CONJUGATE GRADIENT ALGORITHM AND FRACTALS

MOHAMED LAMINE SAHARI AND ILHEM DJELLIT

This work is an extension of the survey on Cayley's problem in case where the conjugate gradient method is used. We show that for certain values of parameters, this method produces beautiful fractal structures.

1. Introduction

It is well known that the basins of attraction of different roots have fractal boundaries when Newton's method is used to determine the complex roots of polynomials. The simplest example was given by Cayley in 1879 [2]; it is the case for the equation

$$z^3 - 1 = 0. \tag{1.1}$$

The works of Cayley have been taken then by the French mathematician Julia [6] to found a theory on sets that will carry his name. The survey of this type of sets has been thrown back there after 1975 by another French mathematician, Mandelbrot [7], with the help of computers and melting; that is what is known by the fractal geometry. In this work we are going to give another approach precisely to Cayley's problem; we are going to study the case where an algorithm of the generalized conjugate gradient (GCG) is used and one is going to show that the qualitative and geometric aspects of the basins of attraction for each solution of (1.1) depend on certain parameters, sometimes, driving to beautiful fractal structures. We consider the nonlinear equation

$$f(z) = 0, \tag{1.2}$$

where f is complex function with variable z. To solve (1.2), we can use the conjugate gradient method (CG) defined by a recurrent sequence of form

$$z_{k+1} = z_k + \lambda_k d_k \tag{1.3}$$

and the search direction d_k is defined by

$$d_0 = -f(z_0), \quad d_k = -f(z_k) + \beta_k d_{k-1}, \quad k > 0. \tag{1.4}$$

The case where f is linear, parameters (λ_k, β_k) are well defined and permit a convergence of the method (1.3)–(1.4) in few iterations. In the nonlinear case, the choice of these parameters is not obvious. The well-know formulas for β_k are Fletcher-Reeves (FR) [4, 5] and Plolak-Ribière (PR) [9], and they are given by

$$\beta_k^{FR} = \frac{|f(z_k)|^2}{|f(z_{k-1})|^2}, \tag{1.5}$$

$$\beta_k^{PR} = \frac{\operatorname{Re}\left(\overline{f(z_k)}[f(z_k) - f(z_{k-1})]\right)}{|f(z_{k-1})|^2}, \tag{1.6}$$

where $\overline{z}$ denotes conjugate of z, $\operatorname{Re}(z) = (1/2)(\overline{z} + z)$, and $|z|^2 = \overline{z} \cdot z$. The choice of step length λ_k is delicate because it is obtained by solving the equation $f(z_k + \lambda_k d_k) = 0$, which is absurd. Therefore we prefer to take λ_k as approximal solution of problem

$$\min_{\lambda > 0} |f(z_k + \lambda d_k)|^2; \tag{1.7}$$

it is called an inexact line search [8], and several methods are proposed among them: Wolfe line search (W-L.S.) and given by: if one puts $\theta(\lambda) \equiv |f(z_k + \lambda d_k)|^2$ and $\theta'(\lambda) \equiv (d/d\lambda)\theta(\lambda)$, the step length $\lambda_k > 0$ satisfies the following conditions:

$$\theta(\lambda_k) \le m_1 \theta'(0)\lambda_k + \theta(0), \qquad \theta'(\lambda_k) \ge m_2 \theta'(0), \tag{1.8}$$

with $0 < m_1 < m_2 < 1$.

Note that in the case where f would be nonlinear, results of convergence (and counter example) are rare. It has been established in [9, 10] the convergence of the two versions FR and PR with a chosen step length after solving exactly the problem (1.7). In [1] Al-Baali proved a global convergence of FR method with strong Wolfe line searches, that is,

$$\theta(\lambda_k) \le m_1 \theta'(0)\lambda_k + \theta(0), \qquad |\theta'(\lambda_k)| \le -m_2 \theta'(0). \tag{1.9}$$

2. Notations and definitions

We denote by $\mathbb{C}$ the complex plan, $\mathbb{R}$ a real line, and $\mathbb{M}$ nonempty open subset of $\overline{\mathbb{C}}^n (n = 1, 2, \ldots)$, where $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$.

Let us consider the complex mapping

$$z_{k+1} = g(z_k, \beta) \equiv g(z_k), \tag{2.1}$$

where $g : \mathbb{M} \to \mathbb{M}$ is holomorphic function and β is complex parameter.

Definition 2.1. A point $\omega \in \mathbb{M}$ is a fixed point of g if

$$g(\omega) = \omega. \tag{2.2}$$

ω is attracting if

$$\left| \frac{d}{dz} g(z) \right|_{z=\omega} < 1. \tag{2.3}$$

ω is repelling if

$$\left| \frac{d}{dz} g(z) \right|_{z=\omega} > 1. \tag{2.4}$$

Definition 2.2. The basin of attraction $\mathcal{A}(\omega)$ of an attractive fixed point $\omega \in \mathbb{M}$, associated with a function g, is defined by

$$\mathcal{A}(\omega) = \left\{ z \in \mathbb{M} : g^k(z) \xrightarrow[k \to \infty]{} \omega \right\}. \tag{2.5}$$

Definition 2.3. Let $\omega \in \mathbb{M}$ be an attractive fixed point of g, then the disk of convergence $\mathcal{D}(z, \beta)$ associated with the map (2.1) is defined by

$$\mathcal{D}(z, \beta) = \left\{ \beta \in \mathbb{C} : g^k(z, \beta) \xrightarrow[k \to \infty]{} \omega \right\}. \tag{2.6}$$

In the literature [3], if $g(z) = z^2 + \beta$ and $\mathbb{M} = \mathbb{C}$, the closure of $\mathcal{D}(0, \beta)$ is not other than a Mandelbrot set and $\partial \mathcal{A}(\omega)$ is the Julia set, ($\partial \mathcal{A}$ denotes a boundary of $\mathcal{A}$).

Definition 2.4. Let $\mathscr{C}_{z_I, z_S}$ be a cone in $\mathbb{C}$ generated by the complex z_I, and z_S is defined by

$$\mathscr{C}_{z_I, z_S} = \left\{ \nu z_I + \xi z_S; \ \nu > 0, \ \xi > 0 \right\}. \tag{2.7}$$

3. Generalized conjugate gradient method (GCG)

Instead of choosing the reals (λ_k, β_k) at every iteration of the CG method, we fix the real or complex parameters, and the sequences (3.1) may be generalized in the following form:

$$\mathcal{S}_{\lambda, \beta} : \begin{cases} z_{k+1} = z_k + \lambda w_k, \\ w_{k+1} = -f(z_k + \lambda w_k) + \beta w_k, \end{cases} \tag{3.1}$$

where $(z_k, w_k, \lambda, \beta) \in \mathbb{C}^4$, $w_0 = 0$, and thus (3.1) is called generalized conjugate gradient method (GCG).

