AAPT PHYSICS EDUCATION
AMERICAN JOURNAL of PHYSICS

Virial theorem and scale transformations

Peter Kleban

Citation: American Journal of Physics **47**, 883 (1979); doi: 10.1119/1.11619
View online: http://dx.doi.org/10.1119/1.11619
View Table of Contents: http://scitation.aip.org/content/aapt/journal/ajp/47/10?ver=pdfcov
Published by the American Association of Physics Teachers

Virial theorem and scale transformations

Peter Kleban
Department of Physics and Astronomy, University of Maine, Orono, Maine 04469
(Received 5 March 1979; accepted 4 May 1979)

In this article the virial theorem, which is useful in classical, quantum, and statistical mechanics, is considered. Most derivations of this result are of one of two types. Either one starts with the function r·p [or operator 1/2 (r·p + p·r)], or else takes a derivative with respect to the parameter of a scale transformation. It is pointed out that these two methods are closely related, since r·p [or 1/2 (r·p + p·r)] is the generator of scale transformations. This is demonstrated explicitly for classical and quantum-mechanical systems. The pedagogical advantages in pointing out this connection are also examined.

I. INTRODUCTION

The virial theorem is an important result that may be found in textbooks and research papers on classical, quantum, and statistical mechanics. It appears in slightly different guises in these different fields. This might be thought to be the reason why there is apparently no single derivation for it. Almost all of the derivations fall into one of two classes: Either one considers the function $\mathbf{r}\cdot\mathbf{p}$ [or operator $\frac{1}{2}(\mathbf{r}\cdot\mathbf{p}+\mathbf{p}\cdot\mathbf{r})$] and an associated time derivative; or else one differentiates with respect to the parameter characterizing a scale transformation. We point out here that there is a fundamental connection between these two methods of derivation. This follows from the fact that $\mathbf{r}\cdot\mathbf{p}$ [or $\frac{1}{2}(\mathbf{r}\cdot\mathbf{p}+\mathbf{p}\cdot\mathbf{r})$] is itself the generator of a scale transformation. This connection can be found, at least implicitly, in the research literature (see Ref. 1, for example). However it seems to have escaped the notice of textbook writers.[2] It may be exploited to good pedagogical advantage in presenting the virial theorem, as we point out below.

In Sec. II, we review the $\mathbf{r}\cdot\mathbf{p}$ derivation in classical mechanics and show how it is connected to an infinitesimal (Canonical) scale transformation. We then point out some pedagogical reasons for making this connection, and briefly mention how it has helped in discovering some new physics. In Sec. III, we review the corresponding derivation in quantum mechanics and show that it is connected in a very similar way to an infinitesimal (unitary) scale transformation. The extension of these arguments to the case of N particles is also mentioned. Section IV summarizes some other derivations of the virial theorem.

II. r·p DERIVATION

For clarity we begin by considering the $\mathbf{r}\cdot\mathbf{p}$ derivation for the case of a single particle when classical mechanics is valid. The standard derivation of the virial theorem in this case may be found in Goldstein.[3] Consider the function

$$G=\mathbf{r}\cdot\mathbf{p}. \tag{1}$$

Then

$$\frac{dG}{dt}=\frac{p^2}{m}-\mathbf{r}\cdot\nabla V, \tag{2}$$

where V is the potential energy. Now consider trajectories for which r^2 and p^2 are both bounded. Integrating (2) over time from 0 to T we find

$$\frac{1}{T}[G(T)-G(0)]=\frac{2}{T}\int_0^T\frac{p^2}{2m}dt-\frac{1}{T}\int_0^T\mathbf{r}\cdot\nabla V. \tag{3}$$

Since G is bounded, if we let $T\to\infty$ we find

$$2\overline{K_e}=\overline{\mathbf{r}\cdot\nabla V}, \tag{4}$$

where K_e is the particle's kinetic energy, and the bar denotes the time average. Equation (4) is the virial theorem in what is perhaps its simplest manifestation. The right-hand side, divided by 2, is called the virial. Equation (4) is often applied to the case of power law potentials

$$V=V_0r^n,\ V_0=\text{const}, \tag{5}$$

with the result

$$2\overline{K_e}=n\overline{P_e}, \tag{6}$$

where P_e is the potential energy. In particular, for the familiar case of a harmonic oscillator, one has $\overline{K_e}=\overline{P_e}$.

