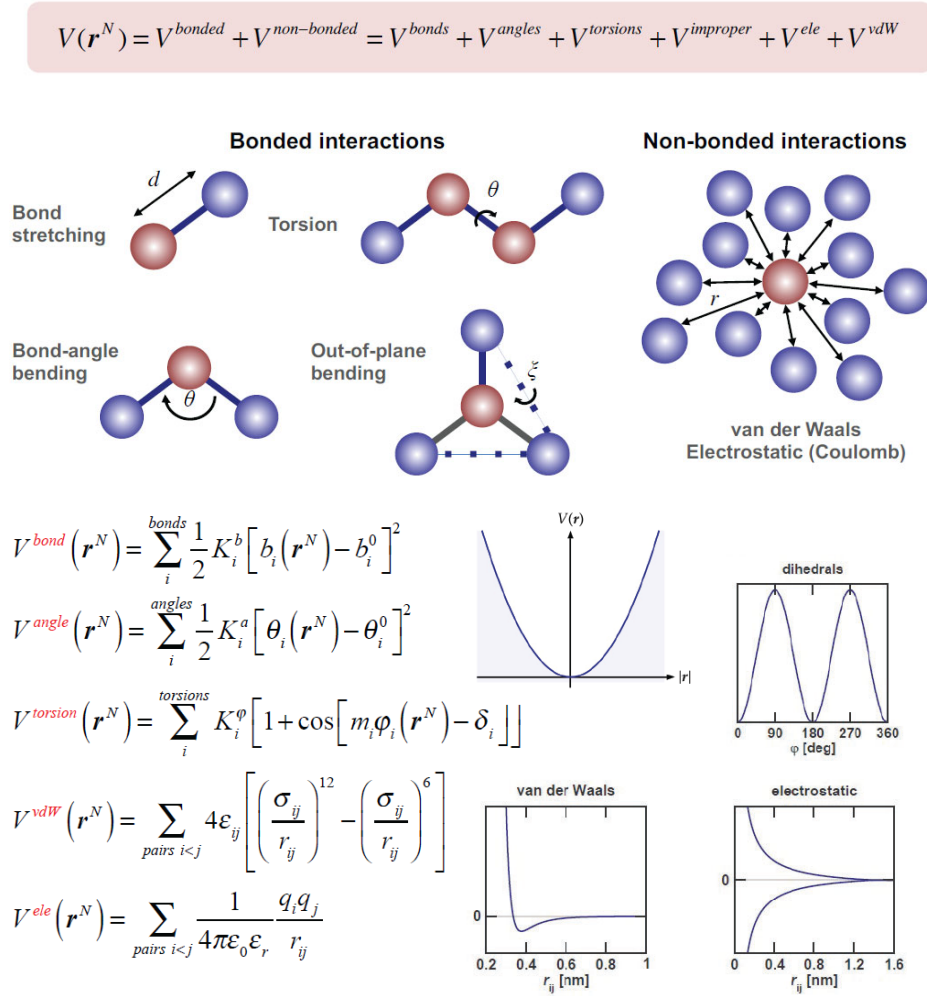# SAMPLE SOLUTION

.

1 Concepts and Understanding (S2022.1)

Note that the answers given below are more detailed (for didactical purposes!) than what was actually expected from the students during the examination.

a. In the grand-canonical ensemble, the system has the following boundary conditions: (*i*) it is open (*i.e.* allowed to exchange molecules with its surroundings); (*ii*) it is isochor (of fixed volume, *i.e.* not allowed to exchange volume work with its surroundings); (*iii*) it is diatherm (*i.e.* allowed to exchange heat with its surroundings). The three independent variables normally used to specify the state of this ensemble (restricting the discussion to a one-component one-phase system) are the chemical potential μ (intensive), the volume V (extensive), and the temperature T (intensive). The corresponding dependent quantities are the number of particles N (extensive), the pressure P (intensive), and the internal energy U (extensive).

b. Covalent (bonded) terms are: bond stretching, bond-angle bending, dihedral torsion and improper-dihedral distortion (out-of-plane bending). Their roles and typical functional forms are shown in the figure below (which also includes the two non-bonded types, electrostatic and van der Waals interactions).