3.1. Convergence analysis. We verify that the map (3.1) would admit some stationary points if $\lambda \neq 0$ and such points $(z^*, 0)$ verify $f(z^*) = 0$; the Jacobian matrix is given by

$$J_{\mathcal{S}_{\lambda, \beta}}(z, w) = \begin{bmatrix} 1 & \lambda \\ -f'(z + \lambda w) & -\lambda f'(z + \lambda w) + \beta \end{bmatrix}; \tag{3.2}$$

hence at stationary points $(z^*,0)$, we have

$$J_{\mathscr{S}_{\lambda,\beta}}(z^*,0) = \begin{bmatrix} 1 & \lambda \\ -f'(z^*) & -\lambda f'(z^*) + \beta \end{bmatrix}. \tag{3.3}$$

The eigenvalues $\mu_1 \equiv \mu_1(z^*,\lambda,\beta)$ and $\mu_2 \equiv \mu_2(z^*,\lambda,\beta)$ of $J_{\mathscr{S}_{\lambda,\beta}}(z^*,0)$ are

$$\begin{aligned} \mu_1 &= \frac{1}{2}\left[1 - \lambda f'(z^*) + \beta + \sqrt{\left(1 - \lambda f'(z^*) + \beta\right)^2 - 4\beta} \right], \\ \mu_2 &= \frac{1}{2}\left[1 - \lambda f'(z^*) + \beta - \sqrt{\left(1 - \lambda f'(z^*) + \beta\right)^2 - 4\beta} \right]. \end{aligned} \tag{3.4}$$

PROPOSITION 3.1. *If GCG method converge, then*

$$|1 - \lambda f'(z^*) + \beta| < 2, \quad |\beta| < 1. \tag{3.5}$$

Proof. If $|\beta| > 1$ or $|1 - \lambda f'(z^*) + \beta| > 2$, then we deduce that $(z^*,0)$ is repulsive. Indeed let us put $|\mu|_{\max} = \max(|\mu_1|,|\mu_2|)$; it follows that

$$\begin{aligned} |\beta| &= |\det(J_{\mathscr{S}_{\lambda,\beta}}(z^*,0))| = |\mu_1| \cdot |\mu_2| \le |\mu|_{\max}^2, \\ |1 - \lambda f'(z^*) + \beta| &= |\operatorname{Tr}(J_{\mathscr{S}_{\lambda,\beta}}(z^*,0))| = |\mu_1 + \mu_2| \le 2|\mu|_{\max}. \end{aligned} \tag{3.6}$$

$\square$

Unfortunately, the conditions (3.5) are not sufficient, for example if we take $1 - \lambda f'(z^*) + \beta = 1$ and $\beta = 3/4$, we obtain $\mu_1 = (1/2)(1 + i\sqrt{2}) \implies |\mu_1| = \sqrt{3/2} > 1$. Nevertheless, we can find some strict conditions as in the following case.

PROPOSITION 3.2. *Let $f : \mathbb{C} \to \mathbb{C}$ be holomorphic function and z^* a solution of $f(z) = 0$, then if $|\beta| < \gamma^2$ and $|1 - \lambda f'(z^*)| < \gamma - \gamma^2$, where $\gamma \le 2/\sqrt{5} + 1$ (inverse of gold number), the fixed point $(z^*,0)$ is attractive for $\mathscr{S}_{\lambda,\beta}$.*

Therefore in neighborhood of z^*, GCG method generates a sequence $\{z_k\}_{k=0,1,2\ldots}$ such that $z_k \to z^*$.

Proof. Denote $\alpha = |1 - \lambda f'(z^*) + \beta|$, for $|\beta| < \gamma^2$ and $|1 - \lambda f'(z^*)| < \gamma - \gamma^2$, we can have $\alpha < \gamma$.

For $i = 1,2$, we have then

$$2|\mu_i| \le \alpha + \sqrt{\alpha^2 - 4\beta} < \gamma(1 + \sqrt{5}) \le 2 \tag{3.7}$$

and thus

$$|\mu_i| < 1, \quad i = 1,2. \tag{3.8}$$

□

Remark 3.3. We can use $|1 - \lambda f'(z^*)| < \gamma - \gamma^2$ as a condition to satisfy for convergence of GCG, however if $\lambda > 0$ (the case of a linear search), it is necessary that

$$f'(z^*) \in \mathscr{C}_{z_I, z_S} \tag{3.9}$$

with $z_I = \delta - i\Delta$ and $z_S = \delta + i\Delta$, where $\delta = 1 - (\gamma - \gamma^2)^{3/2}$ and $\Delta = (\gamma - \gamma^2)\sqrt{1 - (\gamma - \gamma^2)}$.

But, as we will see further the condition (3.9) is still not true.

4. Application to Cayley's problem

Cayley asked about the problem of characterization of the basins of attraction of zeroes for the cubic complex polynomial

$$f(z) = z^3 - 1. \tag{4.1}$$

The three roots of the equation $f(z) = 0$ are

$$z_k = e^{(2/3)(k-1)\pi i}, \quad k = 1, 2, 3. \tag{4.2}$$

Let us begin to analyze an equivalent problem that consists in minimizing the square of the following residual:

$$\min_{z \in \mathbb{C}} |f(z)|^2. \tag{4.3}$$

To solve the problem (4.3), return to solve the following nonlinear equation:

$$\overline{f'(z)} \cdot f(z) = 0. \tag{4.4}$$

Apply the GC method on this last equation, we obtain the following sequence in $\mathbb{C}$:

$$z_{k+1} = z_k + \lambda_k \left[-\overline{f'(z)} \cdot f(z) + \frac{\beta_k}{\lambda_{k-1}} (z_k - z_{k-1}) \right], \tag{4.5}$$

