Chapter 1
Cycles in Newton-Raphson preconditioned by Schwarz (ASPIN and its cousins)

Conor McCoid and Martin J. Gander

1.1 Introduction

ASPIN [3], RASPEN [5], and MSPIN [8] rely on various Schwarz methods to precondition either Newton-Raphson or inexact Newton. While *a priori* convergence criteria have been found for the underlying Schwarz methods, so far none exist for their combination with Newton-Raphson.

Like in the linear case when combining a Krylov method and a Schwarz method, there is an equivalence between preconditioning Newton-Raphson with a Schwarz method and accelerating that same Schwarz method with Newton-Raphson [6]: A domain is first subdivided into subdomains, the problem solved on each subdomain, and the resulting formulation iterated through Krylov and Newton-Raphson, respectively.

We examine cycling behaviour in alternating Schwarz in one dimension that has been accelerated by applying Newton-Raphson. We begin by presenting the algorithm for alternating Schwarz and how it is accelerated by Newton-Raphson. Suppose we seek to solve the boundary value problem

$$F(x,u,u',u'') = 0, \quad x \in [a,b], \quad u(a) = A, \quad u(b) = B$$

for some function $F(x,u,v,w)$. Then an iteration of alternating Schwarz with subdomains (a,β) and (α,b), $\alpha < \beta$, is comprised of the following three steps:

$$
\begin{aligned}
(1) \quad & F(x,u_1,u_1',u_1'') = 0, & u_1(a) = A, & \quad u_1(\beta) = \gamma_n, \\
(2) \quad & F(x,u_2,u_2',u_2'') = 0, & u_2(\alpha) = u_1(\alpha), & \quad u_2(b) = B, \\
(3) \quad & \gamma_{n+1} = u_2(\beta) = G(\gamma_n). &
\end{aligned}
$$

Conor McCoid
University of Geneva, e-mail: conor.mccoid@unige.ch

Martin J. Gander
University of Geneva, e-mail: martin.gander@unige.ch

The function $G(\gamma)$ thus represents one iteration of alternating Schwarz in substructured form. The process is repeated until convergence, ie.

$$(G \circ G \circ \cdots \circ G)(\gamma) = G^n(\gamma) \approx G^{n+1}(\gamma) = (G \circ G^n)(\gamma).$$

This is naturally a fixed point iteration applied to the function $G(\gamma)$.

To accelerate the method one applies Newton-Raphson to the function $f(\gamma) = G(\gamma) - \gamma$, which has a root at the fixed point. If the fixed point is unique, this is the only root of $f(\gamma)$. To apply Newton-Raphson, one needs to know the value of $G'(\gamma)$, which may be found by adding two new steps, (1') and (2'), to alternating Schwarz:

$$
\begin{array}{llll}
(1) & F(x, u_1, u_1', u_1'') = 0, & u_1(a) = A, & u_1(\beta) = \gamma_n, \\
(1') & J(u_1) \cdot (v_1, v_1', v_1'') = 0, & v_1(a) = 0, & v_1(\beta) = 1, \\
(2) & F(x, u_2, u_2', u_2'') = 0, & u_2(\alpha) = u_1(\alpha), & u_2(b) = B, \\
(2') & J(u_2) \cdot (v_2, v_2', v_2'') = 0, & v_2(\alpha) = 1, & v_2(b) = 0, \\
(3) & \gamma_{n+1} = \gamma_n - \dfrac{u_2(\beta) - \gamma_n}{v_1(\alpha)v_2(\beta) - 1} = \gamma_n - \dfrac{G(\gamma_n) - \gamma_n}{G'(\gamma_n) - 1},
\end{array}
$$

where $v_i(x) = \partial u_i(x)/\partial \gamma$ and $J(u_i)$ is the Jacobian of $F(x, u_i, u_i', u_i'')$.

