Floating-point numbers in Ginger

November 6, 2012

In this section I shall describe a new way to represent floating point numbers in Ginger. We consider $m \times m$ matrix multiplication and require the input entries in the set T=$\{a/b :|\ a\ | \leq 2^{N_a}, b \in \{1, 2, 2^2, 2^3, ..., 2^{N_b}\}\}$. Previously, it was shown that $p > (m+1)^2 \cdot 2^{4(N_a+N_b)}$ is necessary for making θ 1-1 which is required to make the mapping isomorphic from U to $\mathbb{Q}/p$. In this exercise I show that by modifying the definition of θ and the mapped field, one can obtain a better bound on p (i.e., $p > \max\{2m \cdot 2^{2N_a+2N_b}, 2N_a + 4N_b + \log_2 m\}$). I do not make any changes to step $C1$. The changes I propose, with the resulting proofs are described below:
Define θ as follows:

$$\theta : U \to \mathbb{F}$$
$$\frac{a}{b = 2^k} \mapsto (a \ mod \ p, k \ mod \ p)$$

The field $\mathbb{F}$ is the set of equivalence classes on the set $\mathbb{Z}/p \times \mathbb{Z}/p$ under the equivalence relation: $(a,b) \backsim (c,d)$ if $\ell \equiv s \pmod p$ and $r_1 + d \equiv r_2 + b \pmod p$, where $a = \ell \cdot 2^{r_1}$ and $c = s \cdot 2^{r_2}$. We have written a and c in this form by factoring out all the powers of 2. Every integer can be written in this form (for integer greater than 1 this follows from fundamental theorem of arithmetic, and $1 = 1 \cdot 2^0$ and $0 = 0 \cdot 2^k$ where k is arbitrary).
The addition operation is defined by $(a, b) + (c, d) = (a \cdot 2^d + c \cdot 2^b, b + d)$
The multiplication operation is defined by $(a, b) \cdot (c, d) = (a \cdot c, b + d)$
$(1, 0)$ is the multiplicative identity. For any $(a, b) \in \mathbb{F}$ (except additive identity of course) $(a^{-1}, -b)$ is the multiplicative inverse. $(0, 0)$ is the additive identity (in fact $(0, 0) \backsim (0, k)$). For any $(a, b) \in \mathbb{F}$, $(-a, b)$ is the additive inverse. Now we show that θ preserves addition and multiplication rules for $q_1, q_2 \in U$.

$$\theta(q_1 + q_2) = \theta(\frac{a_1}{b_1} + \frac{a_2}{b_2}) = \theta(\frac{a_1 \cdot b_2 + a_2 \cdot b_1}{2^{k_1} \cdot 2^{k_2}}) = (a_1 \cdot b_2 + a_2 \cdot b_1, k_1 + k_2)$$
$$\theta(q_1) + \theta(q_2) = (a_1, k_1) + (a_2, k_2) = (a_1 \cdot 2^{k_2} + a_2 \cdot 2^{k_1}, k_1 + k_2)$$
$$so, \theta(q_1 + q_2) = \theta(q_1) + \theta(q_2)$$

Similarly, the multiplication rule holds:

$$\theta(q_1 \cdot q_2) = \theta(\frac{a_1 \cdot a_2}{b_1 \cdot b_2}) = \theta(\frac{a_1 \cdot a_2}{2^{k_1} \cdot 2^{k_2}}) = (a_1 \cdot a_2, k_1 + k_2)$$
$$\theta(q_1) \cdot \theta(q_2) = (a_1, k_1) \cdot (a_2, k_2) = (a_1 \cdot a_2, k_1 + k_2)$$
$$so, \theta(q_1 \cdot q_2) = \theta(q_1) \cdot \theta(q_2)$$

Claim: θ is a function from U to $\mathbb{F}$

Proof: We need to show that if $q_1 = q_2$ then $\theta(q_1) \equiv \theta(q_2)$

$$q_1 = q_2$$
$$\frac{a_1}{b_1} = \frac{a_2}{b_2}$$
$$a_1 \cdot b_2 = a_2 \cdot b_1$$
$$\ell \cdot 2^{r_1} \cdot 2^{k_2} = s \cdot 2^{r_2} \cdot 2^{k_1}$$

because we can write: $a_1 = \ell \cdot 2^{r_1}$ and $a_2 = s \cdot 2^{r_2}$.

