18 Fekete problem

18.1 General information

The problem is an index 2 DAE from mechanics. The dimension is $8N$, where N is a user supplied integer. The numerical tests shown here correspond to $N = 20$. The problem is of interest for the computation of the elliptic Fekete points [Par95]. The parallel-IVP-algorithm group of CWI contributed this problem to the test set, in collaboration with W. J. H. Stortelder. The software part of the problem is in the file fekete.f available at [MM08].

18.2 Mathematical description of the problem

The problem is of the form

$$M\frac{\mathrm{d}y}{\mathrm{d}t} = f(y), \quad y(0) = y_0, \quad y'(0) = y'_0, \tag{II.18.1}$$

with

$$y, f \in \mathbb{R}^{8N}, \quad 0 \le t \le t_{\mathrm{end}}.$$

Here, $t_{\mathrm{end}} = 1000$, $N = 20$ and M is the (constant) mass matrix given by

$$M = \begin{pmatrix} I_{6N} & 0 \\ 0 & 0 \end{pmatrix},$$

where I_{6N} is the identity matrix of dimension $6N$. For the definition of the function f, we refer to §18.3.

The components $y_{0,i}$ of of the initial vector y_0 are defined by

$$\begin{pmatrix} y_{0,3(j-1)+1} \\ y_{0,3(j-1)+2} \\ y_{0,3(j-1)+3} \end{pmatrix} = \begin{pmatrix} \cos(\omega_j)\cos(\beta_j) \\ \sin(\omega_j)\cos(\beta_j) \\ \sin(\beta_j) \end{pmatrix} \quad \text{for} \quad j = 1, \ldots, N,$$

where

$$\begin{aligned}
\beta_j &= \tfrac{3}{8}\pi \quad \text{and} \quad \omega_j = \tfrac{2j}{3}\pi + \tfrac{1}{13}\pi \quad \text{for} \quad j = 1, \ldots, 3, \\
\beta_j &= \tfrac{1}{8}\pi \quad \text{and} \quad \omega_j = \tfrac{2(j-3)}{7}\pi + \tfrac{1}{29}\pi \quad \text{for} \quad j = 4, \ldots, 10, \\
\beta_j &= -\tfrac{2}{15}\pi \quad \text{and} \quad \omega_j = \tfrac{2(j-10)}{6}\pi + \tfrac{1}{7}\pi \quad \text{for} \quad j = 11, \ldots, 16, \\
\beta_j &= -\tfrac{3}{10}\pi \quad \text{and} \quad \omega_j = \tfrac{2(j-17)}{4}\pi + \tfrac{1}{17}\pi \quad \text{for} \quad j = 17, \ldots, 20,
\end{aligned}$$

and

$$\begin{aligned}
y_{0,i} &= 0 && \text{for} \quad i = 3N+1, \ldots, 6N, \\
y_{0,6N+j} &= \tfrac{1}{2}\langle p_j(0), \widehat{f}_j \rangle && \text{for} \quad j = 1, \ldots, N, \\
y_{0,i} &= 0 && \text{for} \quad i = 7N+1, \ldots, 8N,
\end{aligned}$$

where

$$p_j = \begin{pmatrix} y_{3(j-1)+1} \\ y_{3(j-1)+2} \\ y_{3(j-1)+3} \end{pmatrix}, \quad \widehat{f}_j = \begin{pmatrix} f_{3N+3(j-1)+1}((p(0), 0, \ldots, 0)^{\mathrm{T}}) \\ f_{3N+3(j-1)+2}((p(0), 0, \ldots, 0)^{\mathrm{T}}) \\ f_{3N+3(j-1)+3}((p(0), 0, \ldots, 0)^{\mathrm{T}}) \end{pmatrix}, \tag{II.18.2}$$

and $p = (y_1, y_2, \ldots, y_{3N})^{\mathrm{T}}$. The initial derivative vector reads $y'_0 = f(y_0)$. These definitions of y_0 and y'_0 yield consistent initial values. The first $6N$ components are of index 1, the last $2N$ of index 2.