with $\beta_0 = 0$, and for $k > 0$, we choose between β_k^{FR} and β_k^{PR} in accordance with (1.5) or (1.6), and λ_k is obtained by line search. Results of numerical test (**Figure 4.1**) confirm the superiority of the PR method on FR method [7]. Note that we can not apply the analysis for **Section 3** to (4.5), because the complex function $\overline{f'} \cdot f$ is not holomorphic. On the other hand, $|f|^2$ is locally convex and the theory developed in [1, 4, 8, 9] assures the local convergence.
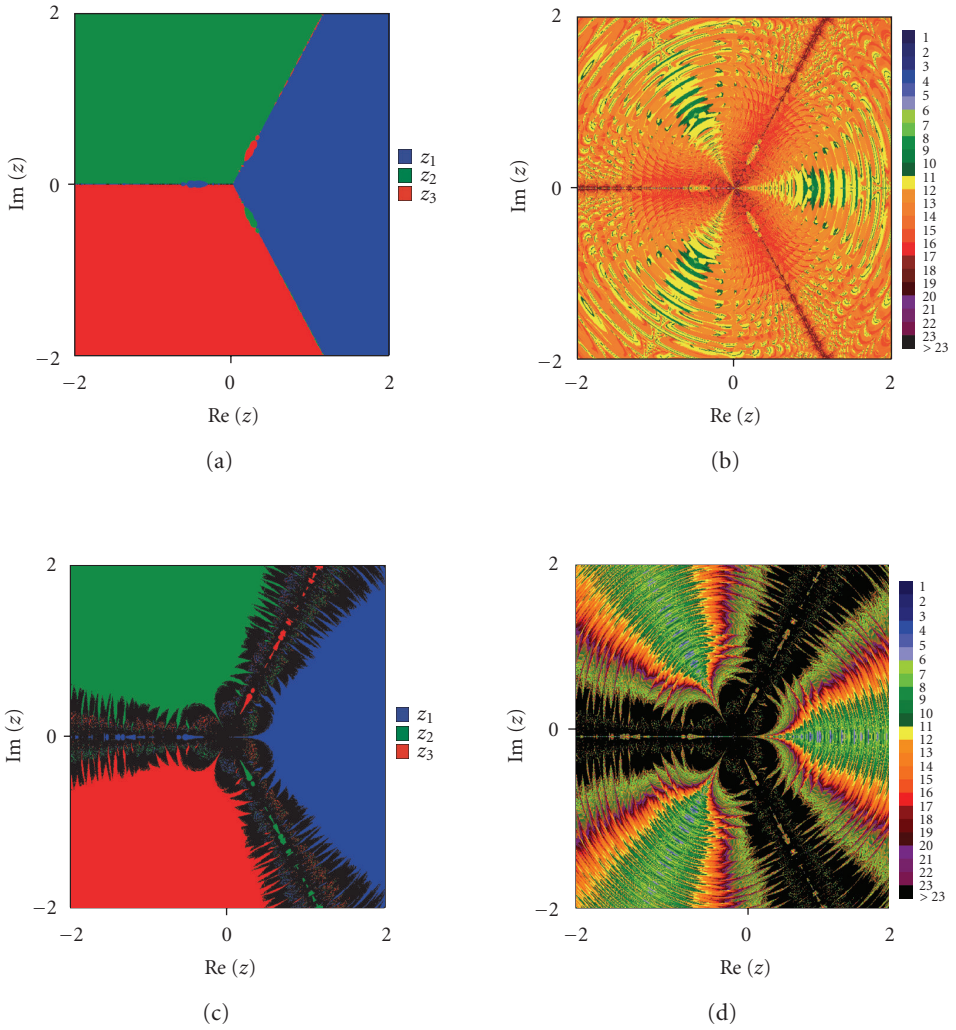
Figure 4.1. Basins of attraction for $z^3 - 1 = 0$. (a)–(b) Using CG-PR method and W-L.S. (c)–(d) Using CG-FR method and W-L.S.

Now the residual is used only to do a line search, the sequence (4.5) will take the following form:

$$z_{k+1} = z_k + \lambda_k \left[1 - z^3 + \frac{\beta_k}{\lambda_{k-1}} (z_k - z_{k-1}) \right]. \tag{4.6}$$

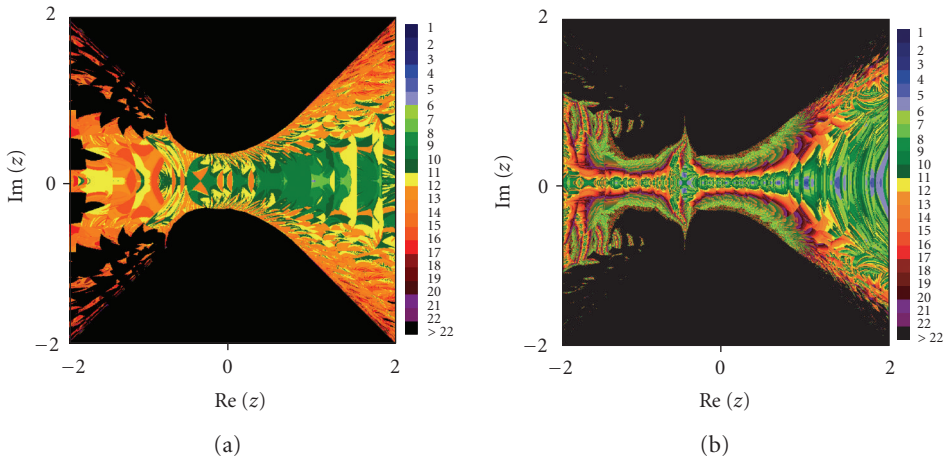In Figure 4.2, we show the basins of solutions of $z^3 - 1 = 0$ with the process (4.6) and $\beta \equiv \beta_k$ chosen making use FR or PR and Wolfe line search.

Figure 4.2. Basins of attraction for $z^3 - 1 = 0$ using (4.6). (a) PR method and W-L.S. (b) FR method and W-L.S.
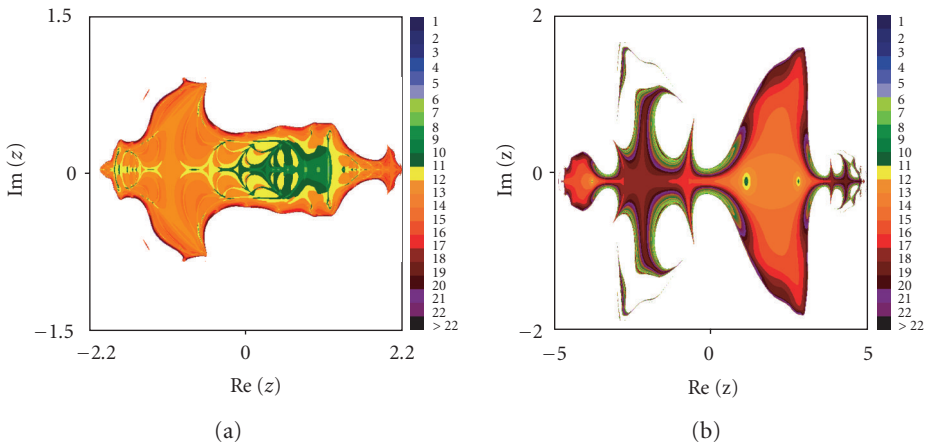
Figure 4.3. Basins of attraction of z_1 using (4.6). (a) With PR method and $\lambda = 0.4$. (b) With FR method and $\lambda = 0.1$.

