

PARALELNÍ PROGRAMOVÁNÍ V ÚLOHÁCH MOLEKULÁRNÍ DYNAMIKY

V. Pelikán, P. Hora, A. Machová¹

Úvod

Modely, které jsou zkoumány v úlohách molekulární dynamiky, obsahují stále větší a větší množství jednotlivých atomů a dnes se již stává poměrně běžným, že takovéto modely dosahují mohutností řádově stovek milionů. Numerické simulace na takovýchto modelech by byly nemyslitelné bez masivního nasazení tzv. *clusterů* a jejich programování s využitím některé z paralelních programovacích technik, jakou je kupříkladu *MPI - Message Passing Interface*. Nedílnou součástí vyhodnocení těchto simulací je vizualizace dosažených výsledků, která sama osobě představuje rovněž ne zcela jednoduchý technický problém. První zkušenosti z 3D simulací metodou molekulární dynamiky byly získány v α -železe [1] a v systému Fe-Cu [2]. Tento příspěvek prezentuje některé výsledky z této oblasti, získané klasickým a paralelním programováním.

Stručný popis problému

Máme *bcc* krystal o nějakém „rozumném“, dostatečně velkém počtu atomů železa s příměsí mědi. Vazby mezi jednotlivými atomy jsou popsány vícečásticovým meziatomovým potenciálem převzatým z [3]. Podíváme-li se podrobněji na níže uvedené funkce (1), které zmíněný potenciál definují, zjistíme, že existuje jistá maximální vzdálenost, za kterou jsou všechny tyto funkce rovny nule a každý atom je tudíž zpravidla ovlivňován pouze svými nejbližšími 14 sousedy. „Prostorová“ derivace tohoto potenciálu udává sílu, která působí na jeden každý atom v daném časovém okamžiku [4]. Tyto síly vstupují do soustavy pohybových rovnic, které jsou řešeny metodou centrálních diferencí [5] a udávají tak polohu jednoho každého atomu v následujícím časovém kroku. Velmi zhruba je možno celý model přirovnat k soustavě hmotných kuliček, navzájem pospojovaných pružinami. Povšimněme si ještě, že tato „soustava pružin“ není v čase konstantní, ale v průběhu výpočtu mohou některé „pružiny“ zanikat anebo naopak nové vznikat podle toho, jak se navzájem příslušné „kuličky“ vzdalují a nebo přibližují. Proto je potřeba, aby simulační program byl ještě navíc vybaven nějakým efektivním mechanismem, který umožňuje rychlé prohledávání okolí jednoho každého atomu [6].

¹ Mgr. Vladimír Pelikán, Ing. Petr Hora, CSc., Ing. Anna Machová, CSc., Akademie věd České republiky, Ústav termomechaniky, Veleslavínova 11, 301 14 PLZEŇ (tel.: 377236415, fax: 377220787, e-mail: Pelikan@cdm.it.cas.cz)

Vícečásticový potenciál

V souladu s [3] definujeme celkovou potenciální energii E soustavy N atomů výrazy:

$$E = \frac{1}{2} \sum_{i \neq j=1}^N V(r_{ij}) - \sum_{i=1}^N \left(\sum_{j \neq i=1}^N \phi(r_{ij}) \right)^{1/2}, \quad (1)$$

kde

$$V(r) = \begin{cases} \frac{1}{r} \sum_{k=1}^4 c_k \cdot \exp(d_k r) & \text{pro } r < x_1, \\ \exp(B_0 + B_1 r + B_2 r^2 + B_3 r^3) & \text{pro } x_1 \leq r < x_2, \\ \sum_{k=1}^6 a_k \cdot H(r_k - r) (r_k - r)^3 & \text{pro } x_2 \leq r, \end{cases}$$

$$\phi(r) = \sum_{k=1}^2 A_k \cdot H(R_k - r) (R_k - r)^3,$$

r_{ij} je vzdálenost mezi i -tým a j -tým atomem, $a_k, r_k, A_k, R_k, B_0, B_1, B_2, B_3, c_k, d_k, x_1, x_2$ jsou tabelované konstanty, které jsou dány ve třech kolekcích v závislosti na typu vazby mezi i -tým a j -tým atomem (Fe-Fe, Fe-Cu, Cu-Cu) a H je Heavisideova funkce.