$$\ell \cdot 2^{r_1+k_2} = s \cdot 2^{r_2+k_1}$$

so this implies $\ell = s$ and $r_1 + k_2 = r_2 + k_1$. To see why this is true first assume both sides are greater than one, so by fundamental theorem of arithmetic we can write them as a product of distinct primes. Now ℓ and s are numbers such that they contain all the other primes except 2.

$$\ell \cdot 2^{r_1+k_2} = s \cdot 2^{r_2+k_1}$$

means that on both sides the primes should be the same, and they should have the same powers. Since ℓ and s do not contain the prime 2, so therefore $r_1 + k_2 = r_2 + k_1$ as it's the power of 2. $\ell = s$ as it contains all the other primes except 2. Also knowing that the powers of 2 on both sides are equal trivially implies $\ell = s$. Now if both sides are equal to 1, then $\ell = s = 1$ and $r_1 + k_2 = r_2 + k_1 = 0$. If both sides are zero, this means a_1 and a_2 were both zero, so this implies $\ell = s = 0$ and $r_1 + k_2 = r_2 + k_1$ since for any given k_1 and k_2 we could choose arbitrary r_1 and r_2 (as we can write $0 = 0 \cdot 2^r$ where r is arbitrary). $\ell = s$ and $r_1 + k_2 = r_2 + k_1$ naturally means $\ell = s \pmod p$ and $r_1 + k_2 = r_2 + k_1 \pmod p$ which implies $\theta(q_1) \equiv \theta(q_2)$ (see definition of the new equivalence relation above) .

Claim: If $p > \max\{2m \cdot 2^{2N_a+2N_b}, 2N_a + 4N_b + \log_2 m\}$ then θ is 1-1 function.

Proof: We need to prove that if $\theta(q_1) \equiv \theta(q_2)$ then $q_1 = q_2$. Suppose for the sake of contradiction that:

$$q_1 \neq q_2$$
$$\frac{a_1}{b_1} \neq \frac{a_2}{b_2}$$
$$a_1 \cdot b_2 \neq a_2 \cdot b_1$$
$$\ell \cdot 2^{r_1+k_2} \neq s \cdot 2^{r_2+k_1}$$

then either $\ell \neq s$ or $r_1 + k_2 \neq r_2 + k_1$

Suppose $\ell \neq s$. $\theta(q_1) \equiv \theta(q_2)$ means that $\ell = s \pmod p$ and $r_1 + k_2 = k_1 + r_2 \pmod p$. Now $\ell \neq s$ implies that $\ell - s = hp$ where h is an integer other than zero (as $\ell = s \pmod p$). So, it follows:

$$| \ell - s | \geq p$$
$$| \ell | + | s | \geq p$$

$| \ell |\leq | \ell \cdot 2^{r_1} |\leq | a_1 | \leq m \cdot 2^{2N_a+2N_b}$ where the last inequality uses the bound on the numerator from Claim $B.1$. Similarly, $| s |\leq | s \cdot 2^{r_2} |\leq | a_2 | \leq m \cdot 2^{2N_a+2N_b}$. Therefore, it follows:

$$2m \cdot 2^{2N_a+2N_b} \geq p \tag{1}$$

which results in a contradiction as $p > \max\{2m \cdot 2^{2N_a+2N_b}, 2N_a + 4N_b + \log_2 m\}$
now suppose $r_1 + k_2 \neq r_2 + k_1$. We know $r_1 + k_2 = r_2 + k_1 \pmod p$. Hence, it follows that:

$$| (r_1 + k_2) - (r_2 + k_1) | \geq p$$

$$| (r_1 - r_2) - (k_1 - k_2) | \geq p$$

$$| r_1 - r_2 | + | k_1 - k_2 | \geq p$$

$| 2^{r_1} |\leq | \ell \cdot 2^{r_1} |\leq | a_1 | \leq 2^{2N_a+2N_b+\log_2 m}$. This implies $r_1 \leq 2N_a + 2N_b + \log_2 m$. Similarly, $r_2 \leq 2N_a + 2N_b + \log_2 m$. Hence, $| r_1 - r_2 | \leq 2N_a + 2N_b + \log_2 m$. Now, $b_1 = 2^{k_1} \leq 2^{2N_b}$(follows from the denominator bound in Claim $B.1$). This implies $k_1 \leq 2N_b$ and $k_2 \leq 2N_b, | k_1 - k_2 | \leq 2N_b$ then immediately follows. Using the above results:

