Cyclic and Chaotic Examples in Schwarz-Preconditioned Newton Methods

Conor McCoid and Martin J. Gander

1 Introduction

In [8] we showed the limits of using Newton's method to accelerate domain decomposition methods in 1D. Methods such as ASPIN [1], RASPEN [4], and MSPIN [7] that use additive (A), restricted additive (RA), and multiplicative (M) Schwarz (S) methods, respectively, to precondition (P) exact (E) or inexact (I) Newton's method (N) cannot always be relied upon to converge, though few examples of divergent, cyclic or chaotic behaviour have been presented to the community. The goal of this paper is to expand the space of counterexamples, showing the extent to which chaotic and cycling behaviour can occur in these methods and that these *SP*N methods are not robust to the pitfalls of Newton's method.

We begin by defining the particular method considered, the alternating Schwarz-preconditioned Newton's method (AltSPN) [8] in a general dimension:

$$(1)\ \begin{cases} F(x,u_1,Du_1,D^2u_1)=0,\ x\in\Omega_1,\\ u_1(\partial\Omega)=h,\\ u_1(\Gamma_1)=\gamma_n, \end{cases} \qquad (2)\ \begin{cases} F(x,u_2,Du_2,D^2u_2)=0,\ x\in\Omega_2,\\ u_2(\partial\Omega)=h,\\ u_2(\Gamma_2)=u_1(\Gamma_2), \end{cases}$$

$$(3)\ \begin{cases} J(u_1)\cdot g_1=0,\\ g_1(\partial\Omega)=0,\\ g_1(\Gamma_1)=I, \end{cases} \qquad (4)\ \begin{cases} J(u_2)\cdot g_2=0,\\ g_2(\partial\Omega)=0,\\ g_2(\Gamma_2)=g_1(\Gamma_2), \end{cases}$$

$$(5)\ \gamma_{n+1}=\gamma_n-(G'(\gamma_n)-I)^{-1}(G(\gamma_n)-\gamma_n)=\gamma_n-(g_2(\Gamma_1)-I)^{-1}(u_2(\Gamma_1)-\gamma_n),$$

where Du and D^2u represent the partial derivatives of $u(x)$ of orders 1 and 2, respectively, $J(u)$ is the Jacobian of F evaluated for the function $u(x)$, Γ_1 is the boundary

Conor McCoid
Université Laval, e-mail: conor.mccoid.1@ulaval.ca

Martin J. Gander
Université de Genève, e-mail: martin.gander@unige.ch

of Ω_1 that lies in Ω_2, and Γ_2 the same for Ω_2. Steps (1) and (3) are solved in the first subdomain Ω_1 while steps (2) and (4) are solved in the second, Ω_2.

The operators g_1 and g_2 are the derivatives of u_1 and u_2, respectively, with respect to γ. Since at $\partial\Omega$ u_1 does not depend on γ, the derivative there is zero. Likewise, since $u_1(\Gamma_1) = \gamma$, $g_1(\Gamma_1)$ is the identity operator. In a discrete setting they are represented by matrices.

The function $G(\gamma) := u_2(\Gamma_1)$ is the fixed point iteration of alternating Schwarz. It represents one iteration of alternating Schwarz using γ as the boundary value of $u_1(\Gamma_1)$. Step (5) applies Newton's method to $G(\gamma)$, thereby accelerating the convergence of the fixed point iteration.

Note that AltSPN possesses identical convergence behaviour to RASPEN, and in fact one may consider RASPEN to be the "gluing together" of two instances of AltSPN, one starting from the first subdomain and another starting from the second, see Section 2 of [4]. See also [2] for substructuring of RASPEN akin to that done here. ASPIN then behaves similarly, though without the restriction in the overlap and using inexact Newton in place of exact Newton [6]. Multiplicative Schwarz is identical to alternating Schwarz in the linear case [5], and so these counterexamples are of interest for MSPIN as well.

As explained in [8] Newton's method displays chaotic or cycling behaviour if it displays both oscillatory convergence and oscillatory divergence in a given domain. In 1D this was found to happen when the fixed point iteration to be accelerated ran parallel to the functions

$$g_C(x) = C\,\mathrm{sign}(x-x^*)\sqrt{|x-x^*|}+x,$$

where $g(x^*) = x^*$ and $C \in \mathbb{R}$ [8]. In higher dimensions the class of functions indicating this behaviour greatly increases and becomes difficult to summarize, but notably if one applies the square root operator to each coordinate of $|x-x^*|$ then the higher dimensional equivalent of $g_C(x)$ remains an indicator of this behaviour.

