How Accurate Does Newton Have to Be?

Carl Christian Kjelgaard Mikkelsen[1(✉)], Lorién López-Villellas[2], and Pablo García-Risueño[3]

[1] Department of Computing Science, Umeå University, 90187 Umeå, Sweden
spock@cs.umu.se
[2] Barcelona Supercomputing Center, Barcelona, Spain
lorien.lopez@bsc.es
[3] Zaragoza, Spain
risueno@unizar.es

Abstract. We analyze the convergence of quasi-Newton methods in exact and finite precision arithmetic. In particular, we derive an upper bound for the stagnation level and we show that any sufficiently exact quasi-Newton method will converge quadratically until stagnation. In the absence of sufficient accuracy, we are likely to retain rapid linear convergence. We confirm our analysis by computing square roots and solving bond constraint equations in the context of molecular dynamics. We briefly discuss implications for parallel solvers.

Keywords: Systems of nonlinear equations · Quasi-Newton methods · approximation error · rounding error · convergence · stagnation

1 Introduction

Let $\Omega \subseteq \mathbb{R}^n$ be open, let $F \in C^1(\Omega, \mathbb{R}^n)$ and consider the problem of solving

$$F(x) = 0.$$

If the Jacobian F' of F is nonsingular, then Newton's method is given by

$$x_{k+1} = x_k - s_k, \quad F'(x_k)s_k = F(x_k). \tag{1}$$

A quasi-Newton method is any iteration of the form

$$y_{k+1} = y_k - t_k, \quad F'(y_k)t_k \approx F(y_k). \tag{2}$$

In exact arithmetic, we expect local quadractic convergence from Newton's method [7]. Quasi-Newton methods normally converge locally and at least linearly and some methods, such as the secant method, have superlinear convergence [5,8]. In finite precision arithmetic, we cannot expect convergence in the strict mathematical sense and we must settle for stagnation near a zero [11]. In

P. García-Risueño—Independent scholar.

R. Wyrzykowski et al. (Eds.): PPAM 2022, LNCS 13826, pp. 3–15, 2023.
https://doi.org/10.1007/978-3-031-30442-2_1

this paper we analyze the convergence of quasi-Newton methods in exact and finite precision arithmetic. In particular, we derive an upper bound for the stagnation level and we show that any sufficiently exact quasi-Newton method will converge quadratically until stagnation. We confirm our analysis by computing square roots and solving bond constraint equations in the context of molecular dynamics.

2 Auxiliary Results

The line segment $l(x, y)$ between x and y is defined as follows:

$$l(x, y) = \{tx + (1 - t)y \ : \ t \in [0, 1]\}.$$

The following lemma is a standard result that bounds the difference between $F(x)$ and $F(y)$ if the line segment $l(x, y)$ is contained in the domain of F.

Lemma 1. *Let $\Omega \subseteq \mathbb{R}^n$ be open and let $F \in C^1(\Omega, \mathbb{R}^n)$. If $l(x, y) \subset \Omega$, then*

$$F(x) - F(y) = \int_0^1 F'(tx + (1 - t)y)(x - y)dt$$

and

$$\|F(x) - F(y)\| \leq M\|x - y\|.$$

where

$$M = \sup\{\|F'(tx + (1 - t)y)\| \ : \ t \in [0, 1]\}.$$

It is convenient to phrase Newton's method as the functional iteration:

$$x_{k+1} = g(x_k), \quad g(x) = x - F'(x)^{-1}F(x).$$

and to express the analysis of quasi-Newton methods in terms of the function g. The next lemma can be used to establish local quadratic convergence of Newton's method.

Lemma 2. *Let $\Omega \subseteq \mathbb{R}^n$ be open and let $F \in C^1(\Omega, \mathbb{R}^n)$. Let z denote a zero of F and let $x \in \Omega$. If $F'(x)$ is nonsingular and if $l(x, z) \subset \Omega$, then*

$$g(x) - z = C(x)(x - z)$$

where

$$C(x) = F'(x)^{-1} \left(\int_0^1 \left[F'(x) - F'(tx + (1 - t)z) \right] dt \right)$$

Moreover, if F' is Lipschitz continuous with Lipschitz constant $L > 0$, then

$$\|g(x) - z\| \leq \frac{1}{2}\|F'(x)^{-1}\|L\|x - z\|^2.$$

The following lemma allows us to write any approximation as a very simple function of the target vector.

