Lattice Dynamics of Silicon with Empirical Many-Body Potentials

E. Roger Cowley
Department of Physics, Camden College of Arts and Sciences, Rutgers, The State University, Camden, New Jersey 08102
(Received 1 April 1988)

Several recently proposed models for the interatomic potential in silicon are used to generate the interatomic force constants. These are then used to calculate the elastic constants and selected normal-mode frequencies. A comparison of the calculated and experimental values provides a test of the models. None of the models tested is completely satisfactory.

PACS numbers: 63.20.Dj, 61.50.Lt

There have recently been proposed several phenomenological potential models for silicon,[1-3] mainly intended for use in simulations of the properties of the condensed phases.[4] The parameters of the models can be adjusted to fit observed structural properties or theoretical values of overlap and bonding energies. Lattice dynamics provides an alternative test of the models, which tests only the properties corresponding to small excursions from equilibrium, but which is an extremely demanding test in that regime. I demonstrate here that none of the models referred to above provide a fully satisfactory description of the lattice vibrations of silicon.

Stillinger and Weber[1] have proposed a potential consisting of a sum of two-body and three-body terms. The two-body potential is written

$$u_2(r_{ij}) = \epsilon f_2(r_{ij}/\sigma),$$

$$f_2(r) = \begin{cases} A(Br^{-p} - r^{-q})\exp[(r-a)^{-1}], & r < a, \\ 0, & r \geq a, \end{cases}$$

where ϵ is chosen to give f_2 depth -1 and σ is chosen to make f_2 a minimum at $r = 2^{1/6}$. With the choice of a given in Ref. 1, the potential is cut off between the first- and second-neighbor distances. The same cutoff is used for the three-body terms which are written

$$u_3(\mathbf{r}_i,\mathbf{r}_j,\mathbf{r}_k) = \sum f_3(\mathbf{r}_i/\sigma,\mathbf{r}_j/\sigma,\mathbf{r}_k/\sigma), \quad f_3(\mathbf{r}_i,\mathbf{r}_j,\mathbf{r}_k) = h(r_{ij},r_{ik},\theta_{jik}) + h(r_{ji},r_{jk},\theta_{ijk}) + h(r_{ki},r_{kj},\theta_{ikj}),$$

where θ_{jik} is the angle between $\mathbf{r}_j$ and $\mathbf{r}_k$ subtended at vertex i. Provided both r_{ij} and r_{ik} are both less than the cutoff a, we have

$$h(r_{ij},r_{ik},\theta_{jik}) = \lambda \exp[\gamma(r_{ij}-a)^{-1} + \gamma(r_{ik}-a)^{-1}](\cos\theta_{jik} + \tfrac{1}{3})^2.$$

Two significant properties of the three-body potential when it is applied to crystalline silicon are that only triplets consisting of an atom and two of its nearest neighbors are counted (and for any such triplet only one h term in f_3 is finite), and that in the ideal structure $\cos\theta_{jik} = -\frac{1}{3}$ for such triplet, so that the three-body terms all vanish. The cohesive energy of the solid thus arises only from the two-body terms, acting only between nearest neighbors, and the three-body triplets include atoms which are, at most, second neighbors.

Tersoff[2] has proposed a more elaborate model potential:

$$E = \sum_i E_i = \tfrac{1}{2} \sum_{i,j \neq i} V_{ij}, \quad V_{ij} = f_c(r_{ij})[A\exp(-\lambda_1 r_{ij}) - B_{ij}\exp(-\lambda_2 r_{ij})], \quad B_{ij} = B_0 \exp(-Z_{ij}/b),$$

$$Z_{ij} = \sum_{k \neq i,j} [w(r_{ik})/w(r_{ij})]^n [c + \exp(-d\cos\theta_{ijk})]^{-1}, \quad w(r) = f_c(r)\exp(-\lambda_2 r).$$

$f_c(r)$ is a cutoff function which again cuts off the potential at a point between the first- and second-neighbor distances in the low-temperature solid. Note that Tersoff defines θ_{ijk} slightly differently than in Ref. 1. The parameters were fitted to the cohesive energy, lattice constant, and bulk modulus of bulk Si, as well as to properties of the dimer and of hypothetical simple cubic and face-centered cubic silicon. The unusual form of Tersoff's potential has the effect of including terms depending on the positions of five atoms, but none are more distant than second neighbors.