Newton's method in any dimension generally requires globalization techniques, such as line searches or trust regions, to guarantee convergence [3, Ch. 6]. In [8] we developed a Newton-like method in 1D with guaranteed convergence, with assumptions on the function usually satsified by SP*N methods. The counterexamples presented here show a need for a generalization of this method into 2D. However, it is not clear how to perform this generalization. We refer to [8] for further discussion.

2 Cycling counterexamples in 1D

We seek a larger space of examples where cycling occurs for AltSPN. To do so, we employ optimization techniques. First we must find a functional that takes a nonlinearity and returns a measure of the chaos that results from applying AltSPN.

We consider the set of problems

$$u''(x) + f(u(x)) = 0, \quad x \in (-1,1) \tag{1}$$

with homogeneous Dirichlet boundary conditions. The function $f(x)$ satsifies $f(0) = 0$ so that $u(x) = 0$ is the solution to the ODE.

The counterexample presented in [8],

$$u''(x) - \sin(\mu u(x)) = 0, \quad u(\pm 1) = 0,$$

makes use of the antisymmetry in $G(\gamma)$ to achieve its cycles. We wish the same for all nonlinearities $f(x)$ in our space of counterexamples.

Proposition 1 *If $f(x)$ is antisymmetric then $G(\gamma)$ is antisymmetric.*

Proof Suppose $\hat{u}_1(x)$ solves step (1) of alternating Schwarz with $u_1(\beta) = \gamma$. Then

$$-\hat{u}_1'' + f(-\hat{u}_1) = f(\hat{u}_1) - f(\hat{u}_1) = 0.$$

Thus, $-\hat{u}_1(x)$ solves step (1) of alternating Schwarz with $u_1(\beta) = -\gamma$.

By the same logic, if $\hat{u}_2(x)$ solves step (2) of alternating Schwarz with $u_2(\alpha) = u_1(\alpha)$ then $-\hat{u}_2(x)$ solves step (2) with $u_2(\alpha) = -u_1(\alpha)$. Thus, $-\hat{u}_2(\beta) = -G(\gamma) = G(-\gamma)$. □

We therefore restrict our search to nonlinearities which are antisymmetric. It is then sufficient for the Newton-Raphson acceleration, represented by $G_{NR}(\gamma)$, to cross the line $y = -\gamma$ for cycles to exist. In the steps of AltSPN, $G_{NR}(\gamma_n)$ is equal to γ_{n+1} found in step (5).

Given that $f(x)$ is antisymmetric it can be decomposed into a Fourier series consisting solely of sinusoids, which may be truncated for our purposes:

$$f(x) = \sum_{k=1}^{N} c_k \sin(\pi k x). \tag{2}$$

Thus, the functional to optimize takes a set $\{c_k\}_{k=1}^N \in \mathbb{R}^N$, passes it through $f(x) \in C(-1,1)$ and $G(\gamma) \in C(\mathbb{R})$ to arrive at a measure for the chaos of the system. We will represent this functional as $L : \mathbb{R}^N \to \mathbb{R}$. There are many ways to define $L(\{c_k\})$, but we shall use

$$L(\{c_k\}) := \|G_{NR}(\gamma) + \gamma\|. \tag{3}$$

Thus, $L(\{c_k\}) = 0$ if and only if $G_{NR}(\gamma) = -\gamma$, and AltSPN cycles between γ and $-\gamma$ for all values of γ.

It is well-known that there always exists a region around the root of a function where Newton-Raphson converges, assuming continuity of $G''(\gamma)$ and $G'(\gamma) \neq 1$. We expect then to find only local minimizers of $L(\{c_k\})$.

We use a gradient descent with line search to optimize the functional L, with restrictions $\gamma \in [-2,2]$ and $\{c_k\}_{k=1}^5 \in \mathbb{R}^5$, with starting condition $c_1 = 1$, $c_i = 0$ for $i = 2,\ldots,5$. The gradient is computed through a centered finite difference stencil in $\mathbb{R}^5$.