If we choose β according to PR or FR and for any $\lambda \equiv \lambda_k > 0$ arbitrary small (Figure 4.3), we remark that we can obtain a local convergence for (4.6) method.

The local convergence is possible with different line searches and varying β in the range $(0,1)$ (Figure 4.4).

It is necessary to note that contrarily to the method (4.5), where there is coexistence of three basins $\mathscr{A}(z_1)$, $\mathscr{A}(z_2)$, and $\mathscr{A}(z_3)$, in Figures 4.2, 4.3, and 4.4, we show that only $\mathscr{A}(z_1)$ is nonempty. Indeed $f'(z_1) = 3 \in \mathscr{C}_{z_I, z_S}$, which implies a local convergence for the
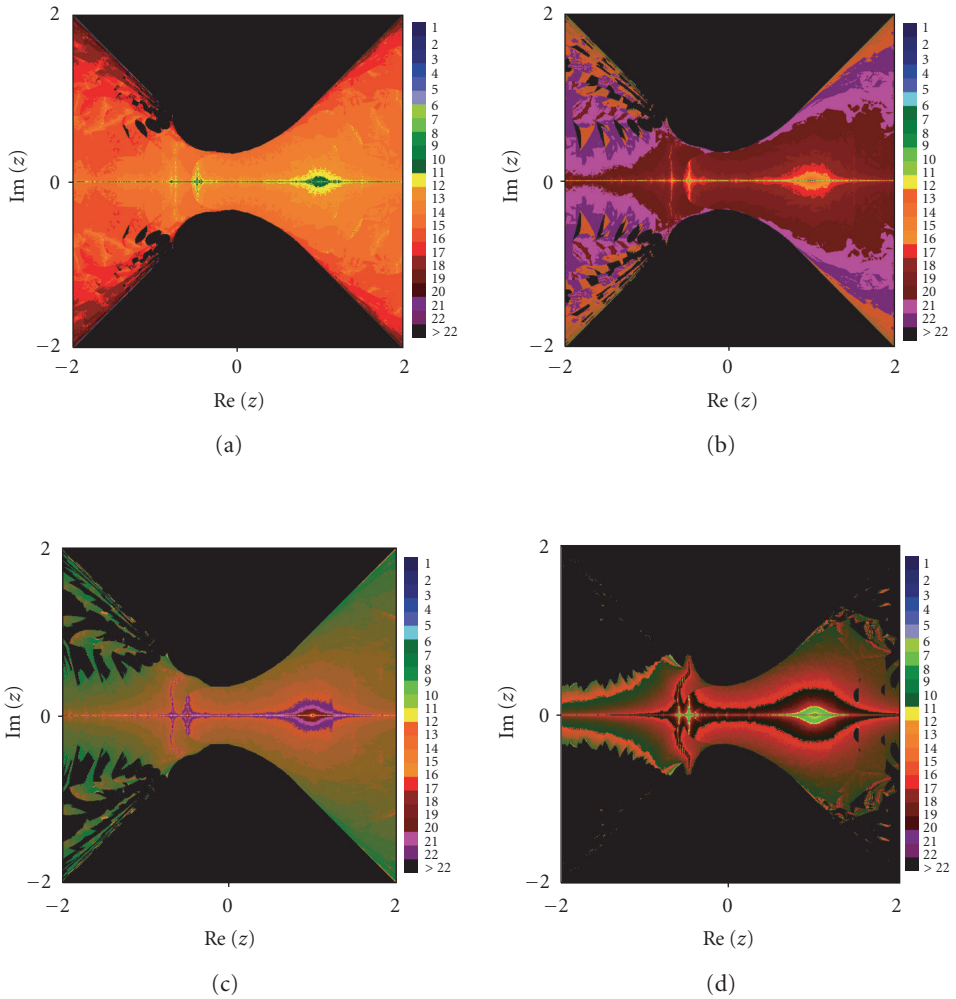
Figure 4.4. Basins of attraction of z_1 using (4.6) and W-L.S. with: (a) $\beta = 0.1$, (b) $\beta = 0.2$, (c) $\beta = 0.3$, (d) $\beta = 0.7$.

solution z_1 (according to Proposition 3.2). On the other hand, an analytical study of real functions $(\lambda, \beta) \to |\mu_1(z_1, \lambda, \beta)| - 1$ and $(\lambda, \beta) \to |\mu_1(z_2, \lambda, \beta)| - 1$ (see Figure 4.5(b)) shows that $|\mu_1(z_1, \lambda, \beta)| > 1$ and $|\mu_1(z_2, \lambda, \beta)| > 1$ for $\lambda > 0$ and $\beta \in \mathbb{R}$; then the points z_2 and z_3 are repulsive in this case.

This fact is confirmed by the bifurcation diagram of GCG method (Figure 4.6); but adding to that it shows the existence of convergence zone for solutions z_2 and z_3 if $\lambda < 0$.