The above derivation raises several questions: Where does the function G come from? To present the theorem as outlined here it is necessary to "pull it out of a hat" with no further explanation. Why does the theorem reduce to such a simple form [Eq. (6)] for power law potentials? [There is also the related question of why the factor of 2 appears on the left-hand side of Eq. (4).] These questions may be answered by pointing out the connection of G with scale transformations. To do this, we first consider a general infinitesimal contact transformation. Any such transformation is characterized by a generating function F that depends on r and p, and an infinitesimal parameter $d\lambda$ that characterizes the "smallness" of the transformation. If we regard the transformation as changing the coordinates $\mathbf{r}$ and $\mathbf{p}$ to new values $\mathbf{r}+d\mathbf{r}$ and $\mathbf{p}+d\mathbf{p}$ (so that we are considering a new point in phase space), we have

$$d\mathbf{r}=d\lambda\{\mathbf{r},F\}$$

and

$$d\mathbf{p}=d\lambda\{\mathbf{p},F\}, \tag{7}$$

where the curly brackets denote Poisson brackets. For an arbitrary function u of $\mathbf{r}$ and $\mathbf{p}$, the change induced by this transformation is then

 0002-9505/79/100883-04$00.50

$$du = d\lambda\{u,F\}. \qquad (8)$$

The function G defined in Eq. (1) and used to derive the virial theorem can be used as a particular choice for F. We find

$$d\mathbf{r} = d\lambda\,\mathbf{r}, \qquad (9a)$$

$$d\mathbf{p} = -d\lambda\,\mathbf{p}. \qquad (9b)$$

Thus G generates a scale transformation. The name refers to one effect of the change. Lengths are multiplied by the scale factor $1 + d\lambda$ as demonstrated by Eq. (9a).[4]

Now we are ready to return to the virial theorem. Note that since G does not explicitly depend on time,

$$\frac{dG}{dt} = \frac{\partial G}{\partial \mathbf{r}} \cdot \dot{\mathbf{r}} + \frac{\partial G}{\partial \mathbf{p}} \cdot \dot{\mathbf{p}} = \{G,H\} = -\{H,G\}, \qquad (10)$$

where H is the Hamiltonian. But by using Eq. (8) for a scale transformation we see that

$$\frac{dG}{dt} = -\frac{dH}{d\lambda}. \qquad (11)$$

Equation (11) is our basic result. It shows that the time derivative of G is equal to minus the derivative of the Hamiltonian with respect to a scale factor. From this point one can complete the derivation of the virial theorem [Eqs. (2)–(4)] unchanged. In this way the virial theorem arises on considering the scale derivative of H and using the fact that it is equal to the time derivative of the function $-G$. In addition, it is now clear on comparing Eqs. (4) and (9b) that the factor 2 multiplying $\overline{K_e}$ arises from the way p^2 changes under the scale transformation. Furthermore, the factor n on Eq. (6) arises because the power law potential of Eq. (5) scales as r^n, and the scale derivative is equivalent to $r(\partial/\partial r)$ for this term, as follows from Eq. (8). Thus the simple form of Eq. (6) is explained very naturally.

The argument just completed is the central point of this work. In what follows, we show that essentially the same reasoning holds for the quantum-mechanical derivation of the virial theorem involving the hermetian operator $\hat{G} = \frac{1}{2}(\mathbf{r}\cdot\mathbf{p} + \mathbf{p}\cdot\mathbf{r})$. We then point out that either of these arguments can be extended to the case of many particles with no difficulty. In this way we complete the task of showing that this type of derivation is closely related to those that proceed directly by differentiating with respect to a length scale. Some examples of this latter type are included, for completeness, in Sec. IV.