1.2 Convergence of generic fixed point iterations and Newton-Raphson

A generic fixed point iteration $x_{n+1} = g(x_n)$ converges when $|g(x_n) - x^*| < |x_n - x^*|$, where x^* is the fixed point. This occurs when $g(x)$ lies between x and $2x^* - x$. The convergence or divergence of the fixed point iteration is monotonic if $\text{sign}(g(x) - x^*) = \text{sign}(x - x^*)$ and oscillatory otherwise. This creates four lines, $y = x$, $y = 2x^* - x$, $y = x^*$ and $x = x^*$, that divide the plane into octants. The four pairs of opposite octants form four regions with distinct behaviour of the fixed point iteration, see left of Figure 1.1 or Figure 5.7 from [7]:

1, $g(x) < x < x^*$ or $g(x) > x > x^*$:   monotonic divergence;
2, $x < g(x) < x^*$ or $x > g(x) > x^*$:   monotonic convergence;
3, $x < x^* < g(x) < 2x^* - x$ or $x > x^* > g(x) > 2x^* - x$:   convergent oscillations;
4, $x < x^* < 2x^* - x < g(x)$ or $x > x^* > 2x^* - x > g(x)$:   divergent oscillations.

If the function $g(x)$ intersects the line $y = x$ at a point other than x^* then there are additional fixed points that the method can converge towards. If it intersects the line $y = 2x^* - x$ then a stable cycle can form. A fixed point iteration is therefore only guaranteed to converge if $g(x)$ lies entirely between the lines $y = x$ and $y = 2x^* - x$, ie. within regions 2 and 3.

Newton-Raphson can make use of this analysis by considering it as a fixed point iteration:
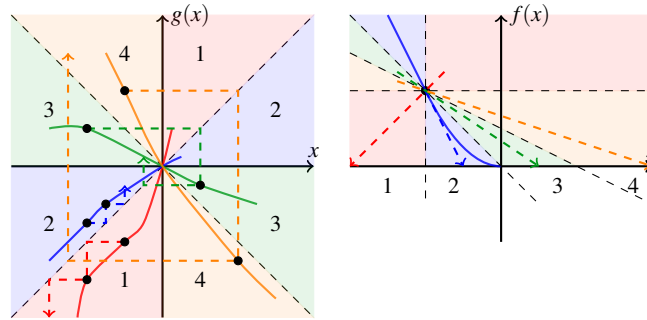
Fig. 1.1 **Left:** Behaviour of the fixed point iteration $x_{n+1} = g(x_n)$, where the origin is the fixed point, $g(0) = 0$. **Right:** Regions of Newton-Raphson, $x_{n+1} = x_n - f(x_n)/f'(x_n)$, where the origin is the root, $f(0) = 0$. The tangent line to $f(x)$ can be traced from $(x, f(x))$ towards the line $y = 0$. Where it lands on this line indicates which fixed point iteration behaviour occurs.

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = g_f(x_n).$$

The borders between the regions no longer depend solely on the value of $f(x)$ but also $f'(x)$. The right of Figure 1.1 shows which type of behaviour Newton-Raphson will have based on where the tangent line points.

As stated, if $g_f(x)$ intersects the line $y = x$ there are additional fixed points, and if it intersects $y = 2x^* - x$ there may be stable cycles. For guaranteed convergence $g_f(x)$ must lie between these lines. Intersections of $g_f(x)$ with $y = x$ occur only if $f(x) = 0$ and $f(x)$ has additional roots or $f'(x) = \infty$. Both circumstances are assumed not to occur. Intersections of $g_f(x)$ with $y = 2x^* - x$ may be represented as a first order ODE:

$$f'_C(x) = -\frac{f_C(x)}{2(x^* - x)}, \quad f_C(x^*) = 0.$$

The solution to this ODE is $f_C(x) = C\sqrt{|x - x^*|}$ where $C \in \mathbb{R}$. If a function $f(x)$ with root x^* is tangential to $f_C(x)$ for any value of C then $g_f(x)$ intersects the line $y = 2x^* - x$. The left of Figure 1.2 shows the functions $f_C(x)$.

A function $f(x)$ that is monotonic with respect to this geometry has guaranteed convergence under Newton-Raphson. That is, if $f(x)$ is nowhere tangential to $f_C(x)$ in a given domain containing x^* for any value of C then $g_f(x)$ converges to the root for any initial guess in that domain. Since $f'_C(x^*) = \infty$ and $f(x^*) = 0$ there is always a region around the root x^* where $f(x)$ crosses all of these lines monotonically. This conforms with the theory on Newton-Raphson.