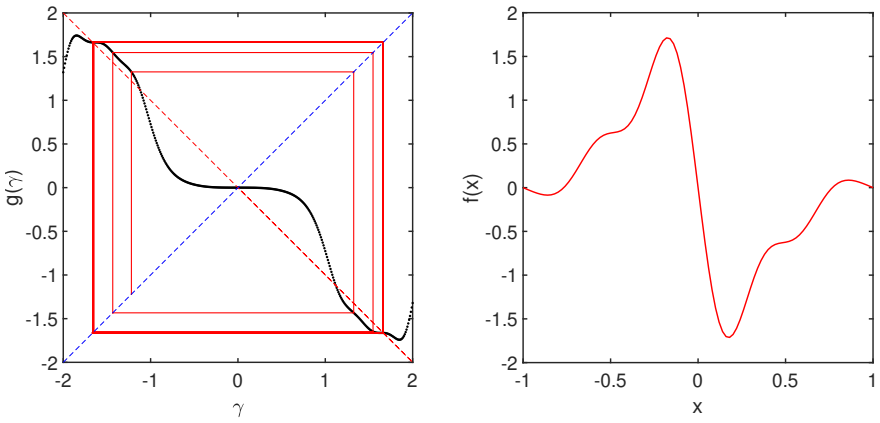

Fig. 1 A counterexample found through optimizing the functional L, equation (3). (Left) AltSPN falls into a stable 2-cycle, as represented by the path of the red lines; the cycle's basin of attraction is most values within $[-2,-1)\cup(1,2]$. (Right) The function $f(x)$ for this counterexample, the sum of five sinusoids.

The coefficients c_k have greater effect with higher k, so that $\partial L/\partial c_5 > \partial L/\partial c_1$. As such, the stencil's step size diminishes with k. The line search takes search direction $\mathbf{p} = \nabla L$ and tests $\{c_k\} + h\mathbf{p}$ for $h = 1$. If $L(\{c_k\} + h\mathbf{p}) > L(\{c_k\})$, then the search is repeated with $h \to 0.5h$, until $L(\{c_k\} + h\mathbf{p}) \leq L(\{c_k\})$.

The domain is discretized with 100 equally spaced points with a 3-point finite difference Laplacian. There is an overlap of 0.4 between the domains, which are symmetric about $x = 0$. This choice of overlap is taken from [8]. We find an exceptional counterexample using this methodology, as presented in Figure 1.

3 Chaotic counterexamples in 2D

We now seek counterexamples in 2D. That is, we seek $f : \mathbb{R} \to \mathbb{R}$ such that using AltSPN to solve

$$\mathscr{L}u(x,y) + f(u(x,y)) = u_{xx}(x,y) + u_{yy}(x,y) + f(u(x,y)) = 0, \quad x,y \in (-1,1)$$

with homogeneous Dirichlet boundary conditions results in cycling behaviour. The two domains are split along the x-axis, so that the first domain is $x \in (-1,\alpha)$, $y \in (-1,1)$ and the second is $x \in (-\alpha,1)$, $y \in (-1,1)$.

The problem with finding cyclic counterexamples in 2D is an issue of dimensionality. If a given direction gives stable cycles, there is no guarantee that every orthogonal direction to these cycles is sufficiently stable to allow these cycles to continue indefinitely. That is, if any point on the cycle is a saddle point, then numerical error can easily eject the iterates from the path of the cycles. The criteria

for numerical cycles is then necessarily stronger than in 1D: Not only must a cycle exist and be stable, there must exist a region around the cycle that is also stable.

As before, $f(x)$ is chosen to be antisymmetric so that $G(\gamma)$ is antisymmetric. Proposition 1 applies to the higher dimensional case by replacing all relevant scalar objects (γ, $G(\gamma)$) with their corresponding vectors (γ, $G(\gamma)$). The decomposition of $f(x)$ into sinusoids and the definition of the functional $L(\{c_k\})$ remain unchanged.