There are several pedagogical advantages to connecting the $\mathbf{r}\cdot\mathbf{p}$ or $\frac{1}{2}(\mathbf{r}\cdot\mathbf{p} + \mathbf{p}\cdot\mathbf{r})$ derivations to scale transformations. As already pointed out, it provides a motivation for introducing the function G (or operator $\hat{G}$) in the first place. In addition, the simple form assumed by the theorem for power law potentials is explained very naturally. Finally one can make the connection with derivations that proceed by differentiating with respect to the length scale from the start. These points are important for a complete understanding of the topic. This kind of clarification can help in research as well, and in fact has led to some interesting and very general inequalities for the energy of many-body systems with power law potentials.[5] One essential step in deriving them came in realizing that since the virial theorem arises from a *first* scale derivative of H [Eq. (11)] one could find something new by considering the *second* scale derivative.

III. $\frac{1}{2}$ (r·p + p·r) DERIVATION AND N PARTICLE CASE

The $\frac{1}{2}(\mathbf{r}\cdot\mathbf{p} + \mathbf{p}\cdot\mathbf{r})$ derivation is very similar to the $\mathbf{r}\cdot\mathbf{p}$ case described in Sec. II. For the case of one particle in an external potential $V(r)$ one considers the operator

$$\hat{G} = \frac{1}{2}(\mathbf{r}\cdot\mathbf{p} + \mathbf{p}\cdot\mathbf{r}). \qquad (12)$$

Let the expectation value of $\hat{G}$ in any single quantum state be denoted by

$$\langle \hat{G} \rangle_t = \int \psi^*(r,t)\hat{G}\psi(r,t)d^3r. \qquad (13)$$

Then, following Wannier,[6] one has

$$\frac{d}{dt}\langle \hat{G} \rangle_t = \frac{1}{i\hbar}\langle [\hat{G},H] \rangle, \qquad (14)$$

where the square brackets denote the commatator and H the hamiltonian operator. For $H = p^2/2m + V(r)$, one finds easily that

$$\frac{d}{dt}\langle \hat{G} \rangle_t = 2\langle K_e \rangle_t - \langle \mathbf{r}\cdot\nabla V \rangle_t, \qquad (15)$$

where K_e is the kinetic-energy operator. Taking time averages as in Eq. (3) one finds, provided that $\langle \hat{G} \rangle_t$ is bounded, that

$$2\langle \overline{K_e} \rangle = \langle \overline{\mathbf{r}\cdot\nabla V} \rangle. \qquad (16)$$

Equation (16), in the general case, is analogous to Eq. (4) and is the statement of the virial theorem for one quantum-mechanical particle. For power law potentials, the analog of Eq. (6) follows simply from this result. When the quantum state is an energy eigenstate, $\langle \hat{G} \rangle_t$ = const and Eq. (16) is true without the bars.

To make the connection with a scale transformation, we note that for an infinitesimal unitary transformation generated by the operator $\hat{F}$, the change induced in an operator $\hat{u}$ is given by

$$d\hat{u} = (1/i\hbar)d\lambda[\hat{u},\hat{F}], \qquad (17)$$

where $d\lambda$ is the infinitesimal parameter of the transformation. For $\hat{F} = \hat{G}$, we then have

$$d\mathbf{r} = d\lambda\,\mathbf{r}, \qquad (18a)$$

$$d\mathbf{p} = -d\lambda\,\mathbf{p}. \qquad (18b)$$

In analogy with Eq. (9). Hence $\hat{G}$ is the generator of infinitesimal unitary scale transformations (see Ref. 4 in this regard). Applying this result to Eq. (14) by setting $\hat{u} = H$ and $\hat{F} = \hat{G}$ we have

$$\frac{d}{dt}\langle \hat{G} \rangle_t = \left\langle \frac{dH}{d\lambda} \right\rangle_t. \qquad (19)$$

If the expectation value is taken in an eigenstate of energy E it is not difficult to show that Eq. (19) becomes

$$\frac{dE}{d\lambda} = -\frac{d}{dt}\langle \hat{G} \rangle = 0. \qquad (20)$$

Combining Eq. (19) or (20) with Eq. (15) or (14) then provides the connection of this derivation of the virial theorem with a unitary scale transformation.