The corresponding geometry for a fixed point function accelerated by Newton-Raphson is skewed such that the line $y = 0$ is aligned to $y = x$, as seen in the right of Figure 1.2. The lines of this figure are the functions $g_C(x) = f_C(x) + x$. A function $g(x)$ must be monotonic in this geometry or Newton-Raphson applied to $g(x) - x$ may exhibit cycling behaviour.
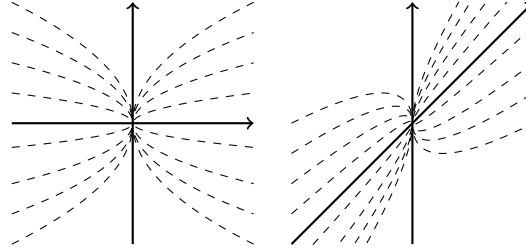
Fig. 1.2 Left: Solutions $f_C(x)$ such that $g_f(x)$ intersects $y = 2x^* - x$ for all x. **Right:** Functions $g_C(x) = f_C(x) + x$ such that $g_f(x)$ for $f(x) = g_C(x) - x$ intersects $y = 2x^* - x$ for all x.

Table 1.1 Conditions for convergent behaviour of Newton-Raphson applied to $g(x) - x$.

$g(x)$ lies in	Necessary condition	Sufficient condition
1	$g'(x) > 1$	
2	$g'(x) < 1$	$g'(x) < 1/2$
3	$g'(x) < 1/2$	$g'(x) < 0$
4	$g'(x) < 0$	

If it is known in which fixed point region of the left of Figure 1.1 $g(x)$ lies then one can find necessary and, in some cases, sufficient conditions for Newton-Raphson to have convergent behaviour based on the slopes of the lines $g_C(x)$. For example, in region 2 the maximum of $g'_C(x)$ is 1. If $g(x)$ lies in region 2 then its slope must therefore be less than 1 everywhere or there will be a point where $g(x)$ runs tangent to $g_C(x)$ for some C. Moreover, the minimum of $g'_C(x)$ is 1/2. If $g(x)$ has a slope less than 1/2 then it cannot run tangent to $g_C(x)$ for any C. The list of these conditions is summarized in Table 1.1.

1.3 The fixed point iteration of alternating Schwarz

We now seek to apply this theory to alternating Schwarz. As stated earlier, we consider alternating Schwarz as a function $G(\gamma)$, taking as input the value of $u_1(\beta)$ and as output the value of $u_2(\beta)$. Under reasonable conditions we can prove a number of useful properties of $G(\gamma)$ without prior knowledge of the fixed point γ^*.

Theorem 1 *If the problem $F(x, u, u', u'') = 0$ for $x \in \Omega$, $u(x) = h(x)$ for $x \in \partial\Omega$ has a unique solution on $\Omega = [a, \alpha]$ and $\Omega = [\beta, b]$ and the continuations of these solutions are also unique, then the function $G(\gamma)$ is strictly monotonic.*

Proof It suffices to show that $G(\gamma_1) = G(\gamma_2)$ implies $\gamma_1 = \gamma_2$. Let u_1^j solve the problem on $[a, \beta]$ with $u_1^j(\beta) = \gamma_j$. Likewise, u_2^j solves the problem on $[\alpha, b]$ with $u_2^j(\alpha) = u_1^j(\alpha)$. Suppose $u_2^1(\beta) = u_2^2(\beta)$. Then both u_2^1 and u_2^2 solve the same problem on $[\beta, b]$. By assumption, this must mean $u_2^1 = u_2^2$ and $u_1^1(\alpha) = u_1^2(\alpha)$. By a

similar argument, this implies u_1^1 and u_1^2 solve the same problem on $[a, \alpha]$. Again by assumption $u_1^1 = u_1^2$ and $\gamma_1 = \gamma_2$.   □

We can even prove that $G(\gamma)$ is restricted to region 2 with additional properties. As an example, we reprove a result from Lui [9].

Theorem 2 (Theorem 2 from [9])

Consider the equation $u''(x) + f(x, u, u') = 0$ *for* $x \in (a, b)$, $u(a) = u(b) = 0$ *under the assumptions that*

- $f \in C^1\left([a, b] \times \mathbb{R} \times \mathbb{R}\right)$,
- $\frac{\partial f(x, v, v')}{\partial u} \leq 0$ *for all* $x \in [a, b]$ *and* $v \in H_0^1([a, b])$,
- $|f(x, v, v')| \leq C(1 + |v'|^\eta)$ *for all* $x \in [a, b]$ *and* $v \in H_0^1([a, b])$ *and some* $C > 0$, $0 < \eta < 1$.