If γ is the discretization of a sinusoid then $G(\gamma) = c\gamma$, where $c \in \mathbb{R}$, up to numerical error. This can be seen by transforming the solution into a Fourier series in the y variable. Suppose

$$\begin{cases} \frac{\partial^2}{\partial x^2} u_1(x,y) + \frac{\partial^2}{\partial y^2} u_1(x,y) + f(u_1(x,y)) = 0, \quad x \in (-1,\alpha), \quad y \in (-1,1), \\ u_1(-1,y) = u_1(x,\pm 1) = 0, \\ u_1(\alpha,y) = C\sin(\pi m y), \end{cases}$$

where $f(x)$ is antisymmetric and can therefore be expressed in the form of equation (2). Use as an ansatz $u_1(x,y) = g(x)\sin(\pi m y)$. Take the Fourier transform of the equation:

$$\begin{aligned} 0 =& \int_{-1}^{1} (g''(x) - m^2\pi^2 g(x))\sin(\pi m y)\sin(\pi k y) + f(g(x)\sin(\pi m y))\sin(\pi k y) dy \\ =& (g''(x) - m^2\pi^2 g(x))\delta_{m,k} C_m + \int_{-1}^{1} \sum_{j=1}^{N} c_j \sin(\pi j g(x)\sin(\pi m y))\sin(\pi k y) dy \\ =& \delta_{m,k}\mathscr{L} g(x) + \sum_{j=1}^{N} c_j \int_{-1}^{1} \sum_{n=0}^{\infty} \frac{(\pi j g(x))^{2n+1}}{(2n+1)!} \sin(\pi m y)^{2n+1} \sin(\pi k y) dy \\ =& \delta_{m,k}\mathscr{L} g(x) + \sum_{j=1}^{N} c_j \sum_{n=0}^{\infty} \frac{(\pi j g(x))^{2n+1}}{(2n+1)!} S(2n+1), \end{aligned}$$

where $S(n)$ is the integral of $\sin(\pi m y)^n \sin(\pi k y)$ over $(-1,1)$. We can find a recursive formula for $S(2n+1)$:

$$\begin{aligned} S(2n+1) =& -\frac{1}{\pi k}\cos(\pi k y)\sin(\pi m y)^{2n+1}\Big|_{-1}^{1} \\ &+ \int_{-1}^{1} \frac{m}{k}(2n+1)\cos(\pi k y)\cos(\pi m y)\sin(\pi m y)^{2n} dy \\ =& (2n+1)\frac{m}{k}\Big(-\frac{1}{\pi k}\cos(\pi m y)\sin(\pi m y)^{2n}\sin(\pi k y)\Big|_{-1}^{1} \\ &+ \frac{m}{k} 2n \int_{-1}^{1} \sin(\pi k y)\left(-\sin(\pi m y)^{2n+1} + \cos(m\pi y)^2 \sin(\pi m y)^{2n-1}\right) dy \Big) \\ =& (2n+1)(2n)\frac{m^2}{k^2}\int_{-1}^{1} \sin(\pi m y)^{2n-1}\sin(\pi k y) - 2\sin(\pi m y)^{2n+1}\sin(\pi k y) dy \end{aligned}$$

$$S(2n+1) = (2n+1)(2n)\frac{m^2}{k^2}\left(S(2n-1) - 2S(2n+1)\right)$$
$$= \frac{(2n+1)(2n)\frac{m^2}{k^2}}{1+2(2n+1)(2n)\frac{m^2}{k^2}} S(2n-1) = C(m,k,n)S(1),$$

where $C(m,k,n)$ is a constant that depends on m, k and n. The value of $S(1)$ is zero unless $k = m$. This proves there exists a function $\tilde{f}_m(x)$ such that