The extension of the above results to the case of N particles is straightforward. All that must be done is to make the replacements

$$\mathbf{r} \to \sum_{i=1}^{N} \mathbf{r}_i,$$

$$\mathbf{p} \to \sum \mathbf{p}_i,$$

$$G \to \sum \mathbf{r}_i \cdot \mathbf{p}_i,$$

$$\hat{G} \to \sum \frac{1}{2} (\mathbf{r}_i \cdot \mathbf{p}_i + \mathbf{p}_i \cdot \mathbf{r}_i),$$

$$H \to \sum \frac{p_i^2}{2m} + V(\mathbf{r}_1 \cdots \mathbf{r}_N),$$

$$\mathbf{r} \cdot \nabla V \to \sum \mathbf{r}_i \cdot \nabla_i V,$$

in Eqs. (1)–(20). There is no change in the physics of the arguments.

IV. OTHER DERIVATIONS OF THE VIRIAL THEOREM

In this section we review a few derivations of the virial theorem that proceed by taking a scale derivative from the start, and one that is unrelated to this or the methods considered above.

In classical statistical mechanics, the length scale enters the free energy via the configurational partition function

$$Q = \int \cdots \int \exp[-\beta V(r_1 \cdots r_N)] d^3 r_1 \cdots d^3 r_N. \tag{21}$$

Here $\beta = 1/k_B T$ and the limits of the integration extend over a box of volume V. For convenience we take this to be a cube of side L. To be able to differentiate with respect to L, which sets the length scale, we make the transformation, following Ref. 6,

$$\mathbf{r}_i = L\mathbf{x}_i. \tag{22}$$

In statistical mechanics it is more meaningful to differentiate with respect to V than L. In particular one has

$$p = kT \frac{\partial \ln Q}{\partial V},$$

where p is the pressure. Since

$$\frac{d \ln L}{d \ln v} = \frac{1}{3}$$

and

$$Q = L^{3N} \int \cdots \int \exp[-\beta V(Lx_1 \cdots Lx_N)] \times d^3 x_1 \cdots d^3 x_N \tag{23}$$

one finds easily that

$$p = \frac{NkT}{V} - \int \cdots \int \sum \mathbf{r}_i \cdot \nabla_i V \exp(-\beta V) d^3 r_1 \cdots d^3 r_N \times (3VQ)^{-1}. \tag{24}$$

If the potential energy V consists of a sum of spherically symmetric pair potentials $\phi(|r_{ij}|)$ and the system is in a fluid state, it is convenient to introduce the pair correlation function $g(r)$ in Eq. (24) with the result

$$P = \frac{NkT}{V} - \frac{2}{3} \pi \frac{N}{V} \int_0^\infty g(r) \phi'(r) r^3 dr. \tag{25}$$

This same result can be obtained via the $\mathbf{r}\cdot\mathbf{p}$ derivation. To do this it is necessary to introduce some sort of "wall" potential-energy term to contain the system and give rise to pressure, use the equipartition theorem to replace $\overline{K_e}$ by $\frac{3}{2}NkT$, and define $g(r)$ and p via time averages, which are equal to the ensemble average values in Eq. (25) by the ergodic theorem.