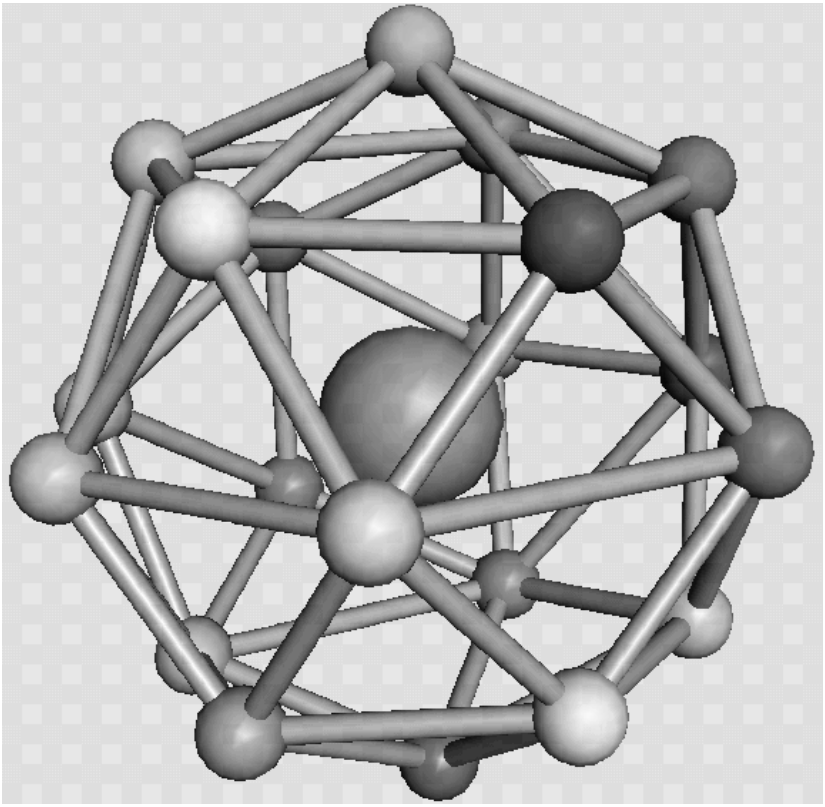

FIGURE II.18.1: *Final configuration for $N = 20$. The large ball is centered at the origin and only added to facilitate the 3-D perception. (Taken from [PSS97] by courtesy of R. van Liere.)*

18.3 Origin of the problem

This problem is of interest for the computation of the elliptic Fekete points. Let us define the unit sphere in $\mathbb{R}^3$ by $\mathcal{S}^2$ and for any configuration $x := (x_1, x_2, \ldots, x_N)^{\mathrm{T}}$ of points $x_i \in \mathcal{S}^2$, the function

$$V(x) := \prod_{i<j} \|x_i - x_j\|_2. \tag{II.18.3}$$

We denote the value of x for which V reaches its global maximum by $\widehat{x} = (\widehat{x}_1, \ldots, \widehat{x}_N)$. The points $\widehat{x}_1, \widehat{x}_2, \ldots, \widehat{x}_N$ are called the elliptic Fekete points of order N. For example, for $N = 4$, the points of the optimal solution form a tetrahedron. But, in case of 8 points, intuition fails; the elliptic Fekete points do not form a cube in this case. A cube where, for example, the upper plane is rotated over 45° with respect to the bottom plane, gives already a larger value of V. It turns out (see e.g. [Par95]) that $\widehat{x}$ is difficult to compute as solution of a global optimization problem. For reasons that will become clear later, we differentiate $\log(V)$ with respect to x_k and apply the method of Lagrange multipliers, to see that $\widehat{x}$ fulfills

$$\nabla_k \log(V(x)) \mid_{x=\widehat{x}} = \sum_{j \neq k} \frac{\widehat{x}_k - \widehat{x}_j}{\|\widehat{x}_k - \widehat{x}_j\|_2^2} = \zeta_k \widehat{x}_k, \tag{II.18.4}$$

where the ζ_k are Lagrange multipliers.

We now discuss the Fekete points from another point of view. Consider on $\mathcal{S}^2$ a number of N particles, on which two forces are invoked: a repulsive force, by which the particles will start to move away from each other, and an adhesion force, by which the particles will reach a stationary state after a certain period of time.

We denote the position in Cartesian coordinates of particle i at time t by $p_i(t)$ and the configuration of N points at time t by $p(t) = (p_1(t), \ldots, p_N(t))^{\mathrm{T}}$. The stationary configuration is assumed to be obtained at $t = t_{\mathrm{stat}}$ and will be denoted by $\widehat{p} := (\widehat{p}_1, \widehat{p}_2, \ldots, \widehat{p}_N)$, where $\widehat{p}_i := p_i(t_{\mathrm{stat}})$. The repulsive force on particle i caused by particle j is defined by

$$F_{ij} = \frac{p_i - p_j}{\|p_i - p_j\|_2^{\gamma}}.$$

Note that the choice $\gamma = 3$ can be interpreted as an electrical force working on particles with unit charge. The adhesion force working on particle i is denoted by A_i and given by

$$A_i = -\alpha q_i.$$

Here, q is the velocity vector and α is valued 0.5.