Odtud dostáváme vztah pro potenciální energii n -tého atomu:

$$E_n = \sum_{i \neq n=1}^N V(r_{in}) - \sqrt{\sum_{i \neq n=1}^N \phi(r_{in})} \quad (2)$$

Pohybové rovnice

Jednotlivé složky *residuální síly*, působící na n -tý atom v důsledku jeho interakce s okolními atomy jsou dány následujícími výrazy [4]:

$$R_{sn} = \sum_{i \neq n=1}^N \left[V'(r_{in}) - \frac{1}{2} \left(\frac{1}{\sqrt{\sum_{j \neq i=1}^N \phi(r_{ji})}} + \frac{1}{\sqrt{\sum_{j \neq n=1}^N \phi(r_{jn})}} \right) \phi'(r_{in}) \right] \frac{s_n - s_i}{r_{in}}, \quad (3)$$

kde

$$r_{in} = \sqrt{(x_i - x_n)^2 + (y_i - y_n)^2 + (z_i - z_n)^2},$$

x_i, y_i, z_i jsou kartézské souřadnice i -tého atomu, x_n, y_n, z_n kartézské souřadnice n -tého atomu a s nabývá po řadě hodnot x, y, z .

Pro každý atom jsou definovány 3 pohybové rovnice:

$$m_n \frac{d^2 s_n}{dt^2} + C \frac{ds_n}{dt} + R_{sn} = F_{sn}, \quad (4)$$

kde t je čas, m_n je hmotnost n -tého atomu, C koeficient útlumu, F_{sn} jsou složky vnější síly, působící na n -tý atom a s opět nabývá po řadě hodnot x, y, z . Koeficient útlumu je různý od nuly pouze v případě, kdy chceme získat statické řešení, např. pro porovnání s modelem kontinua.

Metoda centrálních (středních) diferencí

Zvolíme-li dostatečně malý časový krok Δt a označíme-li symbolem $s_n(i)$ hodnotu veličiny s_n v čase $t_i = t_0 + i\Delta t$, potom podle [5] můžeme psát:

$$\begin{aligned} \left. \frac{ds_n}{dt} \right|_{t=t_i} &\cong \frac{1}{2\Delta t} (s_n(i+1) - s_n(i-1)) \\ \left. \frac{d^2s_n}{dt^2} \right|_{t=t_i} &\cong \frac{1}{\Delta t^2} (s_n(i+1) - 2s_n(i) + s_n(i-1)) \end{aligned} \quad (5)$$

Dosazením (5) do (4) a následnými úpravami dostaneme *rekurentní formuli*:

$$s_n(i+1) = s_n(i-1) + \frac{2}{A_n} (s_n(i) - s_n(i-1)) + \frac{\Delta t^2}{A_n m_n} (F_{sn} - R_{sn}), \quad (6)$$

kde

$$A_n = 1 + \frac{C\Delta t}{2m_n}$$

Stručný popis simulačního programu

Úlohy molekulární dynamiky [1] a [2] jsou postaveny na neustálém ale především mnohónásobném vyčíslování relativně jednoduché rekurentní formule (6), do které vstupují pouze údaje z nejbližšího okolí právě zpracovávaného atomu (převážně se jedná pouze o 14 nejbližších, sousedních atomů). Takovéto úlohy jsou jakoby přímo předurčeny k paralelnímu zpracování a z tohoto důvodu byl na pracovišti ÚT AV ČR učiněn pokus o vytvoření aplikace, která by tyto paralelní výpočty prováděla.

Při počátečních testech paralelního systému *OpenMP* se ukázalo, že tento systém je doposud „nezralý“, jako takový nespolehlivý a tudíž (prozatím ?) nepoužitelný. Z tohoto důvodu byl simulační program vyvinut pod systémem *MPI - Message Passing Interface*, který při úvodních testech vykazoval velmi uspokojivé výsledky.

Simulační program *DynCrack* je napsán v jazyce *Fortran 90* a je určen pro libovolný výpočetní systém, který je vybaven systémem *MPI*.