If values (λ, β) are in convergence area of the diagram in Figure 4.6, the result will be that the basins of attraction of the roots z_2 and z_3 are nonempty, and we remark clearly the fractal structure (Figure 4.7).
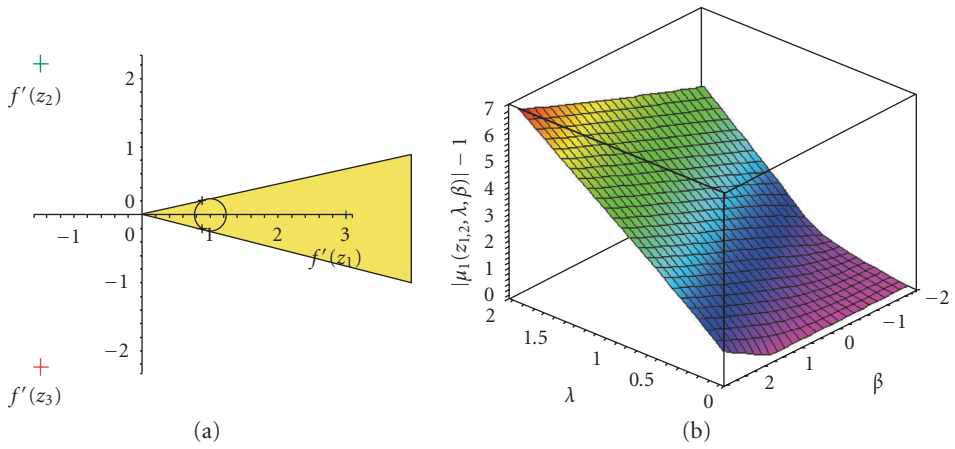
Figure 4.5. (a) The cone $\mathscr{C}_{z_1,z_S}$. (b) Function graph $(\lambda,\beta) \to |\mu_1(z_i,\lambda,\beta)| - 1$ $(i = 1,2)$.
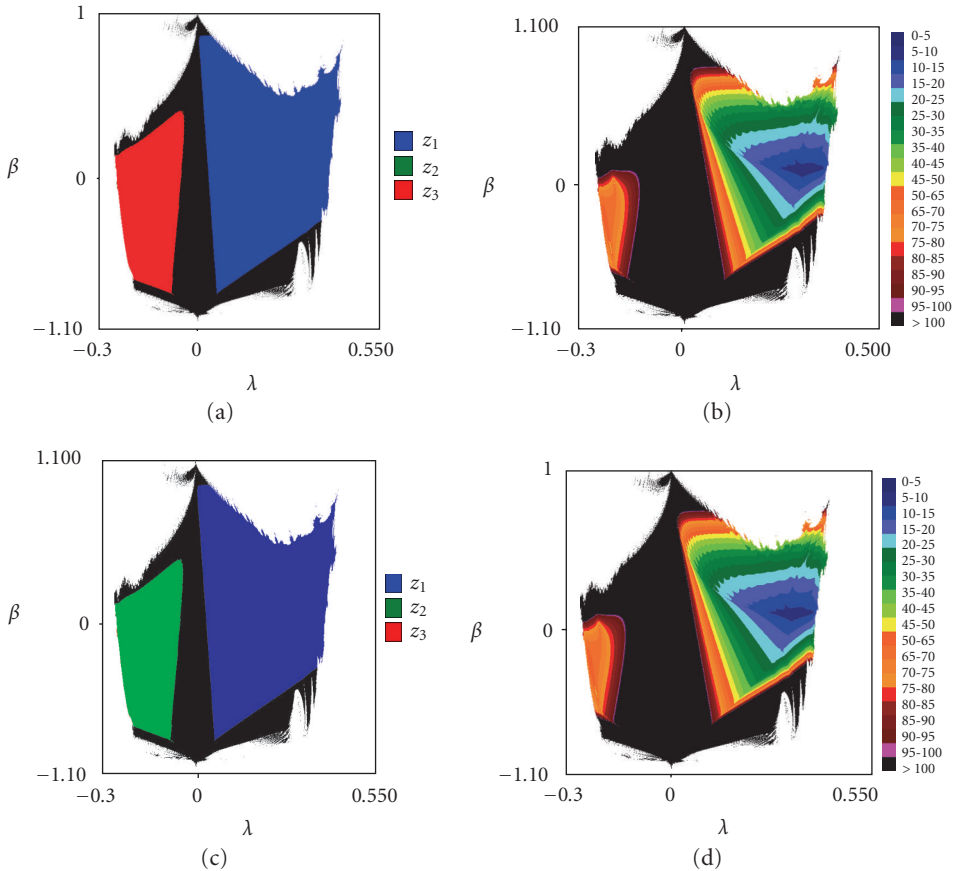

Figure 4.6. Bifurcations diagram of GCG method in (λ,β)-plane with the following starting points: (a)–(b) $z_0 = -1.2 + i \cdot 0.1$, (c)–(d) $z_0 = -1.2 - i \cdot 0.1$.
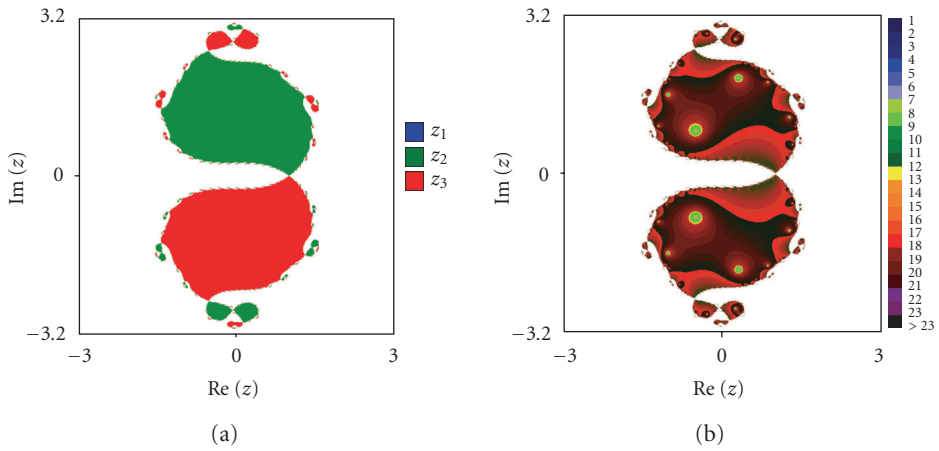
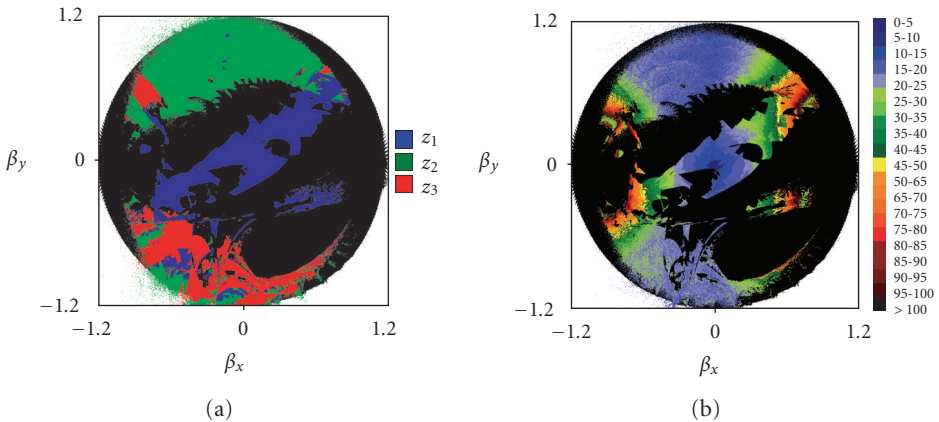Figure 4.7. Basin of attraction for $z^3 - 1 = 0$ using GCG method with $\lambda = -0.2$ and $\beta = -0.2$.