The problem is solved using alternating Schwarz with two subdomains and Dirichlet transmission conditions. Then $G(\gamma)$ *for this problem lies within region 2.*

Proof It suffices to prove that the problem is well posed and $0 < G'(\gamma) < 1$ for all $\gamma \in \mathbb{R}$. The well-posedness of the problem is guaranteed by Proposition 2 from [9]. As Lui points out, this also means the problem is well posed on any subdomain. Using Theorem 1 this gives monotonicity of $G(\gamma)$. Moreover, if $u(x) = 0$ for any $x \in (a, b)$ then the problem would be well posed on the domains $[a, x]$ and $[x, b]$. As such, $u(x)$ has the same sign as γ and $G'(\gamma) > 0$.

Consider the problem in g_1:

$$g_1''(x) + \frac{\partial f}{\partial u} g_1 + \frac{\partial f}{\partial u'} g_1' = 0, \quad x \in [a, \beta], \quad g_1(a) = 0, \quad g_1(\beta) = 1.$$

From the second assumption on f the operator on g_1 satisfies a maximum principle (see, for example, [9]). Therefore, $g_1(x) < 1$ for all $x \in (a, \beta)$. By the same reasoning, $g_2(x) < g_1(\alpha) < 1$ for all $x \in (\alpha, b)$ and $G'(\gamma) < 1$. Incidentally, the same maximum principle applies for the operator on $-g_1$ and $-g_2$, and so $G'(\gamma) > 0$ as we had before.   □

This provides guaranteed convergence of alternating Schwarz. However, it does not guarantee the convergence when one accelerates it through Newton-Raphson. Using Table 1.1 we know that such convergence is assured if $G'(\gamma) < 1/2$ for all γ, but this is not true in all cases and cannot be determined *a priori*.

Take as an example the following second order nonlinear differential equation

$$u''(x) - \sin(au(x)) = 0, \quad x \in (-1, 1), \tag{1.1}$$

with homogeneous Dirichlet boundary conditions. The problem is well posed and admits only the trivial solution $u(x) = 0$. It is easy to see that this equation satisfies the conditions of Theorem 2. Therefore, the alternating Schwarz fixed point iteration, $G(\gamma)$, lies within region 2 and is guaranteed to converge to the fixed point. Sadly, its Newton-Raphson acceleration will not do so for all initial conditions. Take
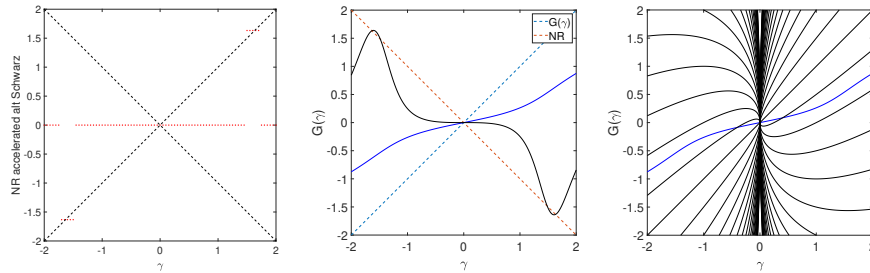
Fig. 1.3 Left: Results of Newton-Raphson accelerated alternating Schwarz as a function of initial condition in solving equation (1.1). The value of a is 3.6 and the subdomains are $\Omega_1 = (-1, 0.2)$ and $\Omega_2 = (-0.2, 1)$. **Middle:** $G(\gamma)$ and its Newton-Raphson acceleration. **Right:** $G(\gamma)$ plotted with the geometry of the right of Figure 1.2.

$a = 3.6$ with an overlap of 0.4 and symmetric regions. The results of the Newton-Raphson acceleration are found in Figure 1.3 (left). While for most initial values of γ the method converges to the correct solution $u = 0$ there are two small intervals where the method enters a stable cycle.