$$2N_a + 2N_b + \log_2 m + 2N_b \geq p$$

$$2N_a + 4N_b + \log_2 m \geq p \tag{2}$$

The above leads to a contradiction as $p > \max\{2m \cdot 2^{2N_a+2N_b}, 2N_a + 4N_b + \log_2 m\}$, so this means θ is a 1-1 function.
Definition B.1, Claim B.3 and Claim B.4 are also applicable with the newer representation. Below are the arguments which justify this:
Definition B.1: The canonical form in the new field would be very smilar to the canonical form in $\mathbb{Q}/p$. An element $(a, b) \in \theta(U)$ is a canonical form of its equivalence class if $a \in [0, 2^{N_a}] \cup [p-2^{N_a}, p)$ and $b \in \{1, 2, 3, ..., N_b\}$. Let (e_a, e_b) be the canonical form of $e \in \theta(U)$, we define θ^{-1} as follows:

$$\theta^{-1} : \theta(U) \to U$$

$$e \mapsto \begin{cases} e_a/2^{e_b} & \text{if } 0 \leq e_a \leq 2^{N_a} \\ (e_a - p)/2^{e_b} & \text{if } p - 2^{N_a} \leq e_a < p \end{cases}$$

Claim B.3: θ^{-1} is well-defined
Proof: For $e \in \theta(U)$, let $e = (a, b) \sim (c, d)$, where (a, b) and (c, d) are both canonical forms. We wish to show that $\theta^{-1}((a, b)) = \theta^{-1}((c, d))$.
We have $\theta^{-1}((a, b)) \in U$ and $\theta^{-1}((c, d)) \in U$, by definition of θ^{-1} and U. Also we have $\theta(\theta^{-1}((a, b))) \sim (a, b)$, as follows. If $a \in [0, 2^{N_a}]$, then $\theta^{-1}((a, b)) = a/2^b$ and $\theta((a, 2^b)) = (a, b)$. If $a \in [p - 2^{N_a}, p)$, then $\theta^{-1}((a, b)) = (a-p)/2^b$ and $\theta((a-p)/2^b) = (a - p \, mod p, b) \sim (a, b)$. Likewise, $\theta(\theta^{-1}((c, d))) \sim (c, d)$. Now let $u_1 = \theta^{-1}((a, b))$ and $u_2 = \theta^{-1}((c, d))$. Assume for the sake of contradiction that $u_1 \neq u_2$; then $\theta((u_1)) \nsim \theta((u_2))$, by 1-1 property of the θ mapping (as proved above). Thus $(a, b) \sim \theta(\theta^{-1}((a, b))) \nsim \theta(\theta^{-1}((c, d))) \sim (c, d)$, a contradiction

Claim B.4: An element in $\theta(U)$ cannot have two canonical representations (a, b) and (c, d) with $a \in [0, 2^{N_a}]$ and $c \in [p - 2^{N_a}, p)$.
Proof: Take $(a, b) \sim (c, d)$ where $a \in [0, 2^{N_a}]$ and $c \in [p - 2^{N_a}, p)$ where $b, d \geq 0$. Because θ^{-1} is a function (claim B.3), so $\theta^{-1}((a, b)) = \theta^{-1}((c, d))$. However $\theta^{-1}((a, b)) = a/2^b \geq 0$ and

$\theta^{-1}((c,d)) = (c-p)/2^d < 0$, which is a contradiction.

Implementation details:
The computation in new field is isomorphic to computation in $\mathbb{Z}/p$, via the following map:

$$f : \mathbb{F} \to \mathbb{Z}/p$$

$$(a,b) \mapsto a2^{-b}$$

f preserves addition and multiplication rules:
$f((a,b)) \cdot f((c,d)) = a2^{-b} \cdot c2^{-d} = ac \cdot 2^{-b-d}$
$f((a,b) \cdot (c,d)) = f((ac, b+d)) = ac \cdot 2^{-b-d}$
Hence, $f((a,b) \cdot (c,d)) = f((a,b)) \cdot f((c,d))$
Similarly, for addition:
$f((a,b)) + f((c,d)) = a2^{-b} + c2^{-d}$
$f((a,b) + (c,d)) = f((a2^d + c2^b, b+d)) = (a2^d + c2^b) \cdot 2^{-b-d} = a2^{-b} + c2^{-d}$
Hence $f((a,b) + (c,d)) = f((a,b)) + f((c,d))$