We can compute the configuration of the particles as function of time, given that the particles cannot leave the unit sphere, as solution of the DAE system

$$\begin{aligned} p' &= q, && \text{(II.18.5)}\\ q' &= g(p,q) + G^{\mathrm{T}}(p)\lambda, && \text{(II.18.6)}\\ 0 &= \phi(p), && \text{(II.18.7)} \end{aligned}$$

where $G = \partial\phi/\partial p$ and $\lambda \in \mathbb{R}^N$. The function $\phi : \mathbb{R}^{3N} \to \mathbb{R}^N$ represents the constraint, which states that the particles remain on the unit sphere:

$$\phi_i(p) = p_{i,1}^2 + p_{i,2}^2 + p_{i,3}^2 - 1.$$

The function $g : \mathbb{R}^{6N} \to \mathbb{R}^{3N}$ is given by $g = (g_i)$, $i = 1, \ldots, N$, where

$$g_i(p,q) = \sum_{j \neq i} F_{ij}(p) + A_i(q).$$

The term $G^{\mathrm{T}}(p)\lambda$ in (II.18.6) represents the normal force which keeps the particle on $\mathcal{S}^2$.

Since we know that the speed of the final configuration at $t = t_{\mathrm{stat}}$ is 0, we can substitute $q = 0$ and $p = \widehat{p}$ in formula (II.18.6), thus arriving at

$$0 = \sum_{j \neq i} F_{ij}(\widehat{p}) + G^{\mathrm{T}}(\widehat{p})\lambda \ ,$$

which is equal to

$$\sum_{i \neq j} \frac{\widehat{p}_i - \widehat{p}_j}{\|\widehat{p}_i - \widehat{p}_j\|_2^{\gamma}} = -2\lambda_i \widehat{p}_i \ . \qquad \text{(II.18.8)}$$

Comparing (II.18.4) and (II.18.8) tells us that computing $\widehat{p}$ for $\gamma = 2$ gives the local optima of the function V in (II.18.3). In [PSS97], it is showed that computing $\widehat{p}$ by solving the system (II.18.5)–(II.18.7) and then substituting $x = \widehat{p}$ in (II.18.3), results in values of V that are very competitive with those obtained by global optimization packages. For more details on elliptic Fekete points, we refer to [Par95] and [SS93].

The DAE system mentioned before is of index 3. To arrive at a more stable formulation of the problem, we stabilize the constraint (see [BCP89, p. 153]) by replacing (II.18.5) by

$$p' = q + G^{\mathrm{T}}(p)\mu, \qquad \text{(II.18.9)}$$

where $\mu \in \mathbb{R}^N$, and appending the differentiated constraint

$$0 = G(p)q. \qquad \text{(II.18.10)}$$

TABLE II.18.1: *Reference solution at the end of the integration interval.*

$y(1)$	−0.4070263380333202	$y(7)$	0.7100577833343567
$y(2)$	0.3463758772791802	$y(8)$	0.1212948055586120
$y(3)$	0.8451942450030429	$y(9)$	0.6936177005172217
$y(4)$	0.0775293475252155	$y(10)$	0.2348267744557627
$y(5)$	−0.2628662719972299	$y(11)$	0.7449277976923311
$y(6)$	0.9617122871829146	$y(12)$	0.6244509285956391

The system (II.18.9), (II.18.6), (II.18.7), (II.18.10) is now of index 2; the variables p and q are of index 1, the variables λ and μ of index 2. We cast the system in the form (II.18.1) by setting $y = (p, q, \lambda, \mu)^{\mathrm{T}}$ and $f(y) = f(p, q, \lambda, \mu) = (q + G^{\mathrm{T}}\mu, g + G^{\mathrm{T}}\lambda, \phi, Gq)^{\mathrm{T}}$, where p_i is in Cartesian coordinates.