In quantum mechanics, we can derive the virial theorem by using the Schrödinger representation. Assume, for definiteness, that we have a single particle in a cube of side L acted on by a potential $V(r)$. Then, if it is in an eigenstate of energy E,

$$\frac{dE}{dL} = \left\langle \frac{dH}{dL} \right\rangle.$$

Now

$$H = -(\hbar^2/2m)\nabla_r^2 + V(r).$$

If we use the transformation[7]

$$\mathbf{r} = L\mathbf{x}$$

we see that (compare Ref. 4)

$$H = -(\hbar^2/2mL^2)\nabla_x^2 + V(Lx)$$

and

$$L\frac{dE}{dL} = \left\langle L\frac{dH}{dL} \right\rangle = \left\langle -2\left(-\frac{\hbar^2}{2m}\nabla_r^2\right) + rV'(r) \right\rangle. \tag{26}$$

For a system in a bound state we can let $L \to \infty$ and Eq. (26) reduces to Eq. (16) (without the bars, which are unnecessary for the expectation value in an energy eigenstate). If the system is in its ground state,

$$L\frac{dE}{dL} = 3V\frac{dE}{dv} = -3pV,$$

where p is the pressure, so Eq. (26) becomes

$$pV = \frac{2}{3}\langle K_e \rangle - \frac{1}{3}\langle rV'(r) \rangle, \tag{27}$$

which is the zero temperature version of Eq. (25) for one particle. The extention of Eq. (27) to many particles is straightforward. One may also extend Eq. (27) to the many-particle case at finite temperature. This can be done by using the unitary scale transformation to take the scale derivative of the average energy at fixed entropy. Details of this method are given in Refs. 1 and 5.

Finally we mention, for completeness, perhaps the simplest way of obtaining the virial theorem: by integrating Newton's second law.[8] Consider one particle for convenience. Then

$$m\ddot{\mathbf{r}} = -\nabla V(r)$$

so

$$m\mathbf{r} \cdot \ddot{\mathbf{r}} = -\mathbf{r} \cdot \nabla V. \tag{28}$$

Taking a time average of Eq. (28) and integrating the left-hand side by parts results in

$$\frac{1}{T}[m\mathbf{r} \cdot \dot{\mathbf{r}}]_0^T - \frac{1}{T}\int_0^T m\dot{r}^2 dt = -\frac{1}{T}\int \mathbf{r} \cdot \nabla V dt,$$

so that

$$2\overline{K_e} = \overline{\mathbf{r} \cdot \nabla V(r)}. \tag{29}$$

ACKNOWLEDGMENTS

We would like to acknowledge useful discussions with Professors K. Brownstein and N. Comins.

Note added in proof. Some closely related conclusions have recently been published by H. A. Gersch, Am. J. Phys. **47**, 555 (1979).

[1]R. D. Puff and N. S. Gillis, Ann. Phys. **46**, 364 (1968).

[2]We are in fact unaware of any work where our basic result, Eq. (11) appears explicitly.

[3]H. Goldstein, *Classical Mechanics* (Addison-Wesley, Reading, MA, 1950).

[4]Taken as a condition on the generator of the transformation, Eq. (9a) is not sufficient to uniquely determine it. The generator G also has the property of leaving any function of **r** and **p** with dimensions of angular momentum (such as G itself) invariant. This choice appears more natural in quantum mechanics (see Sec. III) where a scale transformation should change **r** by the factor $d\lambda$ and leave h invariant. The latter requirement implies immediately that $d\mathbf{p} = -d\lambda\mathbf{p}$ since $\mathbf{p} = (\hbar/i)\nabla$.

[5]P. Kleban and R. D. Puff, Phys. Lett. A **31**, 340 (1970).

[6]G. H. Wannier, *Statistical Physics* (Wiley, New York, 1966).

[7]Our argument here is essentially the same as that given by R. P. Feynman, *Statistical Mechanics* (Benjamin, Reading, MA, 1972).

[8]See E. Fermi, *Notes on Thermodynamics and Statistics* (University of Chicago, Chicago, 1966). This method is the one used in the original publication of the virial theorem in English by R. Clausius, Philos. Mag. Ser. 4, XL, 123 (1870).