Estimating the Memory Time Induced by Exciton-Exciton Scattering

V. M. Axt and T. Kuhn
Institut für Festkörpertheorie, Westfälische-Wilhelms Universität, Wilhelm-Klemm-Strasse 10, 48149 Münster, Germany

B. Haase,* U. Neukirch,† and J. Gutowski
Institut für Festkörperphysik, Universität Bremen, P.O. Box 330440, D-28334 Bremen, Germany
(Received 3 September 2003; published 14 September 2004)

It is shown that lower bounds for the effective memory time induced by two-pair correlations can be estimated by monitoring changes of the shape of excitonic four-wave-mixing spectra. Experimentally we demonstrate a memory time of at least 540 fs for a ZnSe single quantum well. Microscopic calculations reveal that this lower bound is not sharp. Interactions retarded by more than 800 fs are shown to influence the dynamics, reflecting the presence of a long time tail in the memory kernel.

DOI: 10.1103/PhysRevLett.93.127402 PACS numbers: 78.67.De, 42.65.Sf, 71.35.Gg, 71.45.Gm

Recent studies have clearly revealed that correlations arising from the scattering of two excitons by no means can be considered as small corrections to the nonlinear response of semiconductor nanostructures [1–18]. For example, the huge increase of the gain up to 5000 observed in microcavities after insertion of large numbers of quantum wells [19] has been explained by exciton-exciton scattering correlations [16]. These correlations also contribute significantly to a change of the sign of the optical Stark shift [11]. Without them the excitonic emission in four-wave-mixing (FWM) signals is overestimated by more than an order of magnitude compared with the biexcitonic emission [13]. In view of their decisive role for these and other features it is worthwhile to have a closer look on how two-pair correlations affect nonlinear optical signals. In fact, the storing of part of the excitation in two-pair correlations effectively supplies a kind of reservoir which can act back on the optically visible polarizations as long as it prevails [17]. Unlike a thermal reservoir no equilibrium or quasiequilibrium is established. In contrast, the properties of the reservoir are to some extent tunable by varying the excitation conditions. The possibility of acting back at later times is equivalent to providing a memory for the polarization. Typically, past values of the polarization have a lesser influence on the present dynamics the longer they are retarded. This implies that there is a characteristic time scale indicating how long a retardation can be before no further impact on the current dynamics is detectable. We shall refer to this time scale as the *effective memory time* of the scattering process. Another way to express the same basic physics is to state that exciton-exciton scattering processes are not instantaneous as assumed, e.g., in Boltzmann type theories that make use of the Markov approximation [20]. Instead, they have a finite duration of the order of the memory time. Although the existence of a memory structure has been shown to be of utmost importance [10,11,13–16] and in spite of the principal physical relevance of the retarded nature of interaction processes, memory times induced by exciton-exciton scattering have not yet been experimentally determined for any semiconductor structure.

In this Letter it is explained how a lower bound for the memory time induced by two-pair correlations can be derived by monitoring the delay time dependence of the line shape of FWM spectra. By using our scheme we are able to obtain for the first time an experimental measure indicating for how long retarded parts of the exciton-exciton interaction can have a substantial effect on the present dynamics.

The theoretical basis for our analysis is a microscopic density matrix theory where the polarization is expressed in terms of single-pair transition amplitudes defined as $Y^j_\ell \equiv \langle \hat{Y}^j_\ell \rangle = \langle d_j c_\ell \rangle$, c_j (d_j) being the destruction operators for electrons (holes) in Wannier states. Two-pair correlations are described by the correlated two-pair amplitudes $\bar{B}^{jj'}_{\ell\ell'} \equiv \langle \hat{Y}^j_\ell \hat{Y}^{j'}_{\ell'} \rangle - Y^j_\ell Y^{j'}_{\ell'} + Y^{j'}_\ell Y^j_{\ell'}$. A closed set of equations for Y and $\bar{B}$ is obtained by invoking the *dynamics controlled truncation* (DCT) scheme [21] and concentrating on coherent dynamics at low densities. Points to note are (i) the DCT equations in the coherent limit for Y and $\bar{B}$ yield a representation of the semiconductor response that is exact up to third order in the applied laser fields provided that incoherent bath interactions, e.g., with phonons, can be neglected [21]; (ii) $\bar{B}$ accounts for all types of two-pair coherences including biexcitonic transitions and transitions to correlated exciton-exciton scattering states; (iii) the equations are nonperturbative with respect to the Coulomb interaction and thus avoid divergences that occur in the perturbative approach based on the second Born approximation [9]. We shall use the memory kernel representation of the equations in the form derived in Ref. [3], where the effect of $\bar{B}$ on Y is captured by a memory kernel, thus highlighting most clearly the memory structure induced by two-pair correlations. Written in the basis of pair wave functions (i.e., wave functions for bound and unbound excitons) the