Claim: f is a function from $\mathbb{F}$ to $\mathbb{Z}/p$
Proof: we need to prove that if $(a,b) \sim (c,d)$ then $f((a,b)) = f((c,d))$
assume that $(a,b) \sim (c,d)$, where $a = \ell \cdot 2^{r_1}$ and $c = s \cdot 2^{r_2}$ then:

$$\ell \equiv s \, mod \, p$$

$$r_1 + d \equiv r_2 + b \, mod \, p$$

now $\ell \equiv s \bmod p$ implies $\ell = s$, because $\ell \neq s$ means $\mid \ell - s \mid \geq p$ which as shown in (1) leads to a contradiction. So, $\ell = s$
Similarly, $r_1 + d \equiv r_2 + b \pmod p$ implies $r_1 + d = r_2 + b$, because $r_1 + d \neq r_2 + b$ means $\mid r_1 + d - (r_2 + b) \mid \geq p$ which as shown in (2) leads to a contradiction. So, $r_1 + d = r_2 + b$. Furthermore, $r_1 + d = r_2 + b$ implies $2^{r_1+d} = 2^{r_2+b}$
$\ell = s$ and $2^{r_1+d} = 2^{r_2+b}$, then leads to $\ell \cdot 2^{r_1+d} = s \cdot 2^{r_2+b}$, or $\ell \cdot 2^{r_1+d} = s \cdot 2^{r_2+b} \bmod p$.

$$\ell \cdot 2^{r_1+d} = s \cdot 2^{r_2+b} mod \, p$$

$$a2^d = c2^b mod \, p$$

$$a2^{-b} = c2^{-d} mod \, p$$

Hence, f((a,b))=f((c,d)).
Claim: f is 1-1 from $\mathbb{F}$ to $\mathbb{Z}/p$ if $p > 2m \cdot 2^{2N_a + 4N_b}$
Proof: we need to prove that if $f((a,b)) = f((c,d))$ then $(a,b) \sim (c,d)$

$$f((a,b)) = f((c,d))$$

$$a2^{-b} \equiv c2^{-d} \, mod \, p$$

$$a2^d \equiv c2^b \, mod \, p$$

$a2^d \equiv c2^b \, mod \, p$ implies $a2^d = c2^b$ because $a2^d \neq c2^b$ results in a contradiction as shown below:

$$\mid a2^d - c2^b \mid \geq p$$

$$\mid a2^d \mid + \mid c2^b \mid \geq p$$

$\mid a2^d \mid \leq \mid a \mid\mid 2^d \mid \leq m \cdot 2^{2N_a+4N_b} (as \mid a \mid \leq m \cdot 2^{2N_a+2N_b} and \mid 2^d \mid \leq 2^{2N_b})$. Similarly, $\mid c2^b \mid \leq \mid c \mid\mid 2^b \mid \leq m \cdot 2^{2N_a+4N_b}$. Hence, it follows:

$$\mid a2^d \mid + \mid c2^b \mid \leq 2m \cdot 2^{2N_a+4N_b}$$

$$2m \cdot 2^{2N_a+4N_b} \geq p$$

The above inequality results in a contradiction given $p > 2m \cdot 2^{2N_a+4N_b}$
writing $a = \ell \cdot 2^{r_1}$ and $c = s \cdot 2^{r_2}$ and substituting then gives:

$$\ell 2^{r_1+d} = s2^{r_2+b}$$

which implies $\ell = s$ and $r_1 + d = r_2 + b$, or $\ell \equiv s \bmod p$ and $r_1 + d \equiv r_2 + b \bmod p$. So, $(a, b) \sim (c, d)$. So, computation in the field $\mathbb{F}$ is isomorphic to computation in the field $\mathbb{Z}/p$.

Definition B.2(GINGER-Q protocol): Let Ψ be a computation over $\mathbb{F}$, and let Ψ' be the same computation, expressed over $\mathbb{Z}/p$. The GINGER-Q protocol for verifying Ψ is defined as follows:

1. $V \rightarrow P$: a vector x, over the domain $\mathbb{F}$
2. $P \rightarrow V$: $y = \Psi(x)$.
3. $P \rightarrow V$: x' and y'. P obtains x',y' (which are vectors in $\mathbb{Z}/p$) by applying f elementwise to x and y.
4. V checks that for all $(a, b) \in \{x \cup y\}$ and the corresponding element $c \in \{x' \cup y'\}$, $c2^b \equiv a \bmod p$. This confirms that P has applied f correctly. If the check fails, V rejects.
5. V engages P using the existing GINGER implementation, to verify that $y' = \Psi'(x')$.