Zdrojová verze programu *DynCrack* sestává ze dvou částí:

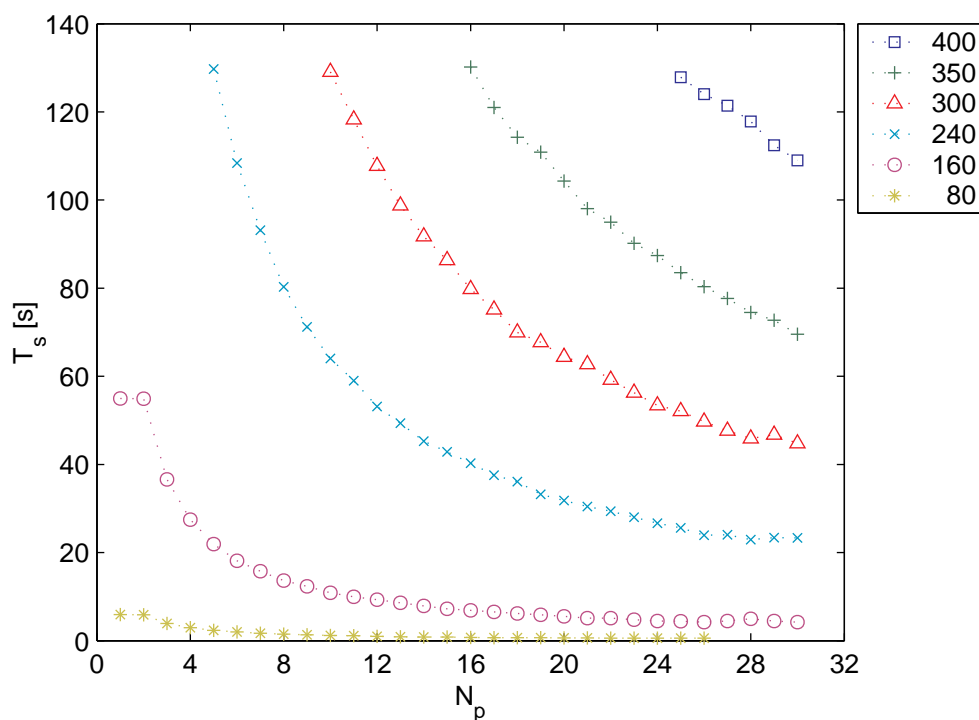
Potencial.FeCu.f výpočet funkcí V a ϕ z [3], jejich derivací a definice některých konstant, popisujících fyzikální strukturu *bcc* Fe-kryystalu.

DynCrack.f vlastní program, který je představován jediným modulem, obsahujícím jak hlavní program, tak kolekci pomocných procedur.

Základní parametry simulované soustavy jsou „natvrdo“ definovány v hlavním programu. Patří sem dimenze zkoumaného krystalu, povolení měděné inkluze a maximální velikost jednoho každého MPI-segmentu. Veškeré ostatní údaje jsou zadávány prostřednictvím vstupního souboru a uživatel je může kdykoliv libovolně měnit. Simulace probíhá po jednotlivých

krocích, ve kterých jsou do souboru na disku průběžně ukládány základní statistické údaje, popisující souhrnný stav simulované soustavy. V předem určených krocích jsou dále ukládány podrobnější údaje o předem vybrané množině atomů, které lze využít například pro následné grafické zpracování (vizualizaci). Konečně na závěr výpočtu je uložen podrobný celkový stav celé soustavy, na který je možno později navázat další simulací. Nad tímto souborem je rovněž možno provádět další speciální vyhodnocení, jakým je například výpočet napětí apod.

Program byl testován na *clusteru MINOS* Západočeské univerzity a na následujícím obrázku je zachyceno, jak při těchto výpočtech závisela střední doba zpracování jednoho simulačního kroku na velikosti zkoumaného vzorku a na počtu použitých procesorů.