Figure 4.8. Disks of convergence $\mathscr{D}(-1.2 + i, \beta)$ with W-L.S.

Take $\beta = \beta_x + i \cdot \beta_y$, according to bifurcation diagram in the parametric plane (β_x, β_y), from GCG method, with or without line search, one can detect convergence areas for the three roots z_1, z_2, and z_3 in a disk $\mathscr{D}(z_0, \beta)$ centered in $(0,0)$ and of radius strictly lower than 1 (Figures 4.8, 4.9), which confirms Proposition 3.1.

Exploiting the diagram in Figures 4.9(b), 4.9(c), 4.9(e), and 4.9(f), we can find that values for each basin of attraction $\mathscr{A}(z_1), \mathscr{A}(z_2)$, and $\mathscr{A}(z_3)$ are nonempty (Figure 4.10).

The same thing if we fix λ and β, we can obtain basins of attraction $\mathscr{A}(z_1), \mathscr{A}(z_2)$, and $\mathscr{A}(z_3)$ with fractal boundary (Figure 4.11).

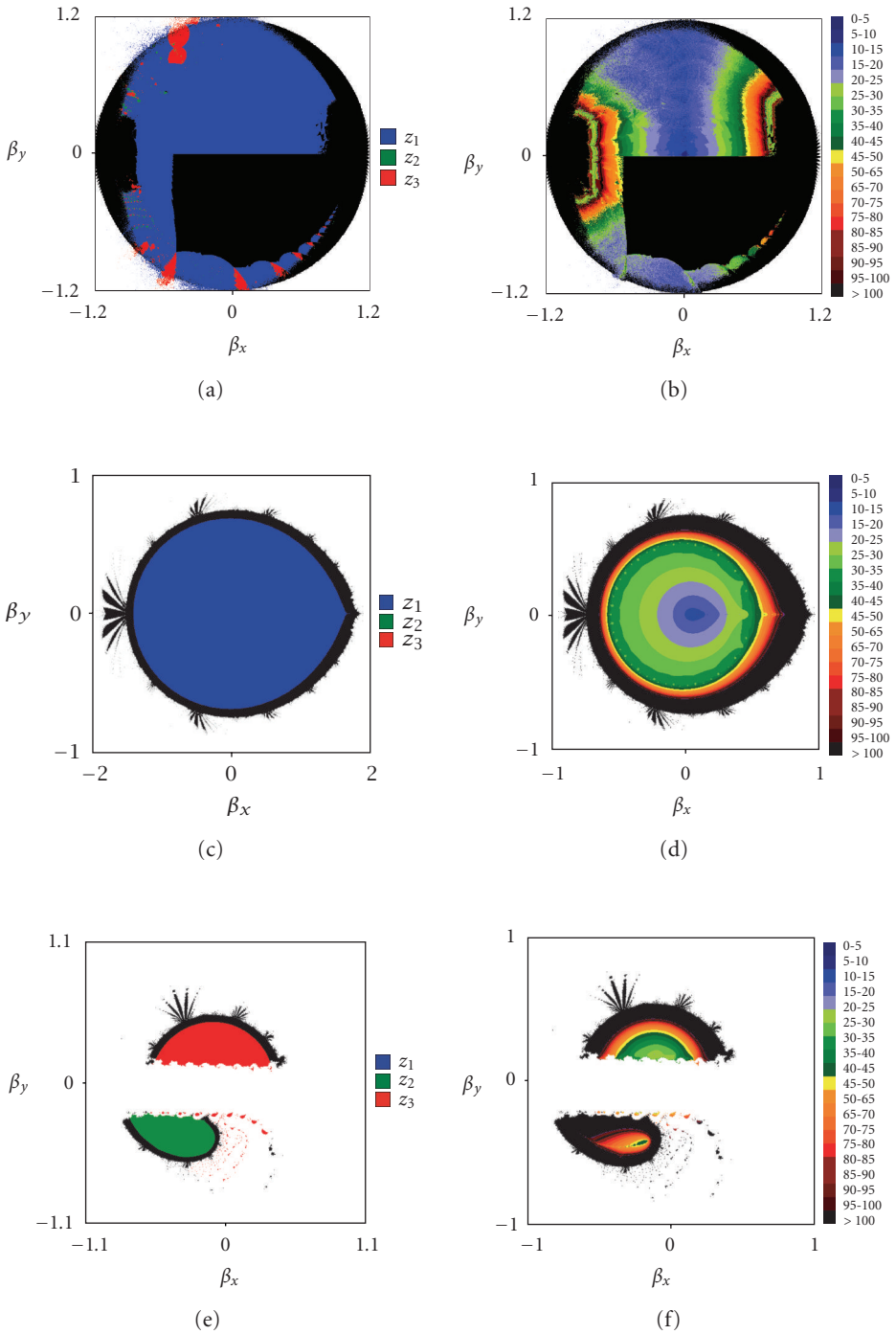In Figure 4.12, we explore one of fractal structure's features that is the self-similarity.