Lemma 3. *Let $x \in \mathbb{R}^n$ be nonzero, let $y \in \mathbb{R}^n$ be an approximation of x and let $E \in \mathbb{R}^{n\times n}$ be given by*

$$E = \frac{1}{x^T x}(y - x)x^T.$$

Then

$$y = (I + E)x, \quad \|E\| = O\left(\frac{\|x - y\|}{\|x\|}\right), \quad y \to x, \quad y \neq x.$$

In the special case of the 2-norm we have

$$\|E\|_2 = \frac{\|x - y\|_2}{\|x\|_2}.$$

Proof. It is straightforward to verify that

$$(I + E)x = x + \frac{1}{x^T x}(y - x)x^T x = x + (y - x) = y.$$

Moreover, if z is any vector, then

$$\|Ez\| \leq \frac{1}{\|x\|_2^2}\|y - x\|\|x^T z\| = \left(\frac{\|x^T\|\|x\|}{\|x\|_2^2}\right)\left(\frac{\|x - y\|}{\|x\|}\right)\|z\|.$$

In the case of the 2-norm, we have

$$\|Ez\|_2 \leq \frac{\|x - y\|_2}{\|x\|_2}\|z\|_2$$

for all $z \neq 0$ and equality holds for $z = x$. This completes the proof.

3 Main Results

In the presence of rounding errors, *any* quasi-Newton method can written as

$$x_{k+1} = (I + D_k)\Big(x_k - (I + E_k)F'(x_k)^{-1}F(x_k)\Big). \tag{3}$$

Here $D_k \in \mathbb{R}^{n\times n}$ is a diagonal matrix which represents the rounding error in the subtraction and $E_k \in \mathbb{R}^{n\times n}$ measures the difference between the computed correction and the correction used by Newton's method. We simply treat the update t_k needed for the quasi-Newton method (2) as an approximation of the update $s_k = F'(x_k)^{-1}F(x_k)$ needed for Newton's method (1) and define E_k using Lemma 3. It is practical to restate iteration (3) in terms of the function g, i.e.,

$$x_{k+1} = (I + D_k)\Big(g(x_k) - E_k F'(x_k)^{-1}F(x_k)\Big). \tag{4}$$

We shall now analyze the behavior of iteration (4). For the sake of simplicity, we will assume that there exist nonnegative numbers K, L, and M such that

$$\forall x \, : \, \|F'(x)^{-1}\| \leq K, \quad \|F'(x) - F'(y)\| \leq L\|x - y\|, \quad \|F'(x)\| \leq M.$$

In reality, we only require that these inequalities are satisfied in a neighborhood of a zero. We have the following generalization of Lemma 2.

Theorem 1. *The functional iteration given by Eq.* (4) *satisfies*

$$x_{k+1} - z = g(x_k) - z - E_k F'(x_k)^{-1} F(x_k) + D_k\left[g(x_k) - E_k F'(x_k)^{-1} F(x_k)\right] \quad (5)$$

and

$$\|x_{k+1} - z\| \leq \frac{1}{2} LK\|x_k - z\|^2 + \|E_k\|KM\|x_k - z\| + \|D_k\|\left(\|z\| + \frac{1}{2}LK\|x_k - z\|^2 + \|E_k\|KM\|x_k - z\|\right). \quad (6)$$

Proof. It is straightforward to verify that Eq. (5) is correct. Inequality (6) follows from Eq. (5) using the triangle inequality, Lemma 1, and Lemma 2. The second occurrence of the term $\|g(x_k)\|$ can be bounded using the inequality

$$\|g(x_k)\| \leq \|z\| + \|g(x_k) - z\|.$$

This completes the proof.