$$\mathscr{L}g(x) + \tilde{f}_m(g(x)) = 0.$$

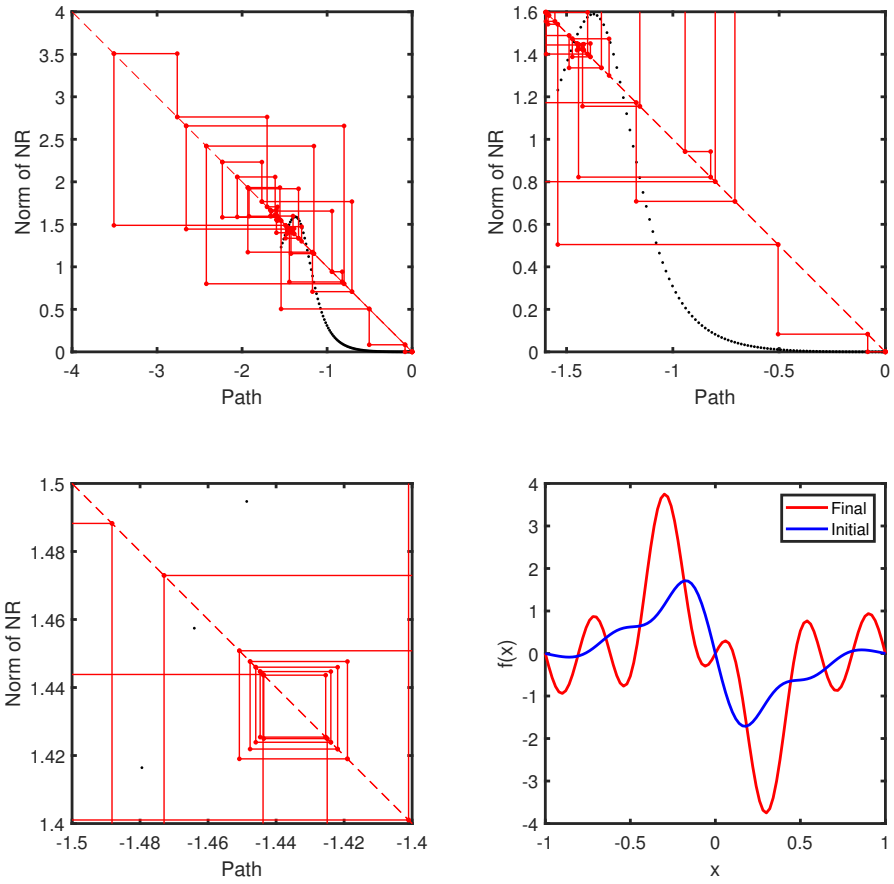

Fig. 2 A chaotic AltSPN sequence in 2D, at three resolutions (top row, bottom left). The solid red lines indicate the path of the fixed point iteration $G_{NR}(\gamma)$. The nonlinearity (bottom right) is found through optimization on the functional L, equation (3). The AltSPN sequence is seeded using a sine wave $\sin(\pi y)$ as boundary condition on the first subdomain, but quickly diverges from this pathway.

Thus, if $g(x)$ satisfies this ODE with boundary conditions $g(-1) = 0$ and $g(\alpha) = C$ then $u_1(x,y)$ solves the PDE. Since $u_1(x,y)$ in this form satisfies the boundary conditions it is the solution to this step of alternating Schwarz. The solution on the second domain, $u_2(x,y)$, then also has the same form by a symmetry argument, and its value at $x = -\alpha$ is a sinusoid of the same period. Therefore, $G(\gamma) = c\gamma$.

As a direct consequence of this, starting with any single sine wave as boundary conditions provides a single parameter pathway for the function $G : \mathbb{R}^N \to \mathbb{R}^N$. We take advantage of this fact and use $c\sin(\pi y)$ as the boundary condition for the first subdomain of AltSPN, varying c between -0.5 and 0. To seed the optimization we use the previously obtained counterexample nonlinearity $f(x)$ from the 1D case. The same optimization method is used. 20 equally spaced points are used in each direction with a 5-point stencil Laplacian.

The resulting nonlinearity does not admit a stable cycle but highlights the chaos that can result from using AltSPN. Figure 2 gives this nonlinearity and an example of a sequence generated by AltSPN. The sequence begins on the sine wave path described previously. It then approaches a nearby stable cycle, having departed from the strict sine wave path to one that closely resembles sine waves (bottom left of the figure). However, the cycle is not numerically stable and is ultimately abandoned.