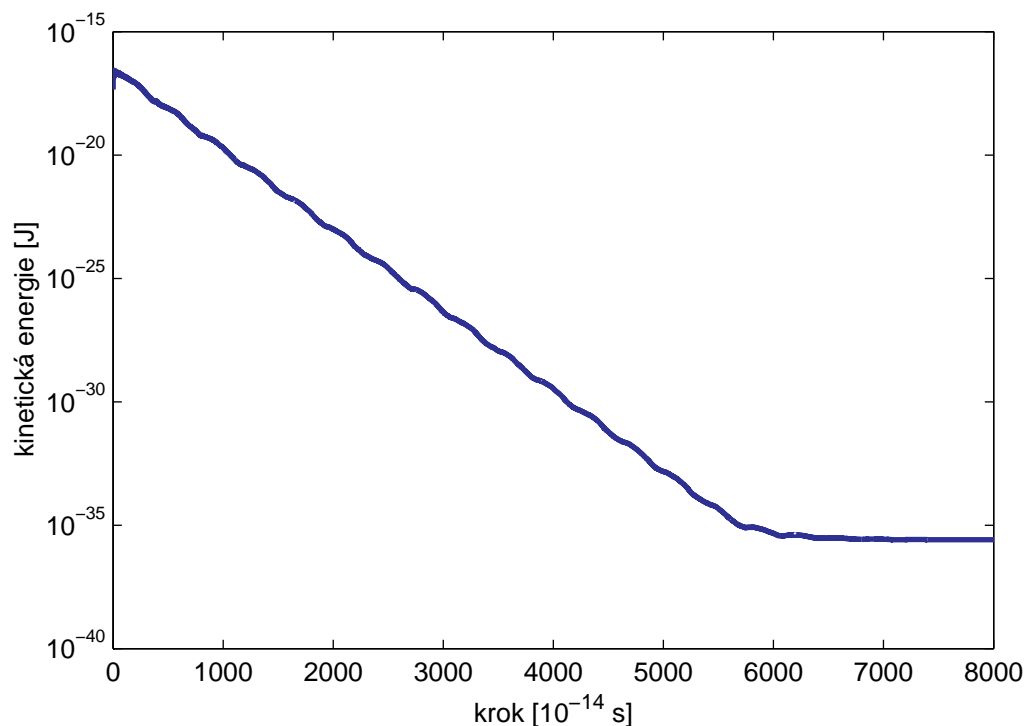


Obr. 1. Závislost doby výpočtu jednoho simulačního kroku (T_s) na velikosti zkoumaného vzorku a na počtu použitých procesorů (N_p).

Zkoumaným vzorkem byly ve všech případech krychle a čísla na pravé straně obrázku udávají počty atomů na hranách těchto krychlí. Pro ilustraci numerické náročnosti těchto výpočtů ještě uvedme, že tyto údaje odpovídají po řadě těmto celkovým množstvím atomů:

Některé typy úloh, zpracovávaných metodou molekulární dynamiky

Prvořadým úkolem každé simulace je vytvoření stabilní, uklidněné soustavy atomů, kdy síly, působící na jednotlivé atomy, jsou v rovnováze a celková kinetická energie soustavy je zanedbatelná (teplota soustavy se blíží absolutní nule). Toho se dosáhne jednoduše tak, že se vygeneruje ideální krystal a výpočet se spustí s vhodně nastaveným útlumem. Proběhne tzv. *povrchová relaxace*. Na následujícím obrázku je zachycen průběh celkové kinetické energie soustavy během tohoto výpočtu. (Zbytková kinetická energie je dána numerickým šumem vzorku.)