It is practical to focus on the case of $z \neq 0$ and restate inequality (6) as

$$r_{k+1} \leq \frac{1}{2} LK(1 + \|D_k\|)\|z\| r_k^2 + \|E_k\|KM(1 + \|D_k\|)r_k + \|D_k\| \quad (7)$$

where r_k is the normwise relative forward error given by

$$r_k = \|z - x_k\|/\|z\|.$$

3.1 Stagnation

We assume that the sequences $\{D_k\}$ and $\{E_k\}$ are bounded. Let D and E be nonnegative numbers that satisfy

$$\|D_k\| \leq D, \quad \|E_k\| \leq E. \quad (8)$$

In this case, inequality (7) implies

$$r_{k+1} \leq \frac{1}{2} LK(1 + D)\|z\| r_k^2 + EMK(1 + D)r_k + D.$$

It is certain that the error will be reduced, i.e., $r_{k+1} < r_k$ when

$$\begin{aligned} D &< r_k - \left(\frac{1}{2}LK(1+D)\|z\|r_k^2 + EMK(1+D)r_k^2\right) \\ &= (1 - EMK(1+D))\, r_k - \frac{1}{2}LK(1+D)\|z\|r_k^2. \end{aligned}$$

This condition is equivalent to the following inequality:

$$D - [1 - EMK(1+D)]\, r_k + \frac{1}{2}LK(1+D)\|z\|r_k^2 < 0.$$

This is an inequality of the second degree. The roots are

$$\lambda_{\pm} = \frac{[1 - EMK(1+D)] \pm \sqrt{[1 - EMK(1+D)]^2 - 2LK(1+D)D\|z\|}}{LK(1+D)\|z\|}.$$

If D and E are sufficiently small then the roots are positive real numbers and the error will certainly be reduced provided

$$\lambda_- < r_k < \lambda_+.$$

It follows that we cannot expect to do better than

$$r_k = \frac{\|z - x_k\|}{\|z\|} \approx \lambda_-.$$

If D and E are sufficiently small, then a Taylor expansion ensures that

$$\lambda_- \approx \frac{D}{(1 - EMK(1+D))^2}$$

is a good approximation. We cannot expect to do better than $r_{k+1} = \lambda_-$, but the threshold of stagnation is not particularly sensitive to the size of E.

3.2 The Decay of the Error

We assume that the sequences $\{D_k\}$ and $\{E_k\}$ are bounded. Let D and E be upper bounds that satisfy (8). Suppose that we are not near the threshold of stagnation in the sense that

$$D \leq Cr_k. \tag{9}$$

for a (modest) constant $C > 0$. In this case, inequality (7) implies

$$r_{k+1} \leq \rho_k r_k, \quad \rho_k = \frac{1}{2}LK(1+D)\|z\|r_k + EKM(1+D) + C. \tag{10}$$

If $C < 1$, then we may have $\rho_k < 1$, when r_k and E are sufficiently small. This explains when and why local linear decay is possible. We now strengthen our assumptions. Suppose that there is a $\lambda \in (0, 1]$ and $C_1 > 0$ such that

$$\|E_k\| \leq C_1 r_k^{\lambda} \tag{11}$$

and that we are far from the threshold of stagnation in the sense that

$$D \leq C_2 r_k^{1+\lambda} \tag{12}$$

for a (modest) constant $C_2 > 0$. In this case, inequality (7) implies

$$r_{k+1} \leq \left[\frac{1}{2} LK(1+D)\|z\| r_k^{1-\lambda} + C_1 KM(1+D) + C_2 \right] r_k^{1+\lambda}. \tag{13}$$

This explains when and why local superlinear decay is possible.

3.3 Convergence

We cannot expect a quasi-Newton method to converge unless the subtraction $y_{k+1} = y_k - t_k$ is exact. Then $D_k = 0$ and inequality (7) implies

$$r_{k+1} \leq \eta_k r_k, \quad \eta_k = \left(\frac{1}{2} LK\|z\| r_k + \|E_k\| KM \right).$$

We may have $\eta_k < 1$ for all k, provided $E = \sup \|E_k\|$ and r_0 are sufficiently small. This explains when and why local linear convergence is possible. We now strengthen our assumptions. Suppose that there is a $\lambda \in (0, 1]$ and a $C > 0$ such that

$$\forall k \in \mathbb{N} \; : \; \|E_k\| \leq C r_k^{\lambda}.$$

In this case, inequality (7) implies

$$r_{k+1} \leq \left(\frac{1}{2} LK\|z\| r_k^{1-\lambda} + CKM \right) r_k^{1+\lambda}.$$

This inequality allows us to establish local convergence of order at least $1 + \lambda$.