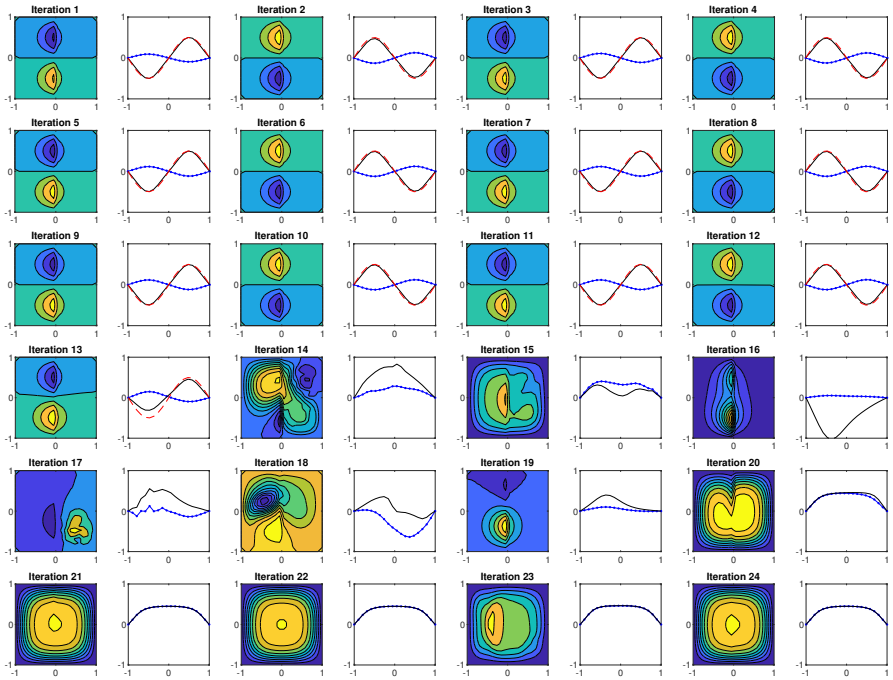

Fig. 3 The first 24 iterations of the chaotic AltSPN sequence. The left figure of each iteration shows the overall solution at that step combined from the two subdomains, and the right of each shows the resulting $G(\gamma)$ in blue and the AltSPN result in black. For comparison, the sine wave $0.5\sin(\pi y)$ is plotted in red for the cycling regime, with sign that alternates with each iteration.

As it leaves this pathway it descends into a choatic regime (top left). It eventually ejects onto a convergent pathway (top right). It is possible these cycles exist on saddlepoints, stable along some pathways but unstable along others. A small numerical error will then shunt the AltSPN sequence away from the cycle, either to a divergent (top left) or convergent (top right) pathway.

Figure 3 provides snapshots of the solution at each of the iterates. The first 24 iterations of AltSPN are shown, giving the nearly cycling regime (iterations 1 to 12) and part of the chaotic regime (13 to 24). The cycling regime is very nearly a cycle of sine waves, as seen by the comparable sine waves in red. A small change from these sine waves in iteration 13 causes chaos to take over until relative stability in iterations 20 through 24. From there, AltSPN eventually converges.

These results show that acceleration cannot be used without consequences for all Schwarz algorithms. As with standard Newton-Raphson, there exist problems for which the sequence diverges, cycles or behaves chaotically.

References

1. Cai, X.-C. and Keyes, D. E. Nonlinearly preconditioned inexact Newton algorithms. *SIAM Journal on Scientific Computing* **24**(1), 183–200 (2002).
2. Chaouqui, F., Gander, M. J., Kumbhar, P. M., and Vanzan, T. Linear and nonlinear substructured restricted additive schwarz iterations and preconditioning. *Numerical Algorithms* **91**(1), 81–107 (2022).
3. Dennis Jr, J. E. and Schnabel, R. B. *Numerical methods for unconstrained optimization and nonlinear equations*. SIAM (1996).
4. Dolean, V., Gander, M. J., Kheriji, W., Kwok, F., and Masson, R. Nonlinear preconditioning: How to use a nonlinear Schwarz method to precondition Newton's method. *SIAM Journal on Scientific Computing* **38**, 3357–3380 (2016).
5. Gander, M. J. Schwarz methods over the course of time. *Electronic transactions on numerical analysis* **31**, 228–255 (2008).
6. Gander, M. J. On the origins of linear and non-linear preconditioning. In: *Domain Decomposition Methods in Science and Engineering XXIII*, 153–161. Springer (2017).
7. Liu, L. and Keyes, D. E. Field-split preconditioned inexact Newton algorithms. *SIAM Journal on Scientific Computing* **37**, A1388–A1409 (2015).
8. McCoid, C. and Gander, M. J. Cycles in Newton-Raphson preconditioned by Schwarz (ASPIN and its cousins). In: *Domain Decomposition Methods in Science and Engineering XXVI*, *Lecture Notes in Computational Science and Engineering*, vol. 145. Springer Cham (2023).