The choice for the initial configuration as defined in §18.2 is a rough attempt to spread out the points over the sphere. To arrive at a consistent set of initial values we choose $q(0) = 0$, yielding $\mu(0) = 0$ and $\phi_i'(0) = \langle 2p_i(0), q_i(0)\rangle = 0$. Consequently,

$$\begin{aligned} \phi_i''(0) &= \langle 2p_i(0), q_i'(0)\rangle \\ &= \langle 2p_i(0), g_i(p(0), q(0)) + 2\lambda_i(0)p_i(0)\rangle. \end{aligned}$$

Requiring $\phi_i''(0) = 0$ gives

$$\lambda_i(0) = -\frac{\langle p_i(0), g_i(p(0), q(0))\rangle}{2\langle p_i(0), p_i(0)\rangle} = -\frac{1}{2}\langle p_i(0), g_i(p(0), q(0))\rangle.$$

The initial derivative vector y_0' can be chosen equal to $f(y_0)$. For $N \le 20$, $t_{\mathrm{stat}} \le 1000$, therefore we chose $t_{\mathrm{end}} = 1000$.

In Figure II.18.1 the final configuration for 20 points is plotted.

18.4 Numerical solution of the problem

All the tests concern the case with $N = 20$. Tables II.18.1–II.18.2 and Figures II.18.2–II.18.6 present the reference solution at the end of the integration interval (first 12 components), the run characteristics, the behavior of the first 6 solution components over the interval $[0, 20]$ and the work-precision diagrams, respectively. In computing the scd values, only the first sixty components were considered, since they refer to the position of the particles. The reference solution was computed using RADAU5, rtol $= 10^{-12}$, atol $= 10^{-12}$, and h0 $= 10^{-12}$. For the work-precision diagrams, we used: rtol $= 10^{-(2+m/16)}$, $m = 0, 1, \ldots, 64$; atol = rtol; h0 = rtol for BIMD, GAMD, MEBDFDAE, MEBDFI, RADAU and RADAU5.

References

[BCP89] K.E. Brenan, S.L. Campbell, and L.R. Petzold. *Numerical Solution of Initial–Value Problems in Differential–Algebraic Equations*. North–Holland, New York–Amsterdam–London, 1989.

[MM08] F. Mazzia and C. Magherini. *Test Set for Initial Value Problem Solvers, release 2.4*. Department of Mathematics, University of Bari and INdAM, Research Unit of Bari, February 2008. Available at http://www.dm.uniba.it/~testset.

[Par95] P.M. Pardalos. An open global optimization problem on the unit sphere. *Journal of Global Optimization*, 6:213, 1995.

TABLE II.18.2: *Run characteristics.*

solver	rtol	atol	h0	mescd	scd	steps	accept	#f	#Jac	#LU	CPU
BIMD	10^{-2}	10^{-2}	10^{-2}	4.12	2.63	30	29	415	29	30	0.2450
	10^{-3}	10^{-3}	10^{-3}	5.36	4.19	43	43	668	43	43	0.3445
	10^{-4}	10^{-4}	10^{-4}	6.69	5.33	65	65	1094	65	65	0.5124
GAMD	10^{-2}	10^{-2}	10^{-2}	4.16	2.99	26	24	526	24	26	0.2147
	10^{-3}	10^{-3}	10^{-3}	4.79	3.78	26	26	947	26	26	0.3006
	10^{-4}	10^{-4}	10^{-4}	5.76	4.45	38	38	1319	38	38	0.4119
MEBDFI	10^{-2}	10^{-2}	10^{-2}	3.56	2.10	60	57	192	15	15	0.1064
	10^{-3}	10^{-3}	10^{-3}	4.58	3.23	129	128	428	18	18	0.1513
	10^{-4}	10^{-4}	10^{-4}	5.81	4.81	218	216	707	23	23	0.2176
PSIDE-1	10^{-2}	10^{-2}		3.66	2.20	73	53	693	16	288	1.3137
	10^{-3}	10^{-3}		4.40	3.19	88	59	779	11	344	1.4357
	10^{-4}	10^{-4}		5.32	4.12	114	75	967	9	448	1.7363
RADAU	10^{-2}	10^{-2}	10^{-2}	3.43	1.97	33	30	274	27	32	0.5065
	10^{-3}	10^{-3}	10^{-3}	4.11	2.65	43	41	315	38	43	0.5993
	10^{-4}	10^{-4}	10^{-4}	5.36	4.29	61	58	442	54	61	0.7662

[PSS97] J.D. Pintér, W.J.H. Stortelder, and J.J.B. de Swart. Computation of elliptic Fekete point sets. Report MAS-R9705, CWI, Amsterdam, 1997. To appear in CWI Quarterly.

[SS93] M. Shub and S. Smale. Complexity of Bezout's theorem III. Condition number and packing. *Journal of Complexity*, 9:4–14, 1993.