Figure 4.9. Disks of convergence: (a)–(b) $\mathcal{D}(1.2 - i, \beta)$ with W-L.S., (c)–(d) $\mathcal{D}(0.0, \beta)$ with $\lambda = 0.2$, (e)–(f) $\mathcal{D}(0.5 + i, \beta)$ with $\lambda = -0.2 - i \cdot 0.1$.
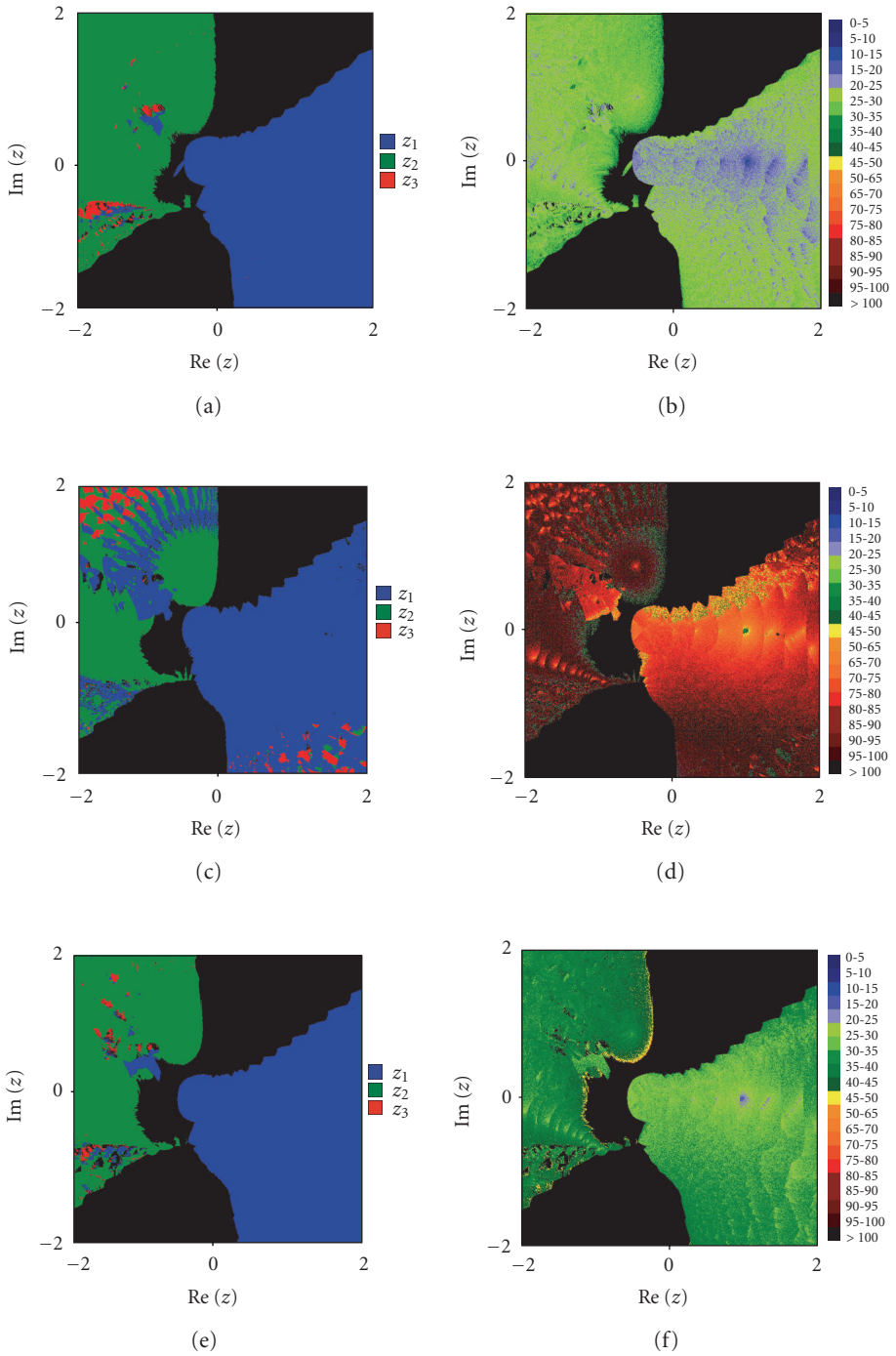
Figure 4.10. Basin of attraction for $z^3 - 1 = 0$ using GCG method and W-L.S. with: (a)–(b) $\beta = 0.360 + i \cdot 0.840$, (c)–(d) $\beta = 0.716 + i \cdot 0.5724$, (e)–(f) $\beta = 0.456 + i \cdot 0.728$.
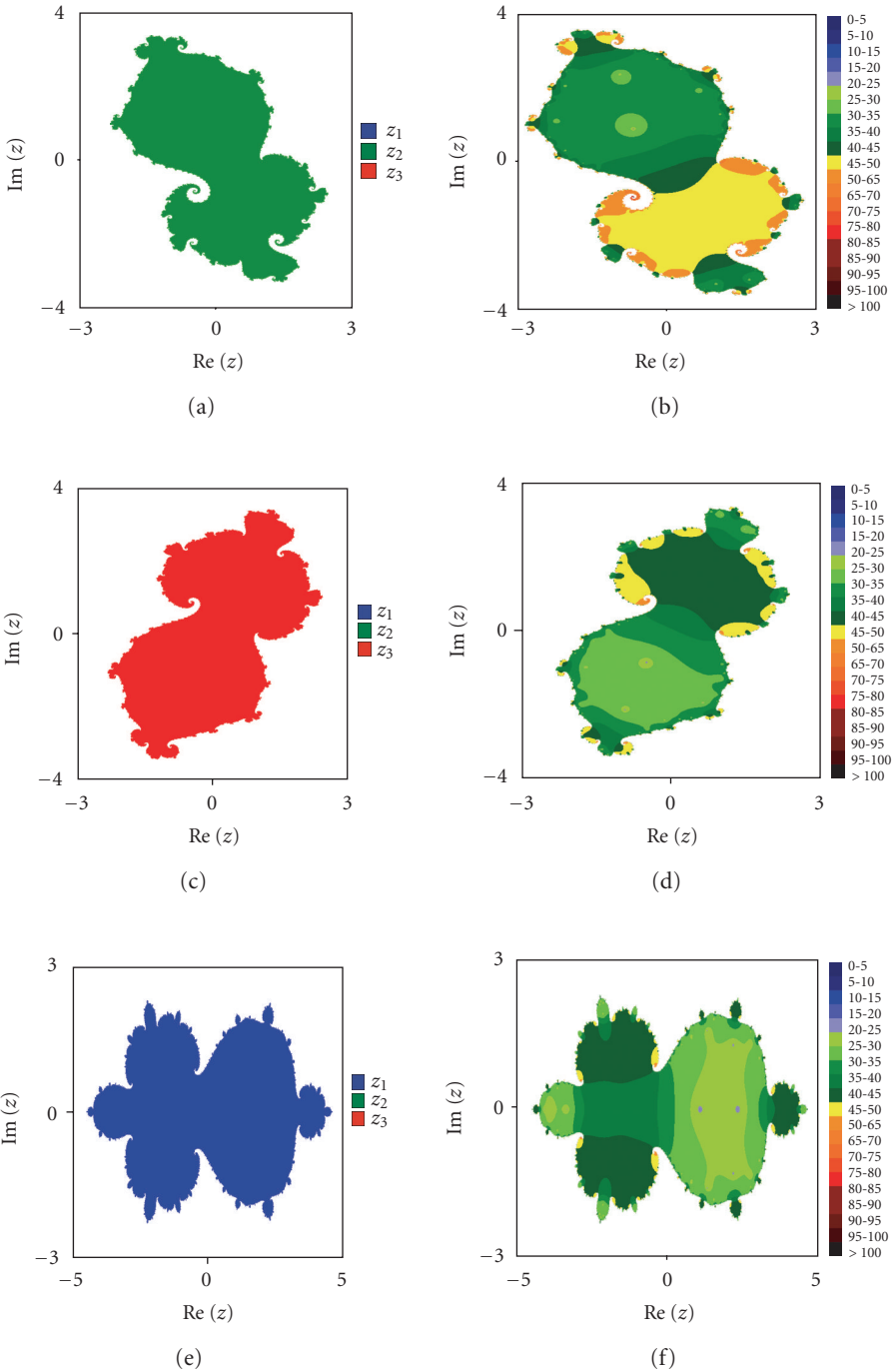
Figure 4.11. Basin of attraction for $z^3 - 1 = 0$ using GCG method with: (a)–(b) $\lambda = -0.2 + i \cdot 0.1$ and $\beta = -0.1 + i \cdot 0.1$, (c)–(d) $\lambda = -0.2 + i \cdot 0.1$ and $\beta = -0.1 - i \cdot 0.1$, (e)–(f) $\lambda = \beta = 0.1$.
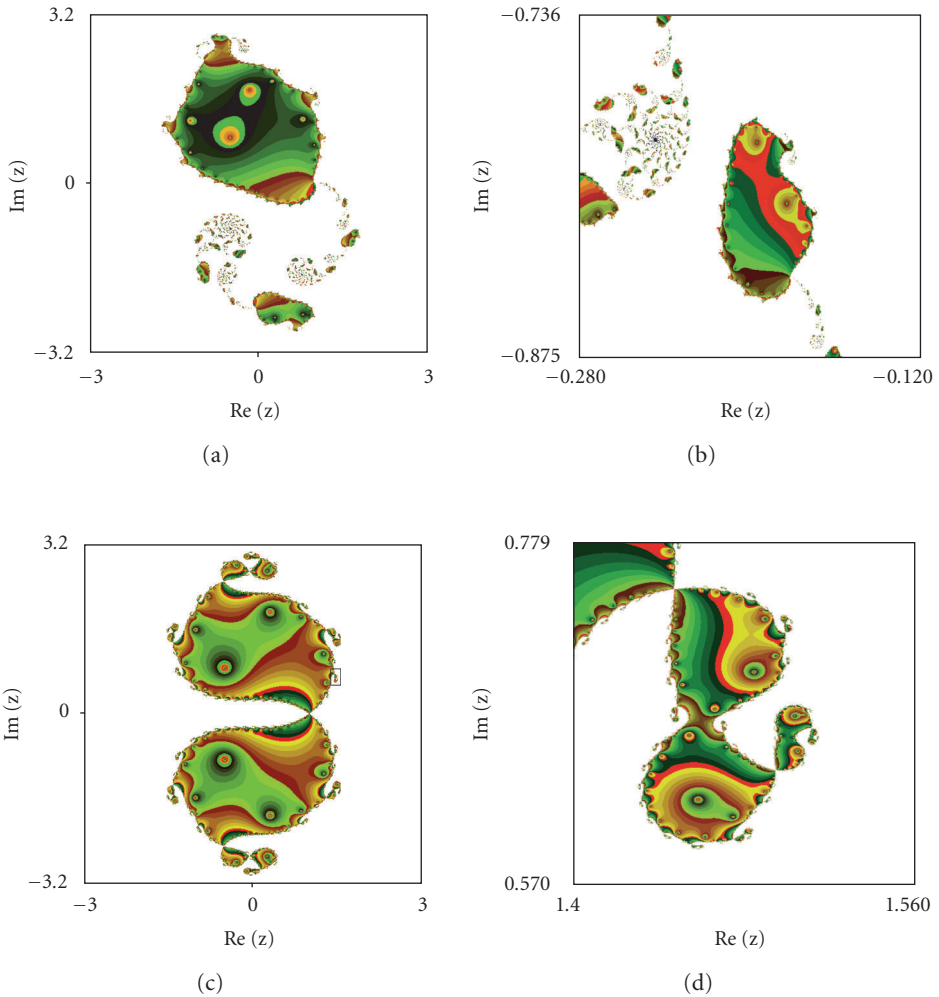
Figure 4.12. Basin of attraction for $z^3 - 1 = 0$ using GCG method: (a) $\lambda = \beta = -0.2 + i \cdot 0.1$, (b) enlargement of the part framed in (a), (c) $\lambda = \beta = -0.2$, (d) enlargement of the part framed in (c).