The function $G(\gamma)$ can be plotted numerically, along with its Newton-Raphson acceleration, see Figure 1.3 (middle), which shows that $G(\gamma)$ does indeed lie within region 2 as predicted by Theorem 2. However, $G(\gamma)$ runs tangential to one of the lines $g_C(\gamma)$, see Figure 1.3 (right), and so its Newton-Raphson acceleration crosses into region 4. Due to symmetry, there is a 2-cycle at each crossing. Depending on the slope of the acceleration as it crosses into region 4 this cycle may be stable.

Where stable cycles exist so too must there be period doubling bifurcation. Changing the value of the parameter a we find that the 2-cycle found in Figure 1.3 (left) becomes two 2-cycles, then two 4-cycles, and so on until it devolves into chaos, see Figure 1.4. With enough chaos the cycles are no longer stable and the
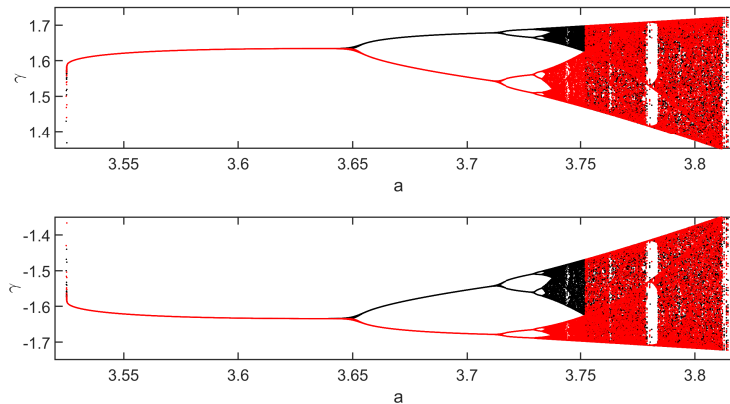

Fig. 1.4 Period doubling bifurcation in the example caused by Newton-Raphson acceleration.
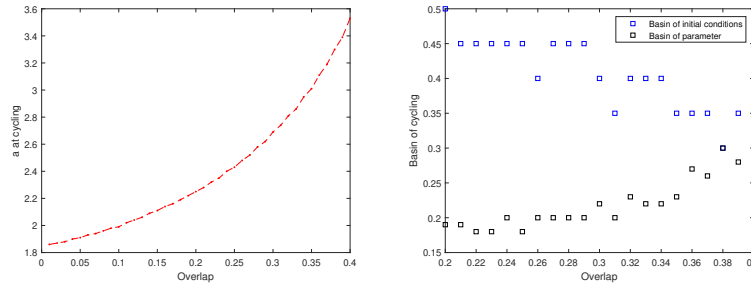
Fig. 1.5 Left: value of *a* at which bifurcation starts. **Right:** width of basin of cycling in γ and *a*.

acceleration exits into a convergent region.

While a change in the parameter *a* is the most obvious way to alter the dynamics, one can also change the size of the overlap. This has a direct effect on the basin of cycling in the spaces of both initial condition γ and the parameter *a*. Figure 1.5 (left) shows a nonlinear relationship between the first value of *a* at which cycling is observed and the size of the overlap. As the overlap grows the parameter *a* must be larger and larger for cycling to occur. Figure 1.5 (right) indicates that the interval of initial conditions that result in cycling shrinks as the overlap grows. Meanwhile, the length of the bifurcation diagram increases, meaning there are more values of *a* with stable cycling.

1.4 Accelerated alternating Schwarz with guaranteed convergence

Given Theorem 2 and the conditions of Table 1.1 one can construct a series of tests to see if the Newton-Raphson acceleration is suitable for a given iteration. We present one further useful trick to strengthen convergence, a correction to Newton-Raphson due to Davidenko and Branin [1, 2, 4]. We replace step (3) in the algorithm with

$$(3*) \quad \tilde{\gamma}_n = \gamma_n - \frac{G(\gamma_n) - \gamma_n}{|G'(\gamma_n) - 1|}.$$

For $G(\gamma)$ within region 2 the Newton-Raphson acceleration will now always march in the direction of the fixed point. It may still overshoot and cycle but the direction will always be correct.