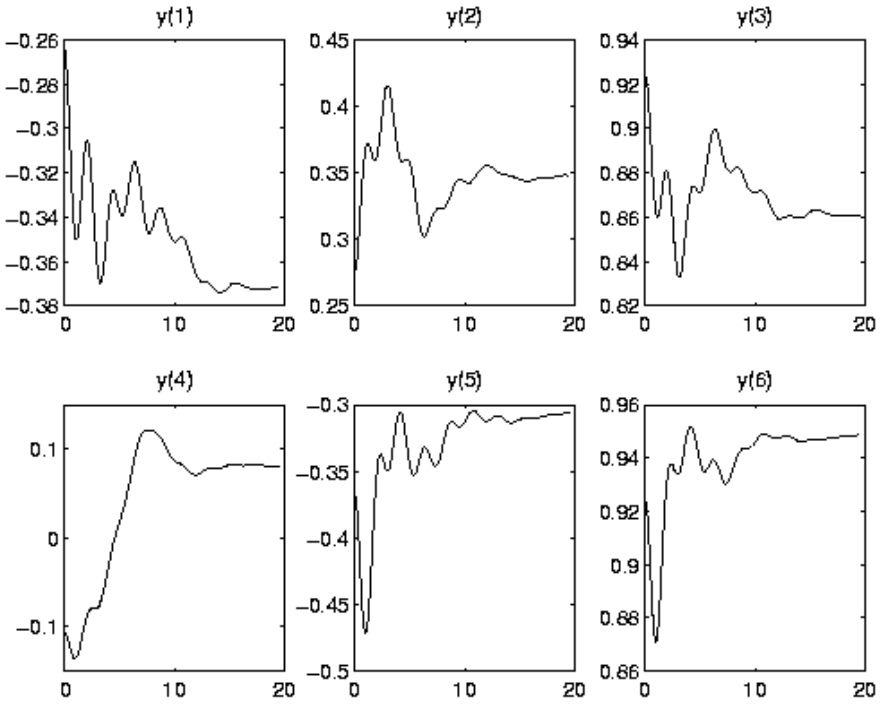

FIGURE II.18.2: *Behavior of the solution over the integration interval.*