$$V(r^N) = V^{bonded} + V^{non-bonded} = V^{bonds} + V^{angles} + V^{torsions} + V^{improper} + V^{ele} + V^{vdW}$$

Bonded interactions

Bond stretching

Torsion

Bond-angle bending

Out-of-plane bending

Non-bonded interactions

van der Waals
Electrostatic (Coulomb)

$$V^{bond}\left(r^N\right) = \sum_i^{bonds} \frac{1}{2} K_i^b \left[b_i\left(r^N\right) - b_i^0 \right]^2$$

$$V^{angle}\left(r^N\right) = \sum_i^{angles} \frac{1}{2} K_i^a \left[\theta_i\left(r^N\right) - \theta_i^0 \right]^2$$

$$V^{torsion}\left(r^N\right) = \sum_i^{torsions} K_i^\varphi \left[1 + \cos\left[m_i \varphi_i\left(r^N\right) - \delta_i \right] \right]$$

$$V^{vdW}\left(r^N\right) = \sum_{pairs\ i<j} 4\varepsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}}\right)^6 \right]$$

$$V^{ele}\left(r^N\right) = \sum_{pairs\ i<j} \frac{1}{4\pi\varepsilon_0\varepsilon_r} \frac{q_i q_j}{r_{ij}}$$

c. Finite-size effects: A typical simulation box (microscopic simulated sample) is smaller than the interaction range of the particles. Surface effects: If a liquid droplet (or protein) is simulated in vacuum, there are many more molecules (or atoms) perturbed by the surface compared to

the number of unperturbed molecules in the bulk (or atoms in the middle) compared to a macroscopic sample. Thus, the perturbed molecules/atoms will influence the properties of the system. Periodic boundary conditions remove the surface to vacuum by mimicking an infinite periodic system (infinite pseudo-crystal), thereby removing the surface effects.

2 Fundamental Equations (S2022.2)

Note that the answers given below are more detailed (for didactical purposes!) than what was actually expected from the students during the examination. Watch out not to overlook the requirements "**explain the meaning** of all the involved symbols" and "**state the SI units** of these quantities".

a. The one-dimensional time-dependent Schrödinger equation in terms of a coordinate x (SI unit: m) and time t (unit: s) reads

$$\hat{\mathcal{H}}(t)\,\Psi(x,t) = i\hbar\,\frac{\partial \Psi(x,t)}{\partial t}\ ,$$

where $\hbar = (2\pi)^{-1}h$, h being the Planck constant (SI unit: J·s), Ψ is the time-dependent wavefunction (complex, SI unit of $\mid \Psi(x,t)\mid^2$: m^{-1}), and $\hat{\mathcal{H}}$ the Hamiltonian operator (SI unit: J). The latter operator is given by

$$\hat{\mathcal{H}}(t) = -\frac{\hbar^2}{2m}\,\nabla^2 + \mathcal{V}(x,t)\ ,$$

where m is the mass (SI unit: kg) and $\mathcal{V}(x,t)$ is the potential energy (SI unit: J). The wavefunction is normalized as

$$\int dx\ \mid \Psi(x,t)\mid^2 = 1 \quad \forall t\quad .$$

Its interpretation is that $dx\mid \Psi(x,t)\mid^2$ is the probability at time t of finding the particle at position x (within an infinitesimal width dx). If the Hamiltonian operator $\hat{\mathcal{H}}$ is time-independent, the time-dependent Schrödinger equation can be simplified to the time-independent one

$$\hat{\mathcal{H}}\psi(x) = E\,\psi(x)\ ,$$

where ψ is the time-independent (stationary) wavefunction (complex, SI unit of $\mid \psi(x)\mid^2$: m^{-1}) and E is the energy (SI unit: J). This is an eigenvalue equation which typically admits a discrete set of real solutions for E. The Hamiltonian $\hat{\mathcal{H}}$ is time-independent when the system is isolated from fluctuating external influences (*e.g.* other molecules moving around, fluctuating fields, electromagnetic waves, ...). Static influences (*e.g.* time-independent external field) are, however, allowed (provided they are included into the Hamiltonian).