Appendix

    (i) Numerical tests are done from implementations in FORTRAN90 with graphic mode.

    (ii) For basins of attraction and bifurcations diagrams, we have used a graphic resolution of 600×600 pixels.

    (iii) The maximal iteration number $n_{\max}$ is fixed to 150 iterations for GC and GCG.

    (iv) For line search, the maximal iteration number $n_{\max}^{\mathrm{LS}}$ is fixed to 50 iterations with these values of parameters: $\lambda_0 = 1.0$, $m_1 = 0.3$, $m_2 = 0.7$.

References

[1] M. Al-Baali, *Descent property and global convergence of the Fletcher-Reeves method with inexact line search*, IMA Journal of Numerical Analysis **5** (1985), no. 1, 121–124.

[2] A. Cayley, *Desiderata and suggestions: no. 3. The Newton-Fourier imaginary problem*, American Journal of Mathematics **2** (1879), no. 1, 97.

[3] K. Falconer, *Fractal Geometry*, Mathematical Foundations and Applications, John Wiley & Sons, Chichester, 1990.

[4] R. Fletcher, *A new approach to variable metric algorithms*, The Computer Journal **13** (1970), no. 3, 317–322.

[5] R. Fletcher and M. J. D. Powell, *A rapidly convergent descent method for minimization*, The Computer Journal **6** (1963/1964), 163–168.

[6] G. Julia, *Mémoire sur l'itération des fonctions rationnelles*, Journal de Mathématiques Pures et Appliquées **81** (1918), 47–235.

[7] B. B. Mandelbrot, *The Fractal Geometry of Nature*, W. H. Freeman, New York, 1983.

[8] J. Nocedal, *Theory of algorithms for unconstrained optimization*, Acta Numerica, 1991, Acta Numer., Cambridge University Press, Cambridge, 1992, pp. 199–242.

[9] E. Polak, *Computational Methods in Optimization. A Unified Approach*, Mathematics in Science and Engineering, vol. 77, Academic Press, New York, 1971.

[10] G. Zoutendijk, *Nonlinear programming, computational methods*, Integer and Nonlinear Programming (J. Abadie, ed.), North-Holland, Amsterdam, 1970, pp. 37–86.

Mohamed Lamine Sahari: Département de Mathématiques, Université Badji-Mokhtar, B.P. 12, Annaba 23.000, Algeria
E-mail address: mlsahari@yahoo.fr

Ilhem Djellit: Département de Mathématiques, Université Badji-Mokhtar, B.P. 12, Annaba 23.000, Algeria
E-mail address: i_djellit@hotmail.com