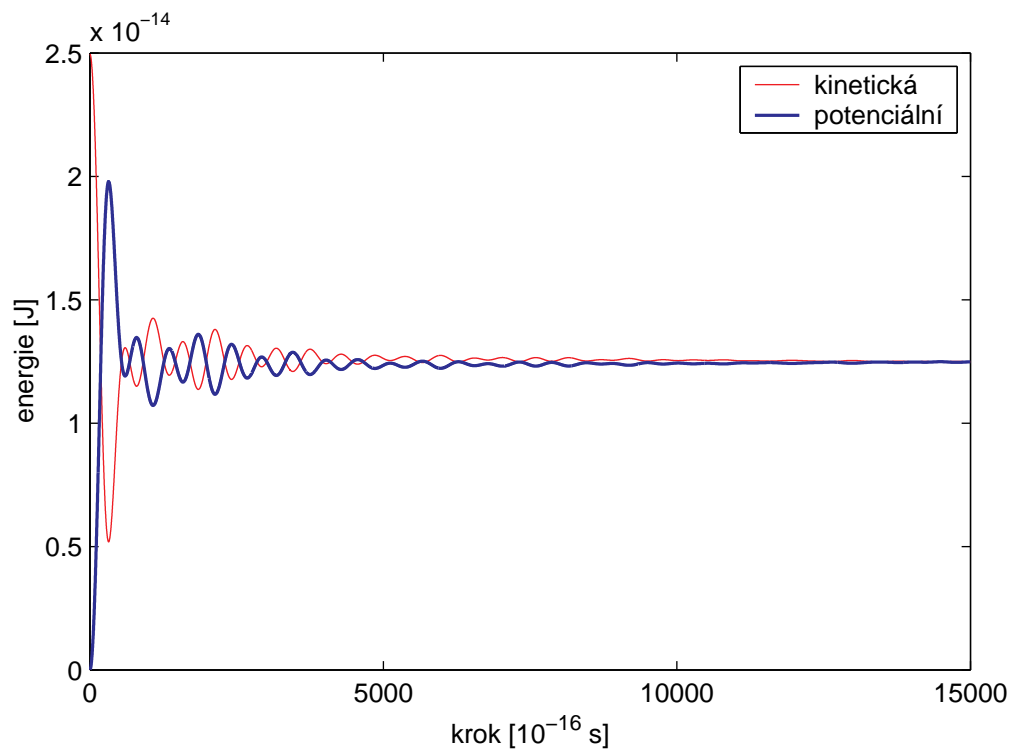


Obr. 2. Kinetická energie celé soustavy v průběhu povrchové relaxace.

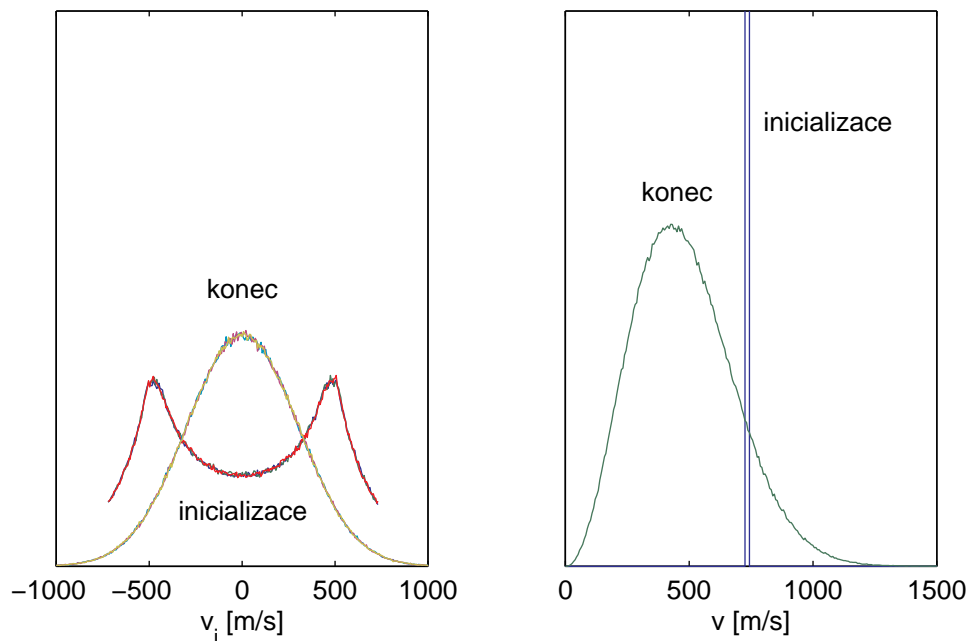
Další úlohou může být například *test teplotní roztažnosti*. Při této úloze je zrelaxovaná soustava skokově zahřátá na teplotu 1200 K tak, že každý atom soustavy je v nultém simulačním kroku „vykopnut“ rychlostí 731.9 m/s náhodným směrem. Tím je systému udělena kinetická energie a teorie praví, že polovina této energie by se měla přeměnit v energii potenciální. Teplota by tedy měla poklesnout na 600 K a nárůst potenciální energie by se měl projevit zvětšením objemu. Na obr. 3 jsou zachyceny průběhy obou energií a na obr. 4 je znázorněno statistické rozdělení dílčích složek vektoru rychlosti resp. absolutní hodnoty vektoru rychlosti jednotlivých atomů na počátku a na konci testu. Na první pohled je patrné, že jakkoliv byla inicializace provedena nešetrně ale především „nefyzikálně“, přesto soustava dospěla do stavu, kdy statistické rozdělení rychlostí je „Gaussovské“ a tedy zcela v souladu s teorií.