b. When using direct counting, the free-energy difference $\Delta G_{U \to F}$ is calculated based on averages $\langle ... \rangle$ over the plain MD (unbiased) trajectory, as

$$\Delta G_{U \to F} = -\beta^{-1}\ln\frac{\langle f(\mathbf{r})\rangle}{\langle 1 - f(\mathbf{r})\rangle}\ ,$$

where $\beta = (k_B T)^{-1}$, with k_B the Boltzmann constant (SI units: J·K^{-1}), T is the absolute temperature (SI units: K), and f is the indicator function for the folded state (unitless). In practice, this amounts to counting the number of configurations where the peptide is folded (numerator in the log) and unfolded (denominator in the log), and calculating $\Delta G_{U \to F}$ from the resulting ratio (equilibrium constant). To improve the accuracy of the calculation, one may rely on the use of a biasing potential $\mathcal{V}_b(\mathbf{r})$ during the MD sampling. This is called umbrella sampling. The desired properties determining the choice of an appropriate biasing potential are: (*i*) the two states should be sampled in approximately equal amounts in the biased sampling (equalization of the relative free energies); (*ii*) the two states should interconvert often during the simulation (reduction of the barrier between the states). When using umbrella sampling, the free-energy difference $\Delta G_{U \to F}$ is calculated based on averages $\langle ... \rangle_b$ over the

biased MD trajectory, as

$$\Delta G_{U \to F} = -\beta^{-1} \ln \frac{\langle f(\mathbf{r}) \exp[+\beta \mathcal{V}_b(\mathbf{r})] \rangle_b}{\langle (1 - f(\mathbf{r})) \exp[+\beta \mathcal{V}_b(\mathbf{r})] \rangle_b} \ .$$

The principle is the same, but the frames are ascribed a relative weight of $\exp[+\beta \mathcal{V}_b(\mathbf{r})]$ (rather than one in direct counting), to correct for the effect of the biasing. The more a configuration was favored (disfavored) during the biased sampling, the more it will be down-weighted (up-weighted) in the averaging.

3 Derivations (S2022.3)

a. The normalization condition reads

$$\int_0^\infty dx\, p(x) = c \int_0^\infty dx\, \exp(-\alpha x) = c\,\left[(-\alpha^{-1})\exp(-\alpha x)\right]_0^\infty = c\,\alpha^{-1}\ .$$

It is satisfied with $c = \alpha$, so that the normalized probability distribution reads

$$p(x) = \alpha\,\exp(-\alpha x)\ .$$

The expectation value (mean) is given by

$$\mu = \int_0^\infty dx\, x\, p(x) = \alpha \int_0^\infty dx\, x\, \exp(-\alpha x)$$
$$= \alpha\,\left\{\left[x\,(-\alpha^{-1})\exp(-\alpha x)\right]_0^\infty - \int_0^\infty dx\,(-\alpha^{-1})\exp(-\alpha x)\right\} = \alpha\,\left\{0 + \alpha^{-2})\right\} = \alpha^{-1}\ ,$$

where the third equality follows from integration by parts (integrand $f \cdot g'$ with $f = x$ and $g' = \exp(-\alpha x)$). So, the mean is equal to the inverse of the decay parameter. If the unit of x is ns, the unit of $p(x)$ is ns^{-1}, the unit of $c = \alpha$ is ns^{-1} as well, and the unit of μ is ns. It is a common mistake to think that a probability distribution is "unitless" or with "arbitrary units". If the distribution is normalized, the normalization condition (see above) imposes that the product of the units of dx and $p(x)$ is unitless (like the resulting "one"). So, the units of $p(x)$ are inverse to the units of x.

b. The classical canonical partition function is defined as

$$Z = h^{-1} \int dx\, dp\, \exp[-\beta \mathcal{H}(x,p)]\ ,$$

where $\beta = (k_B T)^{-1}$, k_B being the Boltzmann constant, and the Planck constant h is introduced for compatibility with the corresponding quantum-mechanical expressions (resulting in a unitless Z). This evaluates to

$$Z = h^{-1}\left\{\int dx\, \exp[-(1/2)\beta k x^2]\right\} \cdot \left\{\int dp\, \exp[-(1/2)\beta m^{-1}p^2]\right\}$$
$$= h^{-1}\left(\frac{2\pi}{\beta k}\right)^{1/2}\left(\frac{2\pi}{\beta m^{-1}}\right)^{1/2} = h^{-1}\beta^{-1}2\pi\left(\frac{m}{\beta k}\right)^{1/2} = \frac{k_B T}{h\nu}\ ,$$