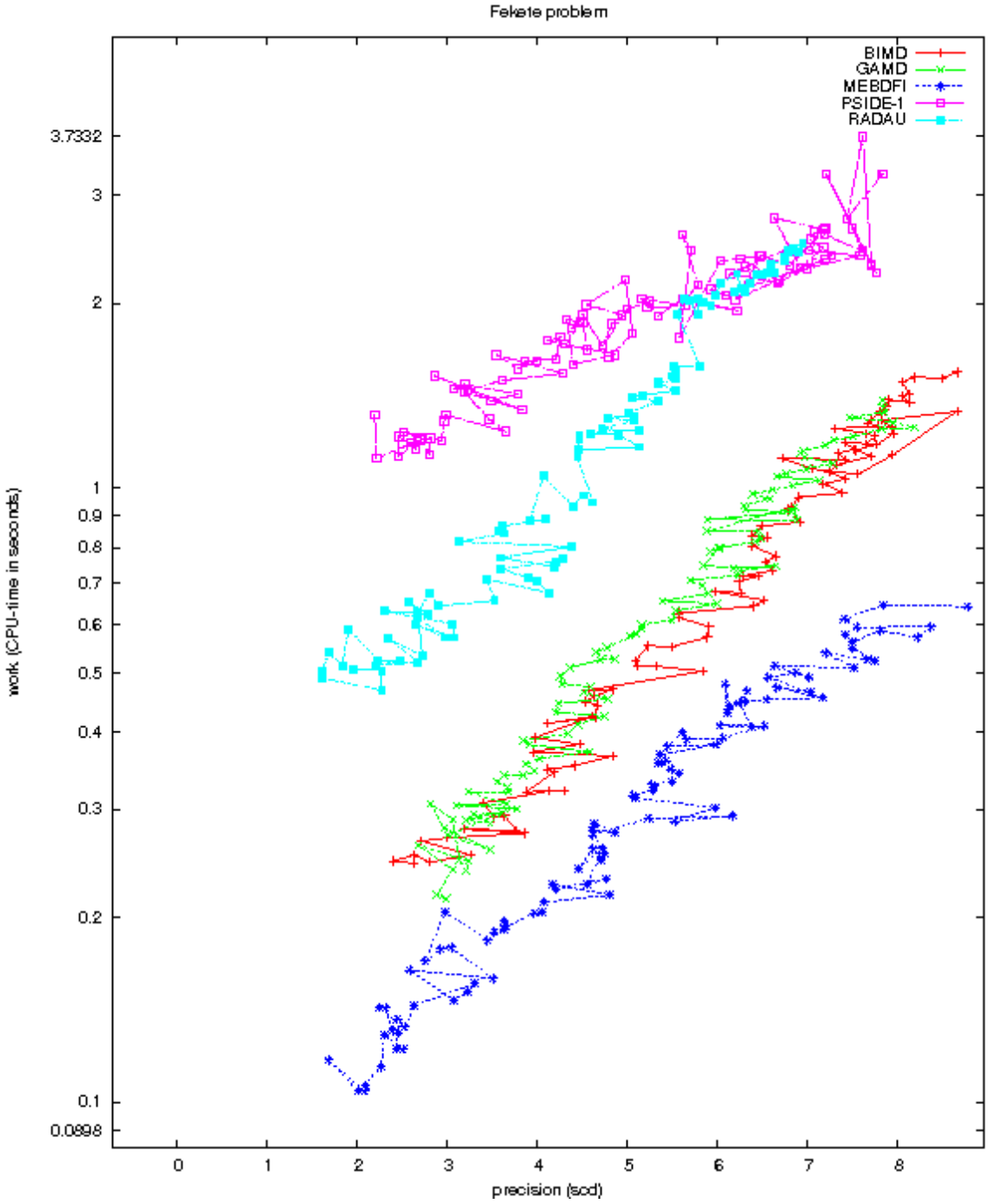

FIGURE II.18.3: *Work-precision diagram (scd versus CPU-time).*

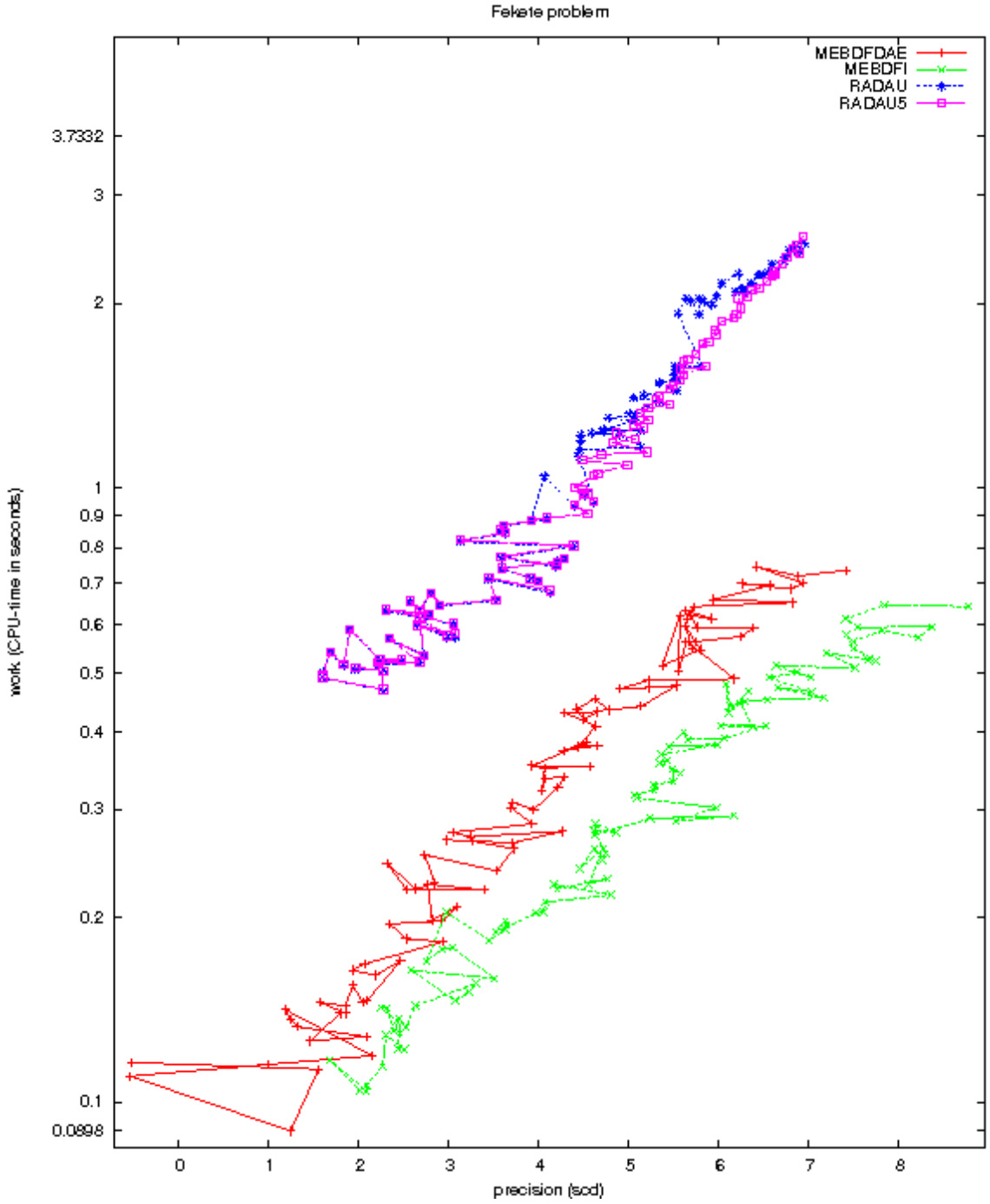