resulting equations read [3]

$$\hbar[-i(\partial_t+\gamma)+\omega_\alpha]y_\alpha=\sum_p[\mathcal{M}_\alpha^p-\sum_{\bar\alpha\bar\alpha'}\mathcal{B}_{\alpha\bar\alpha}^{\bar\alpha' p}y_{\bar\alpha}^* y_{\bar\alpha'}]E^p + Q_\alpha^{\mathrm{Coul}}, \quad (1)$$

where $\hbar\omega_\alpha$ is the energy of an optical transition to the pair state α and y_α is the corresponding component of Y. $\mathcal{B}$ are blocking matrix elements while $\mathcal{M}_\alpha^p$ represent the components of the dipole matrix elements in the direction of the polarization vector of the pth applied pulse with amplitude E^p. γ is a phenomenological damping accounting for residual interactions with the environment. Finally, the Coulomb induced sources Q_α^{Coul} are given by

$$Q_\alpha^{\mathrm{Coul}}=\int_0^\infty d\tau\sum_{\alpha'\bar\alpha\bar\alpha'}\overline{K}_{\alpha\bar\alpha}^{\alpha'\bar\alpha'}(\tau)y_{\bar\alpha}^*(t)y_{\alpha'}(t-\tau)y_{\bar\alpha'}(t-\tau), \quad (2)$$

$$\overline{K}_{\alpha\bar\alpha}^{\alpha'\bar\alpha'}(\tau)=(\mathcal{V}_{\mathrm{H}\alpha\bar\alpha}^{\alpha'\bar\alpha'}-\mathcal{V}_{\mathrm{F}\alpha\bar\alpha}^{\alpha'\bar\alpha'})\delta(\tau)+\mathcal{K}_{\alpha\bar\alpha}^{\alpha'\bar\alpha'}(\tau), \quad (3)$$

where the kernel $\mathcal{K}(\tau)$ takes into account the memory and correlation effects resulting from transitions to two-pair states. $\mathcal{K}(\tau)$ can be determined without having to integrate Eq. (1) from an integral equation [3]. $\mathcal{V}_{\mathrm{H}}$ and $\mathcal{V}_{\mathrm{F}}$ denote the Hartree-Fock Coulomb matrix elements. They represent the mean field (MF) Coulomb nonlinearities. The MF parts are instantaneous interactions and thus do not establish a finite memory time. Only the two-pair correlations lead to retarded interactions coupling the transition density y at time t to its values at all earlier times. Apart from γ, there are no phenomenological parameters in the theory.