Next I describe simple modifications to appendix C(C.3 and C.4).

Claim C.3 $\theta(x_1) - \theta(x_2) \in \theta(S)$ if and only if the numerator in the canonical representation of $\theta(x_1) - \theta(x_2)$ is contained in $[p - 2^{N_a}, p)$.

Proof: We will use the definition of θ^{-1} in the previous appendix. Let $e = \theta(x_1) - \theta(x_2)$. If $e \in \theta(S)$, then $\theta^{-1}(e) = a/2^b$, where $a \in [-2^{N_a}, 0)$ and $2^b \in \{1, 2, 2^2, \ldots, 2^{N_b}\}$. Thus, $\theta(a/2^b) = (p + a, b)$, where $p + a \in [p - 2^{N_a}, p)$, and $\theta(a/2^b) = \theta(\theta^{-1}(e)) \sim e$, so e has a canonical representation of the required form. On the other hand, if $e \sim (a, b)$, where $a \in [p - 2^{N_a}, p)$, then $\theta^{-1}(e) = (a - p)/2^b \in S$, so $e \sim \theta(\theta^{-1}(e)) \in \theta(S)$.

We instantiate step $C3$ with the following constraints $\mathcal{C}_<$:

$$\mathcal{C}_< = \left\{ \begin{array}{ll} A_0((1,0) - A_0) & = (0,1), \\ A_1((2,0) - A_1) & = (0,1), \\ \vdots & \vdots \\ A_{N_a-1}((2^{N_a-1},0) - A_{N_a-1}) & = (0,1), \\ A - (p - 2^{N_a}, 0) - \sum_{i=0}^{N_a-1} A_i & = (0,1), \\ B_0((1,0) - B_0) & = (0,1), \\ B_1((1,0) - B_1) & = (0,1), \\ \vdots & \vdots \\ B_{N_b}((1,0) - B_{N_b}) & = (0,1), \\ \sum_{i=0}^{N_b} B_i - (1,0) & = (0,1), \\ B - \sum_{i=0}^{N_b} B_i \cdot (1,i) & = (0,1), \\ \theta(X_1) - \theta(X_2) - A \cdot B & = (0,1) \end{array} \right\}$$

Lemma C.4 : $\mathcal{C}_<$ is satisfiable if and only if the numerator in the canonical representation of $\theta(x_1) - \theta(x_2)$ is contained in $[p - 2^{N_a}, p)$.

Proof: Assume that $X_3 = \theta(x_1) - \theta(x_2)$ has the required form (a, b). We have $b \in \{0, 1, 2, \ldots, N_b\}$ and $a \in [p - 2^{N_a}, p)$. Now, take $B_{\mathbf{b}} = (1, 0)$ and all other $B_j = (0, 1)$; this satisfies all of the B_i constraints, including $\sum_{i=0}^{N_b} B_i - (1, 0) = (0, 1)$, which requires that exactly one B_i be equal to $(1, 0)$. For B, take $B = (1, b)$, to satisfy $B - \sum_{i=0}^{N_b} B_i \cdot (1, i) = (0, 1)$.

Now, let $a' = a - (p - 2^{N_a})$. The binary representation of a' has bits $z_0, z_1, \ldots, z_{N_a - 1}$. Set $A_i = (z_i, 0)(2^i, 0)$ for $i \in \{0, 1, \ldots, N_a - 1\}$. This will satisfy all of the individual A_i constraints. And, since $\sum_{i=0}^{N_a - 1} A_i = (a', 0)$, we can take $A = (a, 0)$ to satisfy $A - (p - 2^{N_a}, 0) - \sum_{i=0}^{N_a - 1} A_i = (0, 1)$. The remaining constraint is the last one in the list. It is satisfiable because we took $B = (1, b)$ and $A = (a, 0)$, giving $X_3 - (a, 0) \cdot (1, b) = (0, 1)$.

For the other direction, if the constraints are satisfiable, then $X_3 = \theta(x_1) - \theta(x_2)$ can be written as $(a, 0)(1, b)$, where $b \in \{1, 2, \ldots, N_b\}$ and where $a = p - 2^{N_a} + \sum_{i=0}^{N_a - 1} z_i 2^i$, for $z_i \in \{0, 1\}$. This implies that $a \in [p - 2^{N_a}, p)$.