, H. J. Lezec, H. Hiura, J. W. Benett, H. F. Ghaemi, and T. Thio, Nature (London) **382**, 54 (1996); A. Yu. Kasumov, I. I. Khodos, P. M. Ajayan, and C. Colliex, Europhys. Lett. **34**, 429 (1996); L. Langer, V. Bayot, E. Grivei, J. -P. Issi, C. Van Haesendonck, and Y. Bruynseraede, Phys. Rev. Lett. **76**, 479 (1996); S. J. Tans, M. H. Devoret, H. Dai, A. Thess, R. E. Smalley, L. J. Geerlings, and C. Dekker, Nature (London) **386**, 474 (1997); S. Frank, P. Poncharal, Z. L. Wang, and W. A. de Heer, Science **280**, 1774 (1998).

[2] S. Iijima, Nature (London) **354**, 56 (1991).

[3] R. Saito, M. Fujita, G. Dresselhaus, and M. S. Dresselhaus, Phys. Rev. B **46**, 1804 (1992); J. W. Mintmire, B. I. Dunlap, and C. T. White, Phys. Rev. Lett. **68**, 631 (1992); N. Hamada, S. Sawada, and A. Oshiyama, *ibid.* **68**, 1579 (1992).

[4] M. S. Dresselhaus, G. Dresselhaus, and R. Saito, Solid State Commun. **84**, 201 (1992).

[5] M. P. Anantram and T. R. Govindan, Phys. Rev. B **58**, 4882 (1998); T. Kostyrko, M. Bartkowiak, and G. D. Mahan, *ibid.* **59**, 3241 (1999).

[6] T. Ando and T. Nakanishi, J. Phys. Soc. Jpn. **67**, 1704 (1998).

[7] S. Iijima, T. Ichihashi, and Y. Ando, Nature (London) **356**, 776 (1992).

[8] R. Tamura and M. Tsukada, Phys. Rev. B **55**, 4991 (1991).

[9] R. Tamura and M. Tsukada, Z. Phys. D: At., Mol. Clusters **40**, 432 (1997).

[10] R. Saito, G. Dresselhaus, and M. S. Dresselhaus, Phys. Rev. B **53**, 2044 (1996).

[11] L. Chico, V. H. Crespi, L. X. Benedict, S. G. Louie, and M. L. Cohen, Phys. Rev. Lett. **76**, 971 (1996); L. Chico, L. X. Benedict, S. G. Louie, and M. L. Cohen, Phys. Rev. B **54**, 2600 (1996); T. Nakanishi and T. Ando, J. Phys. Soc. Jpn. **66**, 2973 (1997); J. -C. Charlier, T. W. Ebbesen, and Ph. Lambin, Phys. Rev. B **53**, 11 108 (1996); M. Menon and D. Srivastava, Phys. Rev. Lett. **79**, 4453 (1997); V. Meunier, L. Henrard, and Ph. Lambin, Phys. Rev. B **57**, 2586 (1998); B. I. Dunlap, *ibid.* **49**, 5643 (1994); **50**, 8134 (1994).

[12] R. Tamura and M. Tsukada, Phys. Rev. B **49**, 7697 (1994).

[13] R. Tamura, K. Akagi, M. Tsukada, S. Itoh, and S. Ihara, Phys. Rev. B **56**, 1404 (1997); S. Ihara, S. Itoh, K. Akagi, R. Tamura and M. Tsukada, *ibid.* **54**, 14 713 (1996).

[14] R. Tamura and M. Tsukada, Phys. Rev. B **52**, 6015 (1995); D. L. Carroll, P. Redlich, P. M. Ajayan, J. C. Charlier, X. Blase, A. De Vita, and R. Car, Phys. Rev. Lett. **78**, 2811 (1997).

[15] H. Matsumura and T. Ando, J. Phys. Soc. Jpn. **67**, 3542 (1998).

[16] R. Tamura and M. Tsukada, Phys. Rev. B **58**, 8120 (1998).

[17] K. Akagi, R. Tamura, M. Tsukada, S. Itoh, and S. Ihara, Phys. Rev. B **53**, 2114 (1996); Phys. Rev. Lett. **74**, 2307 (1995).

[18] R. Tamura and M. Tsukada, J. Phys. Soc. Jpn. **68**, 910 (1999).

[19] H. Ajiki and T. Ando, J. Phys. Soc. Jpn. **65**, 2976 (1996).

[20] D. P. DiVincenzo and E. J. Mele, Phys. Rev. B **29**, 1685 (1984).

[21] Though the sign $\epsilon/|\epsilon|$ is absent in Ref. 16 and Ref. 18, it is necessary. This difference does not change the transmission rate, but it must be considered when the wave function is calculated.

[22] When eigenchannel defined in Ref. 23 is used, the transmitted wave corresponding to the incident wave ϕ_i is represented by $\sqrt{T_i}\psi_i$, where each set of $\{\phi_i\}$ and $\{\psi_i\}$ is orthonormalized. In our case, the channel number is two and $T_1=T_2$.

[23] M. Brandbyge, M. S. Sorensen, and K. W. Jacobsen, Phys. Rev. B **56**, 14 956 (1997).