To connect the memory with the behavior of FWM spectra we shall analyze a standard two-pulse FWM signal emitted in $2\mathbf{k}_2-\mathbf{k}_1$ direction where $\mathbf{k}_1$ and $\mathbf{k}_2$ denote the directions of the incident pulses. We will show that Eq. (1) implies that under suitable excitation conditions the shape of FWM spectra measured at negative delays $\Delta t<0$, i.e., the $\mathbf{k}_2$ pulse arrives first, can substantially depend on Δt only provided that $|\Delta t|\lesssim\tau_{\mathrm{mem}}+t_{\mathrm{pulse}}$, where τ_{mem} (t_{pulse}) is the effective temporal extent of the memory kernel (pulses). To this end we consider a selective excitation of the $1s$ exciton by short laser pulses and assume low excitation densities such that a perturbative analysis is justified. According to Eq. (1) the linear response contains under these conditions only the $1s$ component which for a pulse approaching the sample at $t=t_0$ is given by $y_{1s}^{\mathrm{lin}}(t-t_0)=e^{-(i\omega_{1s}+\gamma)(t-t_0)}f(t-t_0)$. For ultrafast excitations $f(t-t_0)$ takes, apart from a constant prefactor, the form of the step function $\theta(t-t_0)$. When allowing for a finite pulse duration t_{pulse}, necessary to ensure the selective excitation, the step function is smoothed such that $f(t-t_0)$ is essentially a constant f_0 if $t-t_0\gtrsim t_{\mathrm{pulse}}/2$ and zero for $t-t_0\lesssim -t_{\mathrm{pulse}}/2$. For negative delays larger than the pulse overlap we can concentrate exclusively on the Coulomb source Q_α^{Coul} in Eq. (1) because then the blocking terms do not contribute to the FWM response. Considering a $\mathbf{k}_1$ ($\mathbf{k}_2$) pulse approaching the sample at $t=0$ ($t=\Delta t$) we can evaluate the source Q_α^{Coul} for the FWM signal in $2\mathbf{k}_2-\mathbf{k}_1$ direction by substituting in Eq. (2) the $\mathbf{k}_1$ contribution $y_{1s}^{\mathrm{lin}*}(t)$ for $y_{\bar\alpha}^*(t)$ and the $\mathbf{k}_2$ contributions $y_{1s}^{\mathrm{lin}}(t-\tau-\Delta t)^2$ for $y_{\alpha'}(t-\tau)y_{\bar\alpha'}(t-\tau)$. Because of the factor $y_{1s}^{\mathrm{lin}*}(t)$ the source Q_α^{Coul} is essentially nonzero only for times $t\gtrsim -t_{\mathrm{pulse}}/2$. In addition, the effective temporal extent of the kernel $\overline{K}(\tau)$ is limited by the memory time τ_{mem} which restricts the integration in Eq. (2) to $\tau\lesssim\tau_{\mathrm{mem}}$. Combining both restrictions we effectively have to consider only arguments of $y_{1s}^{\mathrm{lin}}(t-\tau-\Delta t)$ limited to $t-\tau-\Delta t\gtrsim -t_{\mathrm{pulse}}/2-\tau_{\mathrm{mem}}-\Delta t$. Assuming now a sufficiently large negative delay satisfying $\Delta t\lesssim -(\tau_{\mathrm{mem}}+t_{\mathrm{pulse}})$ we obtain $t-\tau-\Delta t\gtrsim t_{\mathrm{pulse}}/2$. This implies that for all times t and τ which give a noticeable contribution to the Coulomb source in Eq. (2) we can replace $y_{1s}^{\mathrm{lin}}(t-\tau-\Delta t)^2$ by $f_0^2e^{-2(i\omega_{1s}+\gamma)(t-\tau-\Delta t)}$. Thus, for all delays $\Delta t\lesssim -(\tau_{\mathrm{mem}}+t_{\mathrm{pulse}})$ the delay dependence of $Q_{1s}^{\mathrm{Coul}}(t,\Delta t)$ reduces to a factor $e^{2\Delta t(i\omega_{1s}+\gamma)}$. Because of the linear structure of Eq. (1) this also holds for the polarization. Obviously such a t-independent factor has no influence on the shape of the corresponding spectra. We conclude that changes of the spectral shape of the FWM emission after selective excitation of the $1s$ resonance in the low density regime are not possible provided the delay is negative and larger than $\tau_{\mathrm{mem}}+t_{\mathrm{pulse}}$. Thus a lower estimate for the memory time τ_{mem} can be obtained by identifying a negative delay $\Delta t_{\mathrm{change}}$ where substantial line shape changes occur and using the inequality $\tau_{\mathrm{mem}}>\Delta t_{\mathrm{change}}-t_{\mathrm{pulse}}$.

Dramatic changes of the $1s$ line have indeed been reported in Ref. [13], where FWM spectra for a few negative values of Δt were recorded and a good agreement with calculations based on Eq. (1) was found. The largest changes occurred for colinear excitation where a sharp dip was observed. The occurrence of the dip can be qualitatively understood from the competition of emissions with different spectral widths whose respective weights and relative phases are determined by the two-pair memory kernel [13]. It was shown that all line shape deformations disappear when the contribution of exciton-exciton scattering states is disregarded [13]. To estimate the memory time we have performed a more systematic study of the delay time dependence of the spectra taking care that the assumptions made in the above analysis are fulfilled.