3.4 How Accurate Does Newton Have to Be?

We will assume the use of normal IEEE floating point numbers and we will apply the analysis given in Sect. 3.2. If we use the 1-norm, the 2-norm or the ∞-norm, then we may choose $D = u$, where u is the unit roundoff. Suppose that Eqs. (11) and (12) are satisfied with $\lambda = 1$. Then inequality (13) reduces to

$$r_{k+1} \leq \left[\frac{1}{2} LK(1+u)\|z\| + C_1 KM(1+u) + C_2 \right] r_k^2.$$

Due to the basic limitations of IEEE floating point arithmetic we cannot expect to do better than

$$r_{k+1} = O(u), \quad u \to 0, \quad u > 0.$$

It follows that we *never* need to do better than

$$\|E_k\| = O(\sqrt{u}), \quad u \to 0, \quad u > 0.$$

4 Numerical Experiments

4.1 Computing Square Roots

Let $\alpha > 0$ and consider the problem of solving the nonlinear equation

$$f(x) = x^2 - \alpha = 0$$

with respect to $x > 0$ using Newton's method. Let r_k denote the relative error after k Newton steps. A simple calculation based on Lemma 2 yields

$$|r_{k+1}| \leq r_k^2/2, \quad |r_k| \leq 2\left(|r_0|/2\right)^{2^k}.$$

We see that convergence is certain when $|r_0| < 2$. The general case of $\alpha > 0$ can be reduced to the special case of $\alpha \in [1, 4)$ by accessing and manipulating the binary representation directly. Let $x_0 : [1, 4] \to \mathbb{R}$ denote the best uniform linear approximation of the square root function on the interval $[1, 4]$. Then

$$x_0(\alpha) = \alpha/3 + 17/24, \quad |r_0(\alpha)| \leq 1/24.$$

In order to illustrate Theorem 1 we execute the iteration

$$x_{k+1} = x_k - (1 + e_k) f(x_k)/f'(x_k)$$

where e_k is a randomly generated number. Specifically, given $\epsilon > 0$ we choose e_k such that $|e_k|$ is uniformly distributed in the interval $[\frac{1}{2}\epsilon, \epsilon]$ and the sign of e_k is positive or negative with equal probability. Three choices, namely $\epsilon = 10^{-2}$ (left), $\epsilon = 10^{-8}$ (center) and $\epsilon = 10^{-12}$ (right) are illustrated in Fig. 1.

In each case, eventually the perturbed iteration reproduces either the computer's internal representation of the square root or stagnates with a relative error that is essentially the unit roundoff $u = 2^{-53} \approx 10^{-16}$. When $\epsilon = 10^{-2}$ the quadratic convergence is lost, but the relative error is decreased by a factor of approximately $\epsilon = 10^{-2}$ from one iteration to the next, i.e., extremely rapid linear convergence. Quadratic convergence is restored when ϵ is reduced to $\epsilon = 10^{-8} \approx \sqrt{u}$. Further reductions of ϵ have no effect on the convergence as demonstrated by the case of $\epsilon = 10^{-12}$. We shall now explain exactly how far this experiment supports the theory that is presented in this paper.

Stagnation. By Sect. 3.1 we expect that the level of stagnation is essentially independent of the size of E, the upper bound on the relative error between the computed step and the step needed for Newton's method. This is clearly confirmed by the experiment.