Dále byla vypočtena střední vzdálenost všech šesti dvojic úhlopříčně protilehlých hran krychle na počátku (318.9273×10^{-10} m) a na konci testu (321.1244×10^{-10} m). Odtud byl vypočten *koeficient lineární teplotní roztažnosti* $11.4 \times 10^{-6} \text{ K}^{-1}$. Souhlas této veličiny s její experimentálně stanovenou hodnotou je opět velmi uspokojivý.

Obě tyto simulace byly z časových důvodů prováděny na krychli $80 \times 80 \times 80$.



Obr. 3. Přeměna energií při skokovém zahřátí na 1200 K.



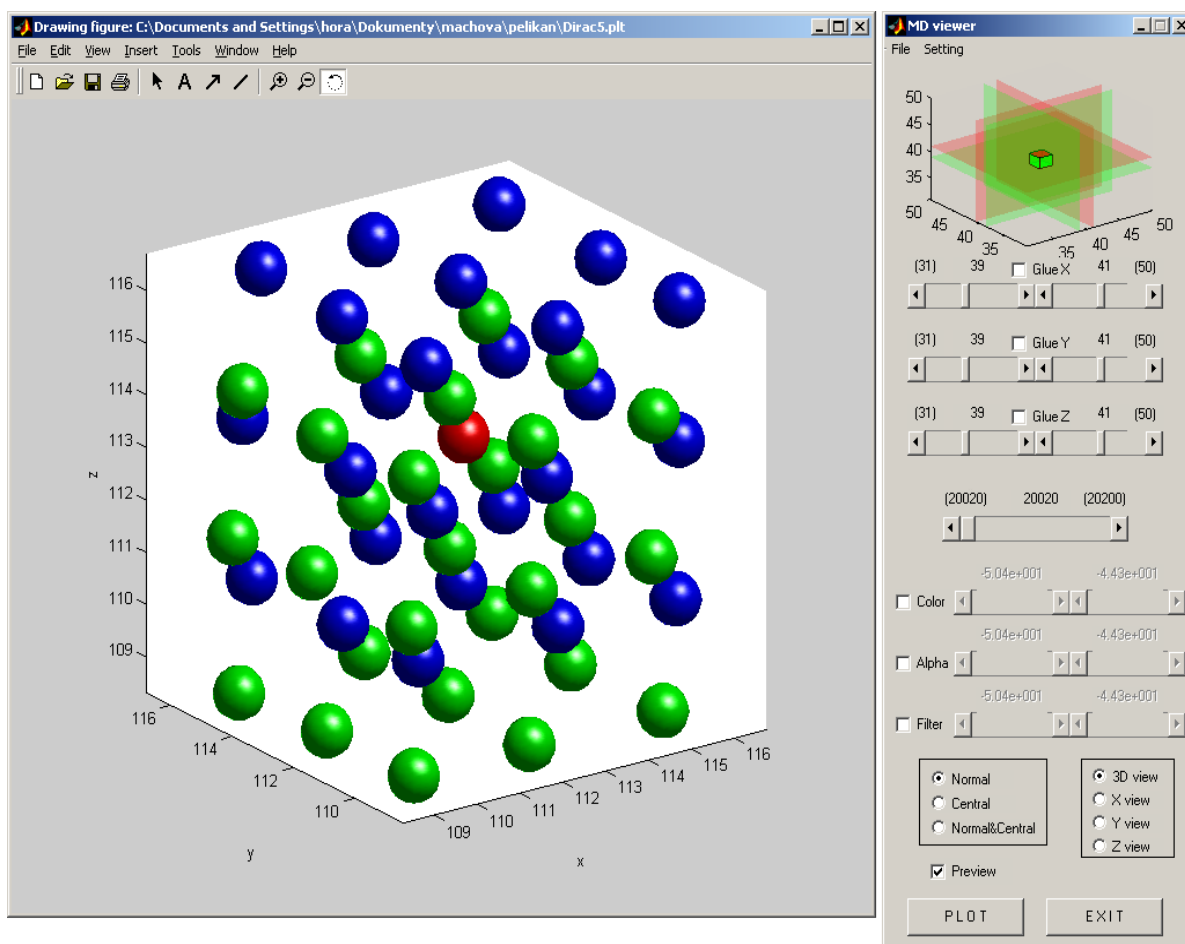
Obr. 4. Statistické rozdělení dílčích složek vektoru rychlosti v_i (vlevo) resp. absolutní hodnoty vektoru rychlosti v (vpravo) jednotlivých atomů na počátku a na konci testu.