For a problem satisfying the conditions of Theorem 2 or similar that guarantees that $G(\gamma)$ lies in region 2 the algorithm proceeds as follows:

1. Select some $\gamma_0 \in \mathbb{R}$. Set $n = 0$.
2. Calculate $G(\gamma_n)$ and $G'(\gamma_n)$. If $G'(\gamma_n) = 1$ then set $\gamma_{n+1} = G(\gamma_n)$, increment *n* and return to step 2. If this is not true, proceed to step 3.

3. Perform step (3*), which is the Newton-Raphson acceleration using the Davidenko-Branin trick. If $|G'(\gamma_n) - 1| \geq 1/2$ then set $\gamma_{n+1} = \tilde{\gamma}_n$, increment n and return to step 2. If this is not true, calculate $\hat{\gamma}_n$, the average of γ_n and $\tilde{\gamma}_n$, and proceed to step 4.
4. Calculate $G(\hat{\gamma}_n)$. If $G(\hat{\gamma}_n) - \hat{\gamma}_n$ has the same sign as $G(\gamma_n) - \gamma_n$ then set $\gamma_{n+1} = \tilde{\gamma}_n$, increment n and return to step 2. If this is not true, set $\gamma_{n+1} = G(\gamma_n)$, increment n and return to step 2.

Each of steps 2, 3 and 4 contain a test of whether Newton-Raphson will converge. In step 2, Newton-Raphson will not converge if the derivative of $G(\gamma) - 1$ is zero. In step 3, convergence is guaranteed if $G'(\gamma) \leq 1/2$ based on Table 1.1. The Davidenko-Branin trick strengthens this and also guarantees convergence if $G'(\gamma) \geq 3/2$.

In step 4 we test the point halfway between the starting value γ_n and the Newton-Raphson acceleration $\tilde{\gamma}_n$, denoted $\hat{\gamma}_n$. Since $G(\gamma)$ is in region 2 if $G(\gamma) > \gamma$ then $\gamma < \gamma^*$ and vice versa. Therefore, we can easily determine whether $\hat{\gamma}_n$ is on the same side of the fixed point as γ_n. If it is, then the fixed point γ^* lies on the same side of $\hat{\gamma}_n$ as $\tilde{\gamma}_n$, and so $\tilde{\gamma}_n$ is closer to γ^* than γ_n. If it is not, then γ^* lies between γ_n and $\hat{\gamma}_n$. Since $\tilde{\gamma}_n$ is on the other side of $\hat{\gamma}_n$ it is further from γ^* than γ_n and we have divergence. In such a case, the fixed point iteration should be used.

Note that while $G(\gamma)$ represents alternating Schwarz in this context, it may be exchanged for any fixed point iteration, in particular any Schwarz method. All that is required for the algorithm to function is for $G(\gamma)$ to be within region 2. For Schwarz methods, this would necessitate a theorem similar to Theorem 2.

References

1. Franklin H Branin. Widely convergent method for finding multiple solutions of simultaneous nonlinear equations. *IBM Journal of Research and Development*, 16(5):504–522, 1972.
2. RP Brent. On the Davidenko-Branin method for solving simultaneous nonlinear equations. *IBM Journal of Research and Development*, 16(4):434–436, 1972.
3. Xiao-Chuan Cai and David E Keyes. Nonlinearly preconditioned inexact Newton algorithms. *SIAM J Sci Comput*, 24(1):183–200, 2002.
4. DF Davidenko. On a new method of numerical solution of systems of nonlinear equations. In *Dokl. Akad. Nauk SSSR*, volume 88, pages 601–602, 1953.
5. V Dolean, M J Gander, W Kheriji, F Kwok, and R Masson. Nonlinear preconditioning: How to use a nonlinear Schwarz method to precondition Newton's method. *SIAM J Sci Comput*, 38:3357–3380, 2016.
6. Martin J Gander. On the origins of linear and non-linear preconditioning. In *Domain Decomposition Methods in Science and Engineering XXIII*, pages 153–161. Springer, 2017.
7. Walter Gander, Martin J Gander, and Felix Kwok. *Scientific computing-An introduction using Maple and MATLAB*, volume 11. Springer Science & Business, 2014.
8. Lulu Liu and David E Keyes. Field-split preconditioned inexact Newton algorithms. *SIAM J Sci Comput*, 37:A1388–A1409, 2015.
9. S. Lui. On Schwarz alternating methods for nonlinear elliptic PDEs. *SIAM J Sci Comput*, 21(4):1506–1523, 1999.