Our experiments were performed at a temperature of 1.8 K on a high quality 4.8 nm ZnSe single quantum well. The sample has been characterized in Ref. [22] revealing a hh-lh splitting of $E_{\text{lh-hh}}=20$ meV and binding energies for the exciton and the biexciton of 25 and 6.6 meV, respectively. To ensure a selective excitation of the $1s$

transition we have used pulses tuned ~4 meV below the 1*s* line of 160 fs full-width-half-maximum (FWHM) of the field amplitude. To accurately calibrate the zero of the time delay stray light from individual pulses has been cross correlated with a reference pulse. By recording the intensity dependence of the FWM signals it was verified that the $\chi^{(3)}$ theory is applicable. FWM spectra measured in the vicinity of the 1*s* resonance are shown in Fig. 1(a) for a series of negative delays. Clearly, varying the delay induces dramatic changes of the spectra. The shape evolves from an almost symmetric line centered at the 1*s* resonance for $\Delta t = 0$ fs into a strongly asymmetric line with more oscillator strength at higher energies. A further increase of $|\Delta t|$ to larger negative values produces a sharp dip. Then the shape again becomes asymmetric but now with more spectral weight on the low energy side. Even when concentrating on the most striking features, substantial changes of the line shape can be attested at least until $\Delta t \approx -700$ fs. The pulse duration can be estimated from the FWHM of the field amplitude to $t_{\text{pulse}} \approx 160$ fs. According to our theory this implies a lower bound for the memory time of $\tau_{\text{mem}} > 540$ fs which can be estimated solely from the measured data.

To illustrate the applicability of our theory we have solved numerically Eq. (1) for conditions matching the experiment. The resulting FWM spectra [cf. Fig. 1(b)] show a good overall agreement with the experiment. Substantial line shape changes can be read off from the theoretical curves, too, at least until $\Delta t \simeq -700$ fs implying the same estimate for a lower bound of the memory time of 540 fs as obtained from the experiment.

It should be noted that our sample exhibits only a moderate inhomogeneous broadening (IB). Otherwise a more careful analysis would be needed to unambiguously extract the exciton-exciton memory time. This can be seen from a common IB model where a Gaussian ensemble of gap energies is considered with correlated shifts of single- and two-pair energies. For low intensities and not too long pulses this results essentially in a multiplication of the unbroadened FWM emission by a factor $\sim \exp[-\sigma(t - 2\Delta t)^2] = \exp(-\sigma t^2)\exp(-4\sigma\Delta t^2) \times \exp(4\sigma t\Delta t)$, where σ determines the width of the IB. The factor $\exp(-\sigma t^2)$ obviously cannot induce a Δt dependence as it does not contain Δt; $\exp(-4\sigma\Delta t^2)$ only scales the emission without changing the shape, but $\exp(4\sigma t\Delta t)$ describes a linear increase of the line width for $\Delta t < 0$. Thus, IB can lead to Δt dependent line shape modifications typically consisting of changes of the width. In our case IB is small and apart form its contribution to the width no impact on the line shapes can be discerned.

The estimate derived above provides a lower bound for the memory time. To get more insight into the relation between the memory time and line shape changes we have performed calculations where we have truncated the upper limit of the memory integral in Eq. (2) at a given cutoff time τ_{cut}. Figure 2(a) [2(b)] shows corresponding results for $\tau_{\text{cut}} = 800$ fs [$\tau_{\text{cut}} = 600$ fs]. First, we note that Figs. 2(a) and 2(b) confirm numerically that changes of the line shape occur only for delay times that stay in the above derived bounds. However, it might have been expected that taking the cut at $\tau_{\text{cut}} = 800$ fs, i.e., at a value much larger than the estimate for the memory time, should have little effects on the results of Fig. 1(b). In fact, we still see characteristic changes of the line shape in Fig. 2(a). But, compared with Fig. 1(b) they occur at earlier delays and also deviate from the shapes found in Fig. 1(b). This indicates that interactions actually retarded by more than 800 fs do still have a noticeable