where the second equality follows from the quoted Gaussian-integral result (with a factor two included, as the given result was for the $[0,\infty]$ interval whereas here, we integrate over $[-\infty,\infty]$). The probability distribution $P(x,p)$ is defined as

$$P(x,p) = (hZ)^{-1}\exp[-\beta\mathcal{H}(x,p)]\ .$$

It has the units of an inverse action (inverse of energy × time) as appropriate for the probability distribution of a two-dimensional quantity (x,p) that has the units of an action. Inserting the expressions for the harmonic oscillator, this gives

$$P(x,p) = \frac{\nu}{k_B T}\cdot\left\{\exp[-(1/2)\beta k x^2]\right\}\cdot\left\{\int dp\, \exp[-(1/2)\beta m^{-1}p^2]\right\}$$

which is factorizable as $P_x(x)\cdot P_p(p)$. The two probability distributions are Gaussians, so that

the $\langle x^2 \rangle^{1/2}$ and $\langle p^2 \rangle^{1/2}$ are simply given by the corresponding standard deviations, *i.e.*

$$\langle x^2 \rangle^{1/2} = (\beta k)^{-1/2} = \left(\frac{k_B T}{k} \right)^{1/2} \quad \text{and} \quad \langle p^2 \rangle^{1/2} = \left(\frac{\beta}{m} \right)^{-1/2} = (m k_B T)^{1/2} \ .$$

The mean-square displacement $\langle x^2 \rangle^{1/2}$ decreases upon increasing k (tighter oscillator → smaller amplitude of motion), increases upon increasing T (more kinetic energy → larger amplitude of motion), and is independent of m. The mean-square momentum $\langle p^2 \rangle^{1/2}$ increases upon increasing m (higher mass → more momentum), increases upon increasing T (more kinetic energy → more momentum), and is independent of k. The expression for $\langle p^2 \rangle^{1/2}$ is also compatible with the equipartition principle, which implies (for a system with one degree of freedom) that

$$\langle \mathcal{K} \rangle = \frac{\langle p^2 \rangle}{2m} = \frac{1}{2} k_B T \ .$$

4 Explicit Calculations (S2022.4)

a. Using the equation for the harmonic potential energy one has

$$V(\theta_o \pm \Delta\theta; \theta_o) = \frac{1}{2}k_\theta(\Delta\theta)^2.$$

Inserting the numerical values $k_\theta = 0.15$ kJ·mol·deg^{-2} and $\Delta\theta = 10$ deg, one gets $V = 7.5$ kJ·mol^{-1}. In comparison, the thermal energy RT at $T = 300$ K is 2.49 kJ·mol^{-1}. So, a bond-angle change of 10 deg corresponds to an energy increase by about $3RT$. As a result, the typical bond-angle fluctuations in molecular dynamics simulations are on the order of 10 deg or less (RT would be $10/3^{1/2} \approx 5.8$ deg).

b. We use the ideal-gas equation of state

$$PV = nRT$$

We know that $V_{end}/V_{beg} = 1.05$ (volume expansion by 5 %) Because the number of mols and the pressure are constant, it follows that $T_{end}/T_{beg} = 1.05$, where T should be an absolute temperature. Denoting T_o the absolute zero on the Celcius scale, one has to solve $(T_o + 40)/(T_o + 25) = 1.05$, which has the solution $T_o = -275$ °C.

5 Algorithms and Implementation (S2022.5)

a. A possible code for the function LennardJonesEnergy is given below

```
double LennardJonesEnergy (int N, double x[], double y[], double z[],
                           double sig, double eps) {
    int i,j; // particle indexes
    double r2, ri6; // for r^2 and 1/r^6
    double fac = pow(sig,6); // for sig^6
    double ene = 0;
    for ( i=0; i<N; i++ ) {
        for ( j=i+1; j<N; j++ ) {
            r2 = (x[j]-x[i])*(x[j]-x[i])
                + (y[j]-y[i])*(y[j]-y[i])
                + (z[j]-z[i])*(z[j]-z[i]);
            ri6 = 1.0/(r2*r2*r2);
            ene += ri6 * ( fac * ri6 - 1.0 );
        }
    }
    return 4.0*eps*fac * ene;
}
```