Error Decay. Since we are always very close to the positive zero of $f(x) = x^2 - \alpha$ we may choose

$$L \approx 2, \quad K|z| \approx 1/2, \quad MK \approx 1,$$

In the case of $\epsilon = 10^{-2}$, Fig. 1 (left) shows that we satisfy inequality (9) with $D = u$ and $C = \epsilon < 1$, i.e.,

$$u \leq \epsilon r_k, \quad 0 \leq k < 5.$$

By Eq. (10) we must have

$$r_{k+1} \leq \rho_k r_k, \quad \rho_k \approx 2\epsilon, \quad 0 < k < 5.$$

This is exactly the linear convergence that we have observed. In the case of $\epsilon = 10^{-8}$, Fig. 1 (center) shows that we satisfy inequality (12) with $C_2 = 1$ and $\lambda = 1$, i.e.,

$$u \leq r_k^2, \quad k = 0, 1.$$

By inequality (13) we must have quadratic decay in the sense that

$$r_{k+1} \leq C r_k^2, \quad C \approx \frac{3}{2}, \quad k = 0, 1.$$

Manual inspection of Fig. 1 reveals that the actual constant is close to 1 and certainly smaller than $C \approx \frac{3}{2}$. By Sect. 3.4 we do not expect any benefits from using an ϵ that is substantially smaller than $\sqrt{u}$. This is also supported by the experiment.

4.2 Constrained Molecular Dynamics

The objective is to solve a system of differential algebraic equations

$$\begin{aligned} q'(t) &= v(t), \\ Mv'(t) &= f(q(t)) - g'(q(t))^T \lambda(t), \\ g(q(t)) &= 0. \end{aligned}$$

Here q and v are vectors that represent the position and velocity of all atoms, M is a nonsingular diagonal mass matrix, f represents the external forces acting on the atoms and $-g'(q)^T\lambda$ represents the constraint forces. Here g' is the Jacobian of the constraint function g. The standard algorithm for this problem is the SHAKE algorithm [10]. It uses a pair of staggered uniform grids and takes the form

$$\begin{aligned} v_{n+1/2} &= v_{n-1/2} + hM^{-1}\left(f(q_n) - g'(q_n)^T\lambda_n\right), \\ q_{n+1} &= q_n + hv_{n+1/2}, \\ g(q_{n+1}) &= 0, \end{aligned} \tag{14}$$

where $h > 0$ is the fixed time step and $q_n \approx q(t_n)$, $v_{n+\frac{1}{2}} \approx v(t_{n+\frac{1}{2}})$, where $t_n = nh$ and $t_{n+\frac{1}{2}} = (n + 1/2)h$. Equation (14) is really a nonlinear equation for the unknown Lagrange multiplier λ_n, specifically

$$g(\phi_n(\lambda)) = 0, \quad \phi_n(\lambda) = q_n + h(v_{n-\frac{1}{2}} + hM^{-1}(f(q_n) - g'(q_n)^T\lambda)).$$

The relevant Jacobian is the matrix

$$A_n(\lambda) = (g(\phi_n(\lambda)))' = g'(\phi_n(\lambda))M^{-1}g'(q_n)^T.$$

The matrix $A_n(\lambda)$ is close to the constant symmetric matrix S_n given by

$$S_n = g'(q_n)M^{-1}g'(q_n)^T$$

simply because $\phi_n(\lambda) = q_n + O(h)$ as $h \to 0$ and $h > 0$. It is therefore natural to investigate if the constant matrix S_n^{-1} is a good approximation of $A_n^{-1}(\lambda)$.

For this experiment, we executed a production molecular dynamics run using the GROMACS [1] package. We replaced the constraint solver used by GROMACS's SHAKE function with a quasi-Newton method based on the matrix S_n. Our experiment was based on GROMACS's Lysozyme in Water Tutorial [6]. We simulated a hen egg white lysozyme [9] molecule submerged in water inside a cubic box. Lysozyme is a protein that consists of a single polypeptide chain of 129 amino acid residues cross-lined at 4 places by disulfide bonds between cysteine side-chains in different parts of the molecule. Lysozyme has 1960 atoms and 1984 bond length constraints. Before executing the production run, we added ions to the system to make it electrically neutral. The energy of the system was minimized using the steepest descent algorithm until the maximum force of the system was below 1000.0 kJ/(mol·nm). Then, we executed 100 ps of a temperature equilibration step using a V-Rescale thermostat in an NVT ensemble to stabilize the temperature of the system at 310 K. To finish, we stabilized the pressure of the system at 1 Bar for another 100 ps using a V-Rescale thermostat and a Parrinello-Rahman barostat in an NPT ensemble. We executed a 100 ps production run with a 2 fs time step using an NPT ensemble with a V-Rescale thermostat and a Parrinello-Rahman barostat with time constants of 0.1 and 2 ps, respectively. We collected the results of the constraint solver every 5 ps starting at time-step 5 ps, for a total of 20 sample points. Specifically, we recorded the normwise relative error $r_k = \|\lambda_n - x_k\|_2/\|\lambda_n\|_2$ as a function of the number k of quasi-Newton steps using the symmetric matrix S_n instead of the nonsymmetric matrix A_n and we recorded $\|E_k\|_2 = \|s_k - t_k\|_2/\|s_k\|_2$ where t_k is needed for a quasi-Newton step and s_k is needed a Newton step. By (10) we have $r_{k+1} \leq \rho_k r_k$, but we cannot hope for more than $r_{k+1} \approx \rho_k r_k$ where $\rho_k = O(\|E_k\|_2)$ and this is indeed what we find in the Fig. 2c until we hit the level of stagnation where the impact of rounding errors is keenly felt.