Vizualizace výsledků MD-simulace

Jednou z vlastností MD-simulací (zejména ve 3D) je jak potřeba ohromného množství operační paměti při vlastním výpočtu, tak potřeba velkého prostoru na disku pro uložení výsledků. Vyhodnotit takové množství dat není bez účinného vizualizačního prostředku vůbec možné. I když existuje řada komerčních i „free“ programů pro vizualizaci, rozhodli jsme se pro tvorbu vlastního nástroje. Vedly nás k tomu následující okolnosti:

- Existující programy jsou pro nás zbytečně složité, neboť většinou slouží pro vizualizaci komplexních chemických struktur.
- Výstupy našich programů mají specifický formát, který by se musel převádět.
- Cizí program (i když s důkladným manuálem a zdrojovým kódem) je cizí program, není nad to, napsat si něco jednoduchého vlastními silami, pokud na to stačí.

Pro tvorbu vlastního vizualizačního nástroje, viz obr. 5, jsme zvolili programové prostředí MATLAB [7], které už má v sobě řadu pro vizualizaci důležitých vlastností zabudovánu, např. řešení viditelnosti, rendering, průhlednost, textury, atd.



Obr. 5. Grafické okno svizualizací MD-simulace.

Možnosti vizualizačního programu:

- stanovení výřezu
- volba uvažované mřížky (základní, centrální nebo obě)
- volba konkrétního kroku simulace

- stanovení pohledu (obecný 3D, ve směru osy X, Y nebo Z)
- zapnutí nebo vypnutí funkce *preview*
- obarvení atomů podle energie
- nastavení průhlednosti atomů podle energie
- odfiltrování atomů podle energie
- nastavení mapy barev a počtu barev v této mapě
- nastavení mapy průhledností a počtu průhledností v této mapě, atd.

Program byl psán a odlaďován v MATLABu 6.5 na IBM ThinkPad R30 sPIII 1GHz, 512MB RAM pod systémem Windows 2000. Na této sestavě lze v „reálném“ čase, tj. odezva do 5 sekund, vykreslovat cca 1000 atomů. Při větším množství atomů je lepší funkci *preview* vypnout. Na sestavě s PIV 2.4 GHz, 512 MB RAM šlo v „reálném“ čase vykreslovat cca 3000 atomů.

Poděkování

Příspěvek vznikl na základě podpory grantu GA AV ČR A2076201 „Simulace křehce-tvárného chování mikrotrhlin metodou molekulární dynamiky v2D a 3D úlohách“ a dále AV0Z2076919.

Literatura

- [1] Pelikán, V., Machová, A.: *Simulace radiačního poškození v železe*. Výzkumná zpráva Z1309-01, ÚT AV ČR, Praha, prosinec 2001.
- [2] Machová, A.: *3D simulace metodou molekulární dynamiky v α -Fe a systému Fe-Cu*. Výzkumná zpráva Z1280/00, ÚT AV ČR, Praha, listopad 2000.
- [3] Ackland, G. J., Bacon, D. J., Calder, A. F., Harry, T.: *Computer simulations of point defect properties in dilute Fe-Cu alloy using a many-body interatomic potential*. Phil. Mag. A 75 (1997), pp. 713-732.
- [4] Finnis, M. W., Sinclair, J. E.: *A simple empirical N-body potential for transition metals*. Phil. Mag. A 50 (1984), pp. 45-55.
- [5] Rektorys, K. a spolupracovníci: *Přehled užití matematiky II*. Praha, SNTL, 1988, pp. 1013-1017.
- [6] Frenkel, D., Smit, B.: *Understanding Molecular Simulations*. Academic Press, New York, 1996.
- [7] <http://www.mathworks.com/>

PARALLEL PROGRAMMING IN MOLECULAR DYNAMICS SIMULATIONS

Molecular dynamics simulation is a well-established technique for modelling complex many-particle systems in diverse areas of physics and chemistry. The computational requirements of simulations of large systems, especially when taking into account long-range interactions, are enormous. They make the use of high-performance parallel computers indispensable and have led to the development of a broad range of advanced algorithms for these machines. There is, however, still a long way to go in order to simulate realistic systems over time scales of typical physical, chemical, and biological processes.

We have developed a classical molecular dynamics code in MPI based on Newtonian dynamics.