FIGURE II.18.4: *Work-precision diagram (scd versus CPU-time).*

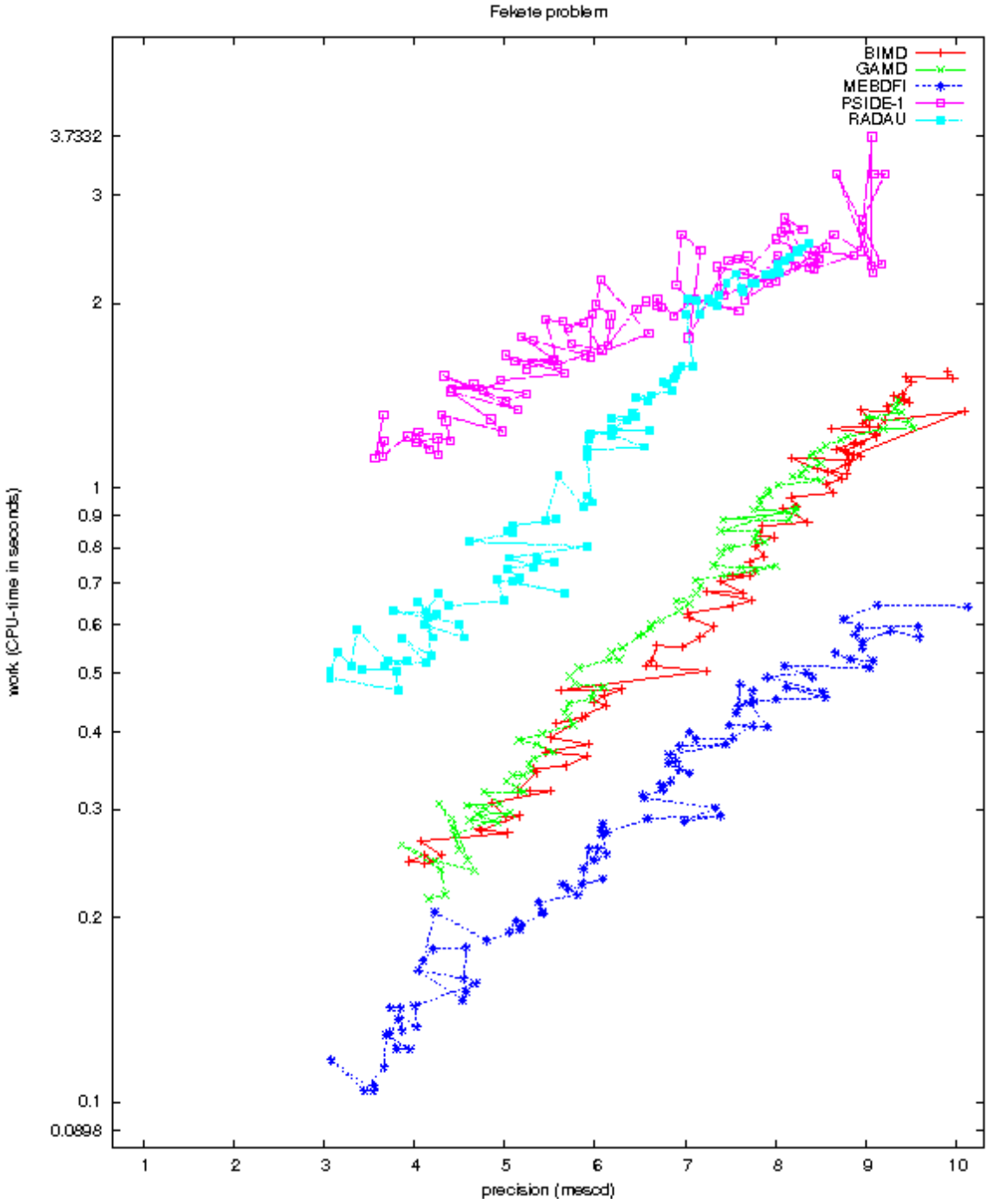

FIGURE II.18.5: *Work-precision diagram (mescd versus CPU-time).*

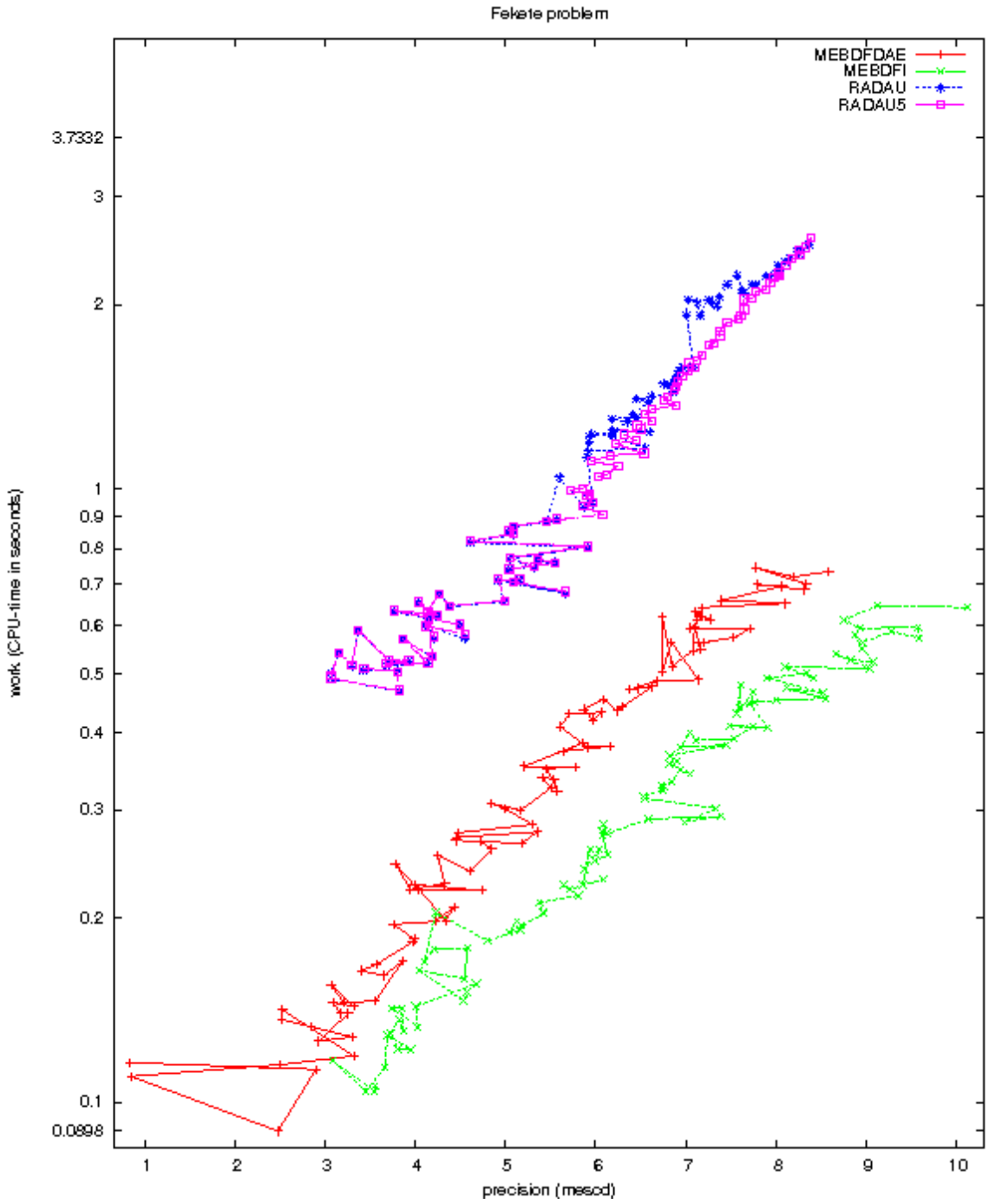

FIGURE II.18.6: *Work-precision diagram (mescd versus CPU-time).*