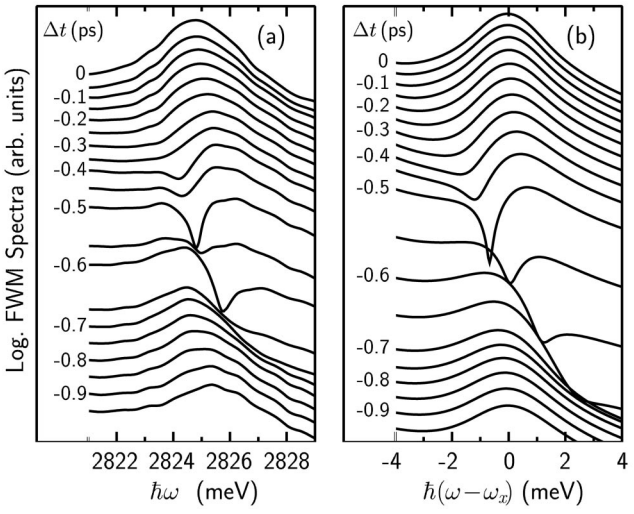

FIG. 1. Excitonic FWM spectra after colinear excitation at negative delay times from $\Delta t = 0$ ps (top) in steps of $\delta\Delta t = -0.05$ ps. (a) Experiment (b) theory. Parameters: damping $\gamma = 1/(2.4\text{ ps})$; pulse area $A = 0.02$; pulse duration 160 fs FWHM of the field amplitude; inhomogeneous broadening = 1 meV. The zero of energy in (b) is at the exciton resonance.

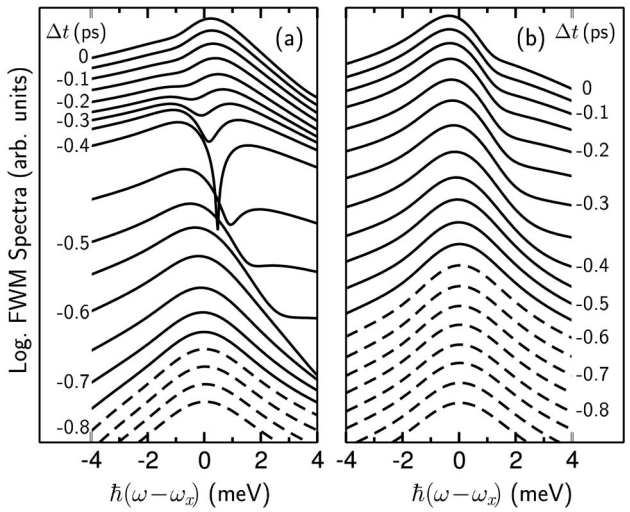

FIG. 2. Excitonic FWM spectra calculated for conditions as in Fig. 1(b) but with a memory time cut to (a) $\tau_{\text{cut}} = 0.8$ ps, (b) $\tau_{\text{cut}} = 0.6$ ps. Dashed curves correspond to delays larger than the truncated memory time.

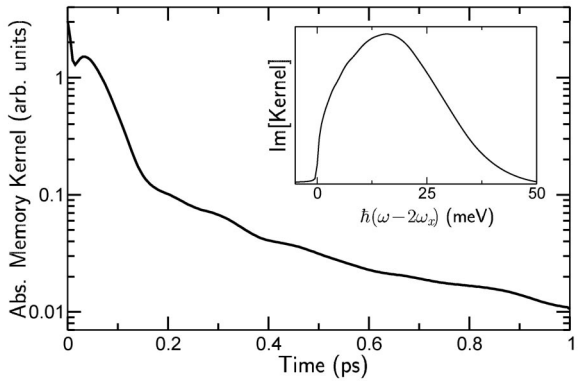

FIG. 3. Absolute value of the memory kernel for exciton-exciton scattering. Shown is the component relevant for couplings of 1s transition amplitudes after colinear excitation. Inset: frequency representation (imaginary part).

influence. Moreover, the strong line deformations almost completely disappear when the memory time is further restricted by setting the cutoff to $\tau_{\mathrm{cut}} = 600$ fs [cf. Fig. 2(b)]. This demonstrates directly that rather long memory times are needed for explaining the observed line shape deformations.