Biswas and Hamann[3] expanded the three-body potential in spherical harmonics:

$$E = \tfrac{1}{2}{\sum_{1,2}}' V_2(1,2) + {\sum_{1,2,3}}' V_3(1,2,3),$$

where

$$V_3(1,2,3) = \sum_l C_l \alpha_l(r_{12})\alpha_l(r_{13})P_l(\cos\theta_1).$$

The only approximation is that the dependence on r_{12} and r_{13} separates into a product. They assumed an ex-

ponential form for the α_l functions and a generalized Morse form for the two-body potential. Values of l up to 6 were included in the sum. A notable feature of their potential is its long range. A cutoff of 10 Å is barely adequate. This corresponds to two-body interactions out to fifteen shells of neighbors, and three-body contributions to force constants for all neighbors within 20 Å.

The lattice dynamics of silicon has been treated many times.[5,6] For forces out to second neighbors, the dynamics is described by two force-constant matrices. The force-constant matrix for a nearest neighbor at position $a(\frac{1}{4},\frac{1}{4},\frac{1}{4})$ is

$$-\begin{pmatrix} \alpha_1 & \beta_1 & \beta_1 \\ \beta_1 & \alpha_1 & \beta_1 \\ \beta_1 & \beta_1 & \alpha_1 \end{pmatrix},$$

and for a second neighbor at $a(\frac{1}{2},\frac{1}{2},0)$ it is

$$-\begin{pmatrix} \alpha_2 & \beta_2 & S_2 \\ \beta_2 & \alpha_2 & S_2 \\ -S_2 & -S_2 & \gamma_2 \end{pmatrix}.$$

Given the values of these six constants, the phonon-dispersion curves can be calculated, and also the elastic constants, from the slopes of the acoustic branches at long wavelengths, as

$$c_{11}=(\alpha_1+8\alpha_2)/a,$$

$$c_{12}=(-\alpha_1+2\beta_1-4\alpha_2+8\beta_2-4\gamma_2)/a,$$

$$c_{44}=(\alpha_1+4\alpha_2+4\gamma_2-\beta_1^2/\alpha_1)/a.$$

The last term in c_{44} arises from the internal strain as the structure is sheared.

It is straightforward to differentiate the model of Stillinger and Weber to obtain (at the spacing corresponding to the minimum of f_2)

$$\alpha_1=F+16G,\quad \beta_1=F-8G,$$

$$\alpha_2=\beta_2=G,\quad \gamma_2=-4G,\quad \delta_2=2G,$$

where F arises from the two-body potential,

$$F=(\epsilon/3\sigma^2)[d^2f_2/dr^2]_{r=2^{1/6}},$$

and G arises from the three-body terms,

$$G=(8\epsilon\lambda/27\sigma^2r^2)\exp[2\gamma/(r-a)],\quad r=2^{1/6}.$$

Note that all of the force constants are expressed in terms of two parameters, and that the second-neighbor force-constant matrix is entirely proportional to the single quantity G. This particular two-parameter model for the force constants appears to have been discussed first by Harrison.[7] The values of the six force constants calculated with use of the parameters of Ref. 1 are given in Table I, as model 1.

Tersoff's model does not lend itself to analytic calculation of the force constants, and most of the values were obtained by numerical differentiation of the energy of a suitable cluster. The values are also given in Table I, as model 2. Force constants for the potential of Biswas and Hamann were also calculated numerically, for the nearest 33 shells of neighbors. Values for the first and second shells are listed in Table I, as model 3.

In Table II are shown the elastic constants calculated for the three models, together with the experimental values.[8] For models 1 and 2 the force constants are zero after the first two shells of neighbors, and the expressions given above were used. For model 3, contributions from 33 shells of neighbors were included. The convergence of the sums is very slow. The bulk modulus and c_{11} can also be calculated by homogeneous deformation, and that method gave values within 2% of the values from the method of long waves. For models 1 and 2, the calculations were made for the experimental room-temperature lattice constant of 5.43 Å, but for model 3 the calculations were made for the lattice constant of 5.354 Å, which minimizes the energy.