5 Related Work

It is well-known that Newton's method has local quadratic convergence subject to certain regularity conditions. The simplest proof known to us is due to Mysovskii [7]. Dembo et al. [2] analyzed the convergence of quasi-Newton methods in terms of the ratio between the norm of linear residual, i.e., $r_k = F(x_k) - F'(x_k)t_k$ and the norm of the nonlinear residual $F(x_k)$. Tisseur [11] studied the impact of rounding errors in terms of the backward error associated

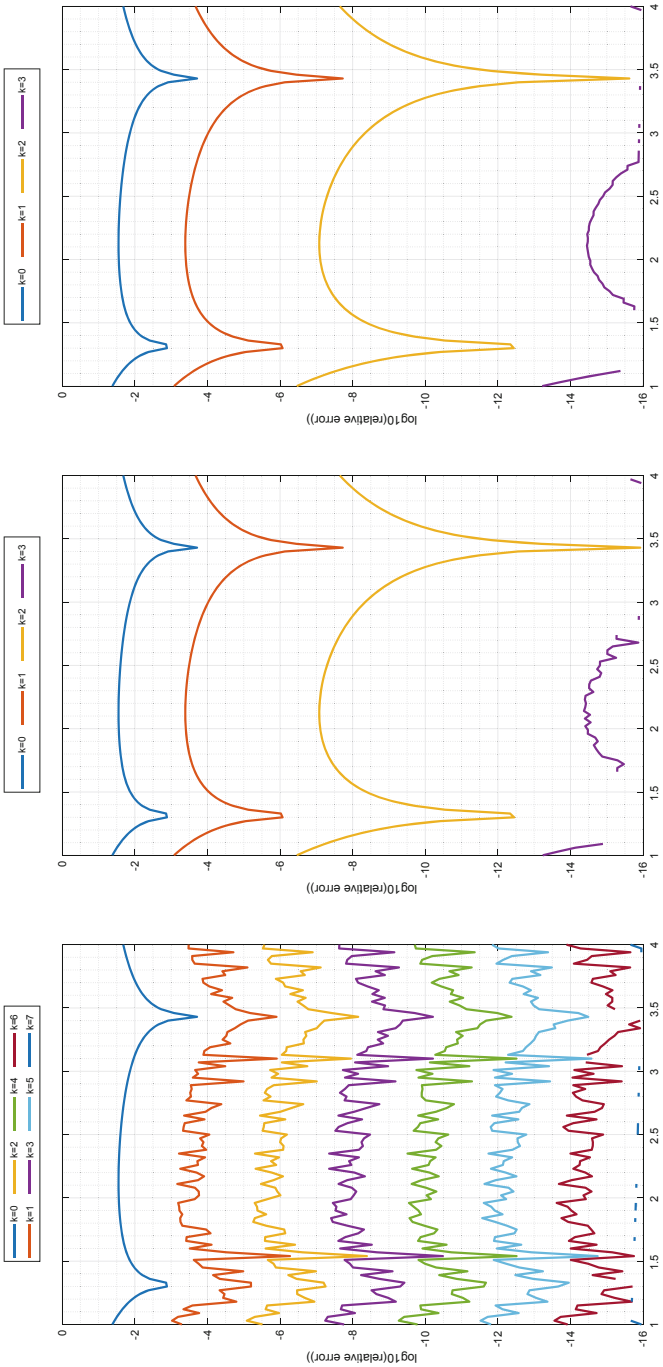
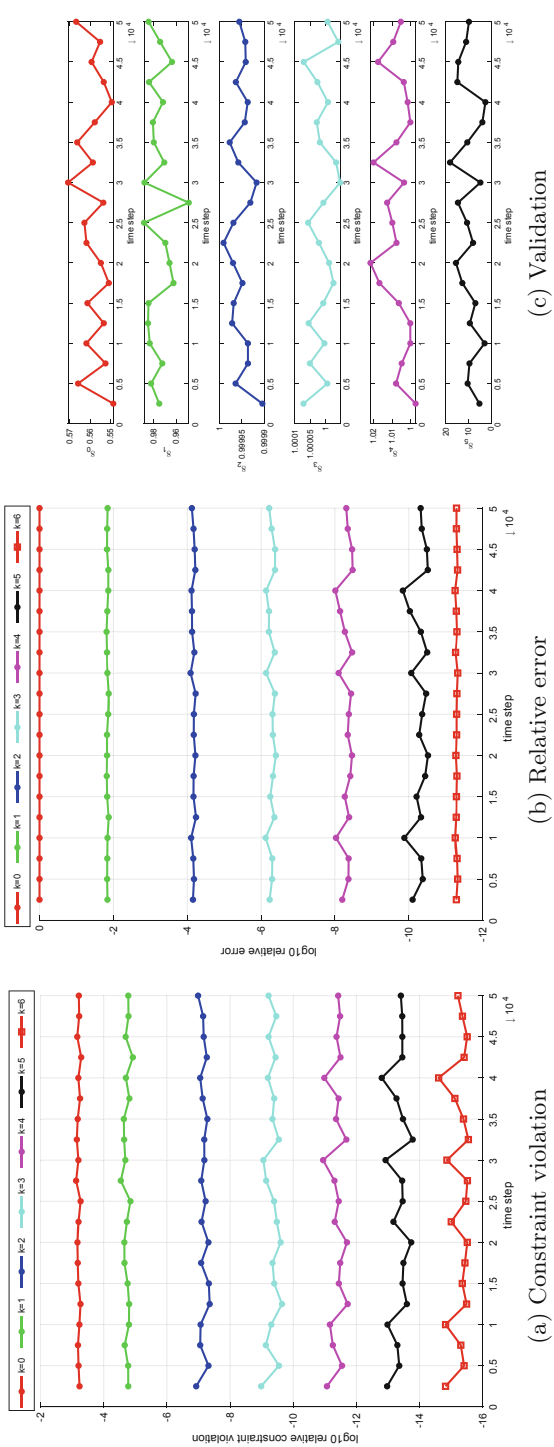


Fig. 1. The impact of inaccuracies on the convergence of Newton's method for a computing square roots. Newton's corrections have been perturbed with random relative errors of size $\epsilon \approx 10^{-2}$ (left), $\epsilon \approx 10^{-8}$ (center) and $\epsilon \approx 10^{-12}$. In each case, the last iteration produces an approximation that matches the computer's value of the square root at many sample points. In such cases, the computed relative error is 0. Therefore, it is not possible to plot a data point and the last curve of each plot are discontinuous.



(a) Constraint violation (b) Relative error (c) Validation

Fig. 2. Data generated during a simulation of lysozyme in water using GROMACS. The GROMACS solver have been replaced with a quasi-Newton method that uses a fixed symmetric approximation of the Jacobian. Figure 2a is mainly of interest to computational chemists. It shows that the maximum relative constraint violation always stagnates at a level that is essentially the IEEE double precision unit roundoff after 6 quasi-Newton steps. The convergence is always linear and the rate of convergence is $\mu \approx 10^{-2}$. Figure 2b shows the development of the relative error r_k between the relevant zero z, i.e., the Lagrange multiplier for the current time step and the approximations generated by k steps of the quasi-Newton method. The convergence is always linear and the rate of convergence is $\mu \approx 10^{-2}$. Figure 2c provides partial validation of a theoretical result. Specifically, the fractions $\nu_k = r_{k+1}/(r_k\|E_k\|_2)$ are plotted for $k = 0, 1, 2, 3, 4, 5$. When ν_k is modest, we have experimental verification that the rate of convergence is essentially $\|E_k\|$.