In the theory we have direct access to the memory kernel and can see how the estimates obtained from the analysis of experimentally accessible data compare with the actual temporal profile of the kernel. Plotted in Fig. 3 is the absolute value of the memory kernel for exciton-exciton scattering. Shown is the component relevant for couplings of 1s transitions after colinear excitation. The kernel exhibits a strongly nonexponential decay with superimposed modulations. The decay can be roughly divided into two stages: a rapid initial drop by about 2 orders of magnitude within $\sim$170 fs and a subsequent decay on a longer time scale. Obviously, it is the slow component which determines the effective memory time deduced from the line shape analysis. And indeed, this component decays by 1 order of magnitude within a time of the order of 800 fs in good agreement with our conclusions derived from Fig. 2. In frequency space the kernel exhibits a broad bandlike continuum (cf. inset of Fig. 3) with a tendency that the fast (slow) temporal component is mostly determined by the spectral behavior above (below) the maximum. Estimates for materials other than ZnSe can be obtained by noting that written in excitonic Bohr units and assuming strong confinement, the memory kernel essentially depends only on the ratio of the electron and hole mass [23]. Thus for materials with identical mass ratios the memory times and exciton periods should have the same ratio provided in Bohr units the well widths and dampings γ coincide.

In conclusion, we have demonstrated that a lower bound for the effective memory time induced by two-pair correlations can be obtained by monitoring changes of the line shape of FWM spectra when the delay time is varied. The measurements reveal, in very good agreement with theory, that dramatic line shape deformations do occur for rather long delays. From our data we are able to demonstrate a memory time of at least 540 fs for a ZnSe single quantum well. To the best of our knowledge this is the first experimental estimate for this fundamentally important quantity. A detailed numerical analysis suggests that the strict lower bound of 540 fs is not sharp. Instead, interactions retarded by more than 800 fs noticeably influence the spectra.

We thank J. Nürnberger and W. Faschinger (University of Würzburg) for providing the samples and the DFG for support through Grant No. Ne 525/4.

*Current address: Bosch GmbH, Stuttgart, Germany.
†Current address: Corning Inc., Corning, NY, USA.

[1] P. Kner *et al.*, Phys. Rev. Lett. **78**, 1319 (1997).
[2] P. Kner *et al.*, Phys. Rev. Lett. **81**, 5386 (1998).
[3] V. M. Axt, K. Victor, and T. Kuhn, Phys. Status Solidi B **206**, 189 (1998).
[4] T. Östreich, K. Schönhammer, and L. J. Sham, Phys. Rev. B **58**, 12 920 (1998).
[5] G. Bartels *et al.*, Phys. Rev. Lett. **81**, 5880 (1998).
[6] T. Aoki *et al.*, Phys. Rev. Lett. **82**, 3108 (1999).
[7] B. Haase *et al.*, Phys. Rev. B **59**, 7805(R) (1999).
[8] A. L. Smirl *et al.*, Phys. Rev. B **60**, 8267 (1999).
[9] N. H. Kwong *et al.*, Phys. Rev. Lett. **87**, 027402 (2001); Phys. Rev. B **64**, 045316 (2001).
[10] D. Birkedal, J. Shah, and L. N. Pfeiffer, Phys. Rev. B **60**, 15 585 (1999).
[11] C. Sieh *et al.*, Phys. Rev. Lett. **82**, 3112 (1999).
[12] T. V. Shahbazyan, N. Primozich, and I. E. Perakis, Phys. Rev. B **62**, 15 925 (2000).
[13] V. M. Axt, B. Haase, and U. Neukirch, Phys. Rev. Lett. **86**, 4620 (2001).
[14] S. Savasta, O. Di Stefano, and R. Girlanda, Phys. Rev. B **64**, 073306 (2001).
[15] T. Aoki *et al.*, Phys. Rev. B **64**, 045212 (2001).
[16] S. Savasta, O. Di Stefano, and R. Girlanda, Phys. Rev. Lett. **90**, 096403 (2003).
[17] D. S. Chemla and J. Shah, Nature (London) **411**, 549 (2001).
[18] S. A. Hawkins *et al.*, Phys. Rev. B **68**, 035313 (2003).
[19] M. Saba *et al.*, Nature (London) **414**, 731 (2001).
[20] T. Kuhn, in *Theory of Transport Properties of Semiconductor Nanostructures*, edited by E. Schöll (Chapman and Hall, London, 1998), pp. 173.
[21] V. M. Axt and A. Stahl, Z. Phys. B **93**, 195 (1994); Z. Phys. B **93**, 205 (1994).
[22] B. Haase *et al.*, J. Cryst. Growth **214**, 856 (2000).
[23] This requires that it is justified to concentrate on a single sublevel and that the electron/hole confinement functions do essentially not dependent on the masses, as applies, e.g., for confinement potentials with infinite barriers.