Stillinger and Weber's model was not directly fitted to any elastic properties, but it gives a reasonable description of all of them (the worst error is 30%, in c_{44}). Tersoff's model was fitted to the bulk modulus, but there is a delicate cancellation between first- and second-neighbor contributions, and the individual elastic constants are less well described. Biswas and Hamann's model gives elastic constants of comparable quality to model 1, but deviates from the experimental values in the prediction of too steep a slope to the dispersion

TABLE I. First- and second-neighbor force constants, in 10^4 dyn/cm.

Force constant	Value Model 1	Model 2	Model 3
α_1	7.314	27.282	14.084
β_1	5.000	27.861	13.578
α_2	0.113	−2.702	−0.597
β_2	0.113	−2.494	−0.412
γ_2	−0.452	3.524	1.512
δ_2	0.226	3.176	1.331

TABLE II. Elastic constants and selected frequencies. Elastic constants are in units of 10^{12} dyn/cm^2 and the normal-mode frequencies are in terahertz.

	Model 1	Model 2	Model 3	Expt.
c_{11}	1.514	1.044	1.71	1.657
c_{12}	0.764	0.597	1.00	0.639
c_{44}	0.564	0.390	0.92	0.796
$\nu_0(q-0)$	17.83	34.43	24.65	15.3
$\nu_{TA}(X)$	5.96	4.81	7.14	4.38

curves while Stillinger and Weber's model is too soft.

To characterize the rest of the dispersion curves, I also give in Table II the values of the zone-center optical frequency ν_0, and the frequency of the zone-boundary transverse acoustic phonon at the X point.[9] The zone-center frequency establishes the overall position of the optical branches. Stillinger and Weber's model gets the value about right, but the other models are much too high. The phonon-dispersion curves of silicon and germanium are characterized by very flat transverse acoustic branches, and the value of the TA(X) frequency tests this behavior. All of the calculated values are too high. It is very unlikely that a model which truncates the forces at two neighbors can give a satisfactory description of the dispersion curves. The flattening of the transverse acoustic branches is very difficult to reproduce unless some kind of electrostatic force is included. Cochran[10] has shown that a shell model for the electrostatic terms leads to long-range effective forces which decay exponentially. The model of Biswas and Hamann is qualitatively in agreement with this.

Perhaps the most surprising result is that three models which all do a reasonable job of describing the structural properties can give such different values for elastic and vibrational properties, especially the Raman frequency, ν_0. This justifies the use of these properties as a filter for the different models. As is often the case, when a model is made more complicated and fitted to additional values, it becomes less reliable in predicting values which were not fitted. The simplest of the three models, that of Stillinger and Weber, gives the best overall description of the lattice dynamics. It is, however, still only partially in agreement with experiment. The use of any of these models to calculate quantities with a significant vibrational contribution should proceed with caution.

[1]F. H. Stillinger and T. A. Weber, Phys. Rev. B **31**, 5262 (1985).

[2]J. Tersoff, Phys. Rev. Lett. **56**, 632 (1986).

[3]R. Biswas and D. R. Hamann, Phys. Rev. Lett. **55**, 2001 (1985).

[4]M. D. Kluge, J. R. Ray, and A. Rahman, J. Chem. Phys. **85**, 4028 (1986).

[5]F. Herman, J. Phys. Chem. Solids **8**, 405 (1959).

[6]N. K. Pope, in *Lattice Dynamics,* edited by R. F. Wallis (Pergamon, New York, 1965).

[7]W. A. Harrison, Ph.D. dissertation, University of Illinois, 1956 (unpublished).

[8]W. A. Harrison, *Electronic Structure and the Properties of Solids* (Freeman, New York, 1980).

[9]G. Nilsson and G. Nelin, Phys. Rev. B **6**, 3777 (1972).

[10]W. Cochran, Proc. Roy. Soc. A **253**, 260 (1959).