with approximating the Jacobians and computing the corrections, as well as the errors associated with computing the residuals. Here we have pursued a third option by viewing the correction t_k as an approximation of the correction s_k needed for an exact Newton step. Tisseur found that Newton's method stagnate at a level that is essentially independent of the stability of the solver and we have confirmed that this is true for quasi-Newton methods in general. It is clear to us from reading Theorem 3.1 of Dennis and Moore's paper [3] that they would instantly recognize Lemma 3, but we cannot find the result stated explicitly anywhere. Forsgren [4] uses a stationary method for solving linear systems to construct a quasi-Newton method that is so exact that the convergence is quadratic. Section 4.1 contains a simple illustration of this phenomenon.

6 Conclusions

Quasi-Newton methods can also be analyzed in terms of the relative error between Newton's correction and the computed correction. We achieve quadratic convergence when this error is $O(\sqrt{u})$. This fact represent an opportunity for improving the time-to-solution for nonlinear equations. General purpose libraries for solving sparse linear systems apply pivoting for the sake of numerical accuracy and stability. In the context of quasi-Newton methods we do not need maximum accuracy. Rather, there is some freedom to pivot for the sake of parallelism. If we fail to achieve quadratic convergence, then we are likely to still converge rapidly. It is therefore worthwhile to develop sparse solvers that pivot mainly for the sake of parallelism.

Acknowledgments. Prof. I. Argyros commented on an early draft of this paper and provided the reference to the work of I. P. Mysovskii. The first author is supported by eSSENCE, a collaborative e-Science programme funded by the Swedish Research Council within the framework of the strategic research areas designated by the Swedish Government. This work has been partially supported by the Spanish Ministry of Science and Innovation (contract PID2019-107255GB-C21/AEI/10.13039/501100011033), by the Generalitat de Catalunya (contract 2017-SGR-1328), and by Lenovo-BSC Contract-Framework Contract (2020).

References

1. Berendsen, H., van der Spoel, D., van Drunen, R.: GROMACS: a message-passing parallel molecular dynamics implementation. CPC **91**(1), 43–56 (1995)
2. Dembo, R.S., Eisenstat, S.C., Steihaug, T.: Inexact Newton methods. SIAM J. Numer. Anal. **19**(2), 400–408 (1982)
3. Dennis, J.E., More, J.J.: Quasi-Newton methods, motivation and theory. SIAM Rev. **19**(1), 46–89 (1977)
4. Forsgren, A.: A sufficiently exact inexact Newton step based on reusing matrix information. TRITA-MAT OS7, Department of Mathematics, KTH, Stockholm, Sweden (2009)
5. Kelley, C.T.: Iterative Methods for Linear and Nonlinear Equations. No. 16 in Frontiers in Applied Mathematics. SIAM, Philadelphia (1995)

6. Lemkul, J.A.: GROMACS Tutorial Lysozyme in Water. https://www.mdtutorials.com/gmx/lysozyme/index.html
7. Mysovskii, I.P.: On the convergence of Newton's method. Trudy Mat. Inst. Steklova **28**, 145–147 (1949). (In Russian)
8. Ortega, J.M., Rheinboldt, W.C.: Iterative Solution of Nonlinear Equations in Several Variables. Computer Science and Applied Mathematics, Academic Press, New York (1970)
9. RSCB: Protein Data Bank. https://www.rcsb.org/structure/1AKI
10. Ryckaert, J.P., Ciccotti, G., Berendsen, H.J.: Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J. Comput. Phys. **23**(3), 327–341 (1977)
11. Tisseur, F.: Newton's method in floating point arithmetic and iterative refinement of generalized eigenvalue problems. SIAM J. Matrix Anal. Appl. **22**(4), 1